UNIVERSITY OF KWAZULU-NATAL
College of Agriculture, Engineering and Science

Liver Segmentation using 3D CT Scans

by

Anura Hiraman
211502750

Supervisors:
Prof. Serestina Viriri
Dr Mandlenkosi Gwetu

A dissertation submitted in fulfilment of the academic
requirements for the degree of Master of Science in Computer Science,

Durban, South Africa, 2018

Declaration of Authorship

I, Anura Hiraman, declare that this thesis titled, 'Liver Segmentation using 3D CT Scans' and the work presented in it are my own.

I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this thesis has previously been submitted for a degree or any qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly stated.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

_____

Anura Hiraman

Declaration - Supervisor

As the candidate's supervisor, I agree to the submission of this dissertation.

_____

Prof. Serestina Viriri

Declaration - Publications

DETAILS OF CONTRIBUTION TO PUBLICATIONS that form part of and/or include research presented in this dissertation

Publication 1:
Anura Hiraman, Serestina Viriri and Mandlenkosi Gwetu, Efficient 3D liver segmentation using enhanced image alignment, IEEE International Conference on Information Communications Technology and Society, pp.1-8, March 2018. DOI: 10.1109/ICTAS.2018.8368753.

Publication 2:
Anura Hiraman, Serestina Viriri and Mandlenkosi Gwetu, Efficient region of interest detection for liver segmentation using 3D CT scans, IEEE International Conference on Information Communications Technology and Society, Accepted November 2018.

Publication 3:
Anura Hiraman, Serestina Viriri and Mandlenkosi Gwetu, Liver segmentation using 3D CT scans, Image Analysis & Stereology Journal, (submitted).

_____

Anura Hiraman

Abstract

Deep learning algorithms have shown ground breaking performance in the field of radiology. One of its primary uses is finding a region of interest and object or lesion detection which is a key part of diagnosis. Segmentation of organs or substructures allows for quantitative analysis of the organ or substructure. This will be used for segmentation of lesions which plays an important role in diagnosis and prognosis of liver diseases or abnormalities. The organ of interest in this research is the liver. Segmentation of the liver from CT scans plays an important role in the study of the liver functions and can assist in the diagnosis of liver diseases. Liver segmentation aims to accurately detect and delineate the liver, separating it from surrounding organs and isolating it for intricate analysis.

A method for liver segmentation using 3D CT scans is proposed in this dissertation. The approach used for liver segmentation employs deep learning techniques. Convolutional neural networks (CNNs) are implemented to separate the liver from its background in the CT image. Firstly, a CNN is used to classify each slice of a 3D scan in order to remove slices that do not belong to the abdomen. This is done to obtain a region of interest that contains the liver that will be further processed during liver segmentation. Furthermore, the abdominal slices are processed by a CNN to segment the liver. The resulting segmented liver slices are then reassembled into a volume for post-processing which involves morphological operations. All 3D scan volumes used for experimental evaluation are taken from the Medical Image Computing and Computer Assisted Intervention (MICCAI) 2007 grand challenge datasets. The results obtained from evaluation indicate the effectiveness of the proposed liver segmentation method in accurately segmenting the liver from a 3D CT image.

Acknowledgements

I would like to thank my supervisors, Prof. Serestina Viriri and Dr Mandlenkosi Gwetu, for their invaluable guidance and feedback throughout this project. I would also like to thank my family for their undying support and motivation.

Contents

List of Tables

List of Figures

Chapter 1

General Introduction

1.1 Introduction

Computed Tomography (CT) was developed by a British engineer named Sir Godfrey Hounsfield and Dr. Alan Cormack [4], and has since had a profound effect on medicine. CT uses sophisticated X-ray technology to aid in the detection of a variety of diseases and conditions and is fast, painless, non-invasive and accurate [2]. CT scanning employs numerous X-ray beams and a set of electronic X-ray detectors which rotate around the patient, measuring the amount of radiation being absorbed throughout the body. A large volume of data is processed and creates 2D cross-sectional images of the body referred to as slices which is reassembled by computer software to produce a detailed multidimensional view of the body's anatomy.

Computed tomography was the first radiologic modality that required the use of computers to scan the internal organs of the body [12]. The use of computer-aided detection (CAD) technology has revolutionized radiology such that many medical procedures have become dependent on it. It is impossible for radiologists to analyse and evaluate voluminous amounts of image data generated from CT and other imaging technologies in a short space of time. Although the expertise of radiologists and trained physicians cannot be entirely replaced by CAD systems, their primary purpose is to expedite and provide validation for diagnosis and interpretation of medical images.

There exist at least four types of CAD approaches to ensure reliable and accurate diagnosis [19]. The first type is meant for the purpose of visualization, quantitative or interactive analysis of regions of interest and image enhancement. The second type assists in feature extraction of an object or region of interest for further analysis, and the third type is to detect or classify regions of interest. The fourth type is to approximate the anatomical or functional properties of regions of interest that can not explicitly be revealed by the medical images. Each of these types of CAD functionalities have been achieved by numerous methods and techniques such as

pattern recognition, mathematical modelling, data mining, biomechanics and deep learning.

Recently, deep learning has achieved state-of-the-art performance in medical image analysis [62]. Deep learning techniques have been applied to medical image analysis with the intention of letting computers learn the features that optimally represent the problem data at hand [68]. The most successful type of model for image analysis to date is convolutional neural networks (CNNs). The main applications that deep learning techniques are used for are classification, detection and segmentation [75]. Within classification, deep learning techniques have been used for image classification where an image is taken as input and a single output is produced stating which class the image belongs to. Another application is object or lesion classification. The use of deep learning for object, region and landmark localization in CAD, entails finding a region, object or lesion of interest, which forms a key part of diagnosis. Segmentation of an organ or its substructures allows for quantitative analysis which may lead to empirical medical findings with respect to the organ of interest. An example of this is the segmentation of lesions which play an important role in diagnosis and prognosis of diseases and abnormalities. Other applications are image registration, content-based image retrieval, image generation and enhancement as well as combining image data with reports.

The organ of interest in this research work is the liver. The focus is on segmenting the liver from abdominal CT scans. Segmentation of the liver from these CT scans plays an important role in the study of the liver functions and can assist in the diagnosis of liver diseases. Accurately detecting and delineating the liver separates the liver from its surrounding organs and isolates it for intricate analysis [43]. Segmenting the liver may reduce the computation time required in analysis as the liver only occupies a portion of the abdominal CT scan. Accurate liver segmentation ensures that the whole liver is analysed and the surrounding organs and tissues are eliminated from the region of interest.

Inspired by the use of deep learning and its success in organ segmentation, an automatic segmentation method using convolutional neural networks for classification and as well as feature extraction for liver segmentation is proposed. CNNs are used to classify each 2D slice into one of three categories; chest, abdomen and pelvis. The abdominal slices make up the liver egion of interest as these slices contain liver tissue. The slices that are classified as chest or pelvis slices are discarded. Furthermore, the slices of the region of interest are processed by a CNN to locate and segment the liver.

1.2   Motivation

The goal of this research work is to obtain a method of liver segmentation using CT scans by implementation of deep learning techniques. Solving the problem of segmentation is dependent on grouping pixels based on certain characteristics such

as colour, intensity or texture. Liver segmentation has its own challenges due to the characteristics of the liver. Despite many years of research and applications of new methods and techniques, liver segmentation remains a challenging task which implies that there is still room for improvement of existing models.

When the liver is affected by tumours or liver diseases, it is important for the infected areas of the liver to be identified timeously. Liver segmentation plays a vital role in detection and classification of tumours and liver lesions. Accurate segmentation can be used for the diagnosis of liver diseases which can have a huge impact on preoperative clinical treatments such as liver transplantation, resection and radiation therapy. Liver segmentation is also used for volume measurements which are used to determine abnormalities of the organ.

1.3 Problem Statement

Currently, it is standard clinical practice to manually segment the liver. However; manual liver segmentation relies heavily on user-interaction to perform segmentation. Manual segmentation is performed via the contouring of pixels along the boundary of the liver on sequential CT or MR slices. After the identification of the liver on each slice, post-processing software is used to generate a liver volume [43]. Manual segmentation is subjective with high intra- and inter-observer variability, poorly reproducible and time consuming. Variability may also be introduced by sharp liver boundaries, window level settings, and computer monitor settings [99]. This makes manual segmentation undesirable in busy clinical practices. The use of semi-automated and automated liver segmentation methods have become more appealing due to the undesirable limitations of manual segmentation. However, as long as there hasn't been a method proposed which can eliminate the limitations of current methods and perform liver segmentation with a 100% accuracy rate there is a need for research in liver segmentation.

Current methods and techniques that have been used for liver segmentation have performed well but still have some limitations. The most common limitations are over-segmentation or under-segmentation due to the occurrence of tumours or lesions close to the liver surface or blood vessels just below the surface which are sometimes not considered to be part of the liver which leads to the detection of incorrect liver boundaries [37, 53, 65, 73, 86]. Another limitation is that there are many methods that rely heavily on user-interaction [25, 32, 40, 83, 41, 87, 94]. The segmentation of the liver from its surrounding organs and tissue has proven to be a challenging task [52]. This is due to low contrast between the liver and surrounding tissues which make liver boundaries fuzzy and difficult to detect, heterogeneity of background structures and great differences in liver shape and appearance. Another factor that makes segmentation difficult is the occurrence of large tumours and other liver pathologies, such as intrahepatic veins, that usually lead to complicated intensity distributions and inhomogeneous appearances [52].

Regions with pathologies are different from healthy liver tissue which may cause under-segmentation or over-segmentation of these tissues.

1.4   Research Objectives

The primary objective of this research is to model a framework for accurate and efficient liver segmentation. This proposed framework uses deep learning techniques as well as some image enhancement to detect and segment the liver while attempting to preserve the boundaries of the liver. The overall goal of this dissertation is to ensure that the proposed method is of comparative or better segmentation ability to those already implemented. The specific objectives are as follows:

- To investigate different methods of liver segmentation currently used in literature and improve the accuracy rate

- To accurately segment the liver using 3D CT scan images

- To model an efficient liver segmentation framework using deep learning

1.5   Contributions of the Dissertation

The main contributions of this dissertation are the use of a novel CT slice alignment algorithm, the use of slice classification to obtain a region of interest for liver segmentation and a new method that includes concatenate layers among convolutional layers in a CNN for liver segmentation. These contributions are described as follows:

- Automatic segmentation of the liver using CT scans. This was achieved by a method that uses CNNs with concatenate layers among convolutional layers for liver segmentation using 3D CT scans. Two concatenate layers are added among the convolution layers in a CNN that performs liver segmentation. The effect of adding these concatenate layers on the segmentation results is analysed and compared to a CNN without concatenate layers. The CNN containing concatenate layers proved to produce better results.

- Model an efficient liver segmentation framework using deep learning. This was achieved by employing a new region of interest detection method for liver segmentation as well as the use of deep learning techniques to perform the liver location and segmentation.

- This methodology managed to improve the liver segmentation results of many methods in the related literature. This was done by the implementation of a novel method to obtain a region of interest for liver segmentation. This

is done by classification of slices to detect slices containing the pelvis and chest so that they can be removed while maintaining the abdomen within which the liver occurs. The effect of deeper networks, networks with many layers, are investigated to determine if they perform better than networks with fewer layers. It was evident that the networks with more convolutional layers performed more accurately.

- A new automatic method for the alignment of serially acquired 2D CT slices is presented. The contribution made by this approach is the use of skeletonization and a randomized alignment algorithm. The images are aligned with their neighbouring slices in a random fashion rather than sequentially.

1.6 Organisation of the Dissertation

The remainder of the dissertation is organized as follows:

- Chapter 2 - Background and Related Work: This chapter investigates the current methods of liver segmentation.

- Chapter 3 - Slice Alignment: This chapter details a 2D CT slice alignment method.

- Chapter 4 - Liver Region of Interest Detection: This chapter looks at locating the liver within the CT scan image volume.

- Chapter 5 - Liver Segmentation: This chapter details the liver segmentation process.

- Chapter 6 - Results and Discussion: This chapter covers the results of the methods implemented in the previous chapters as well as an overview of the dataset used for testing the methods.

- Chapter 7 - Conclusion and Future Work: This chapter concludes the dissertation and discusses future work.

1.7 Conclusion

In this chapter, a general introduction of the topic in this dissertation is described. The motivation, problem statement and main objectives of this research are provided as well as a brief description of the contributions made in this dissertation. The organisation of the dissertation is also detailed.

In the next chapter, the background of the liver anatomy and its appearance in CT scans is described and related literature is discussed.

Chapter 2

Background and Related Work

2.1   Introduction

This research work focuses on the localization and segmentation of the liver in abdominal CT scans using deep learning through the implementation of convolutional neural networks. Various methods have been proposed to achieve liver segmentation however; many of the existing techniques have some kind of limitation [70]. There are interactive methods, semi-automatic methods as well as automatic methods. Interactive methods are dependent on user input to perform liver segmentation whereas automatic methods can perform liver segmentation independently of user input.

Liver segmentation methods can be categorised as anti-learning based methods, classical learning-beased methods and deep learning methods. In the most recent work done in liver segmentation, the use of deep learning is very popular for the purpose of feature extraction [68].In order to gain a better understanding of how to choose an efficient technique for liver segmentation, it is important to have some knowledge about the liver anatomy and its appearance in CT scans. It is also important to conduct a study of related literature.

In this chapter, the anatomy of the liver and the abdominal CT scan are briefly introduced, related literature regarding image classification and the various methods of liver segmentation is discussed as well as a general summary of the drawbacks and limitations of current works.

2.2   Liver and CT

2.2.1   Liver Anatomy

The liver is one of the largest and a vital organ in the human body. It is shaped like a cone and is located in the upper right-hand portion of the abdominal cavity,

below the diaphragm, and above the stomach, right kidney and intestines. More than 500 bodily functions have been associated with the liver and some of them include regulation of most chemical levels in the blood, excretion of bile and filtering blood that comes from the digestive tract [9].

The liver consists of four distinct lobes; the left, right, caudate, and quadrate lobes. The left and right lobes are the largest lobes and the right lobe is about five to six times larger than the tapered left lobe. The small caudate lobe extends from the posterior side of the right lobe and wraps around the inferior vena cava while the small quadrate lobe is inferior to the caudate lobe and extends from the posterior side of the right lobe and wraps around the gallbladder [8]. The liver anatomy is depicted in Figure 2.1.

(a)

(b)
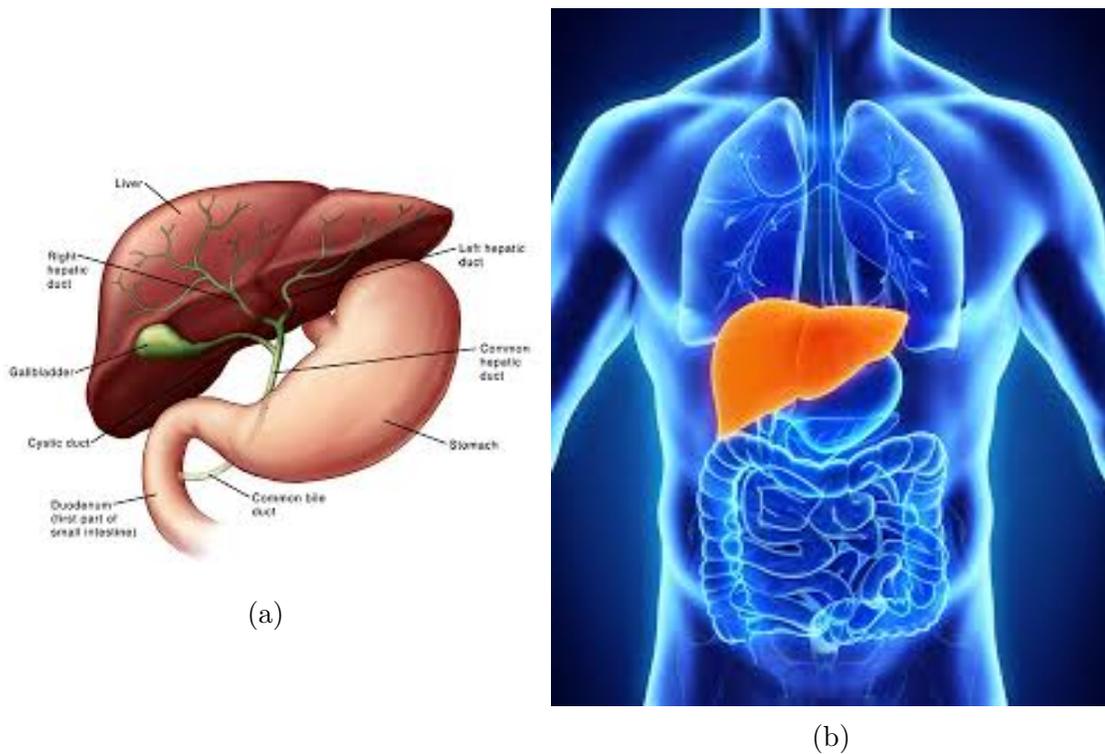
Figure 2.1: Human Liver Anatomy

2.2.2   Liver in Abdominal CT image

3D imaging techniques have made a generous contribution in medicine in areas such as visualization, analysis and diagnostics. Computed Tomography (CT) scans are commonly used for scanning large areas of the body. In this dissertation the type of CT used are abdominal CT scans which are used to view the anatomical structures within the abdominal cavity [1].

A beam of X-rays is aimed and rotated around the body to generate individual slices called tomographic images. These tomographic images contain more detailed

information than conventional X-rays because they gather detail from multiple angles [7]. During the scan different body parts absorb the X-rays in varying degrees and it is this crucial difference in absorption that allows the body parts to be distinguished from one another on an X-ray film or CT electronic image [3].

A series of slices are produced and these can be viewed in two ways, either as individual slices or they can be stacked to produce an image volume. The abdominal CT scan can be viewed from the axial view, sagittal view or coronal view. The slices in this study are from the axial view. The different CT views are depicted in Figure 2.2.

(a) CT views

(b) Sagittal view

(c) Axial view

(d) Coronal view
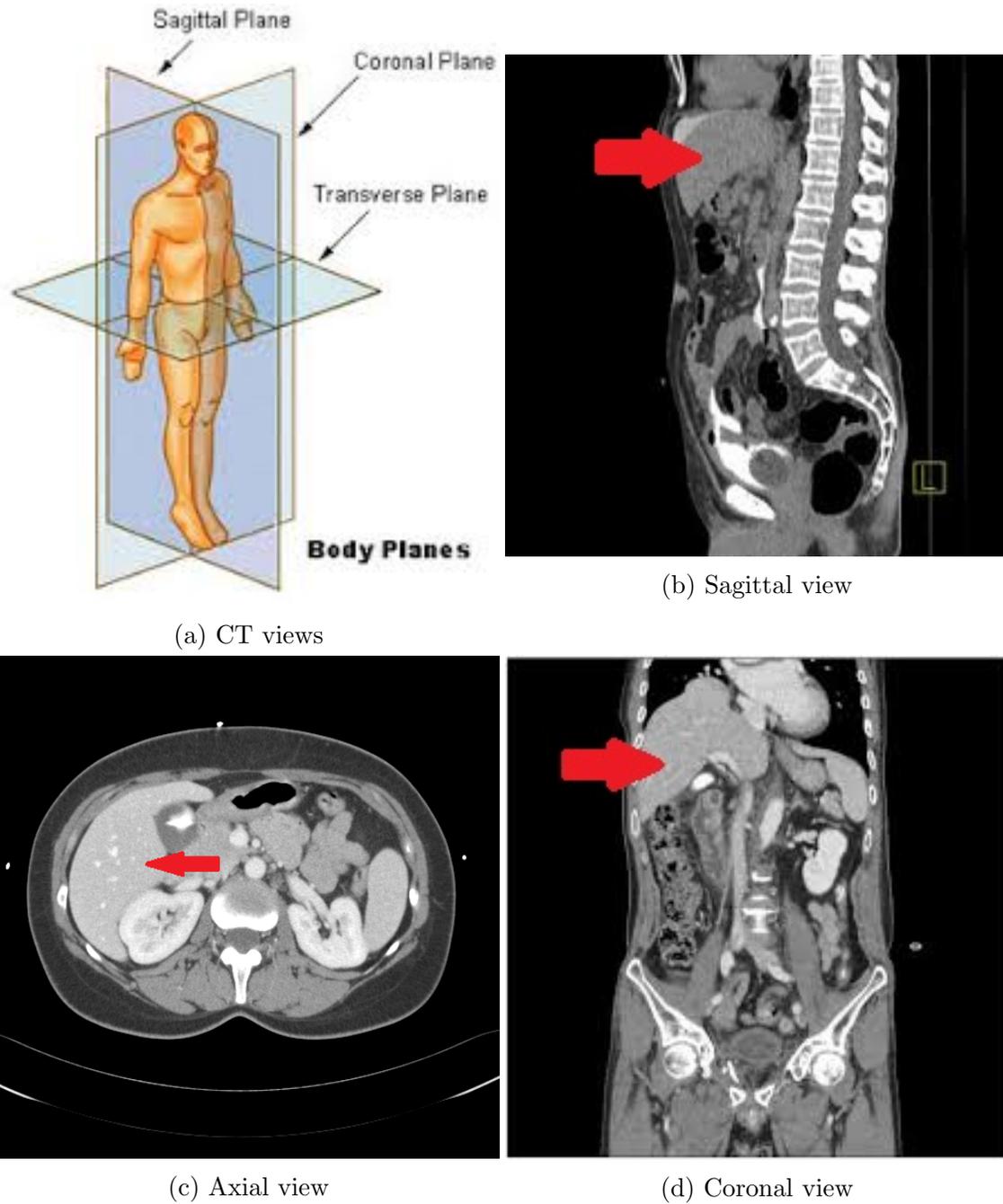
Figure 2.2: Liver in abdominal CT scan

2.3 Related Work

Over the past few decades, a variety of approaches have been proposed for liver segmentation. These approaches can be categorized as semi-automatic, automatic and interactive approaches. Automatic approaches are more desirable as they require no user-interaction but semi-automatic and interactive approaches remain signif-

icant due to the ambiguity of liver segmentation [68]. This section discusses the method used to conduct the literature review and the related works done in slice alignment, liver region of interest detection and liver segmentation.

2.3.1 Overview of Literature Review Method

A systematic literature review (SLR) was the chosen method to conduct this research. The first stage of the SLR is defining the research objectives in Section 1.4 of Chapter 1. The second stage is specifying the methodology and defining eligible criteria. The inclusion and exclusion criteria are shown in Table 2.1. The inclusion criteria include studies done in English, liver segmentation methods, liver segmentation using deep learning techniques and only liver segmentation from the 3D medical imaging modalities. Exclusion criteria are features and factors that do not qualify for inclusion within the study. The third stage of the SLR is retrieving eligible literature. This was done using search terms such as 'Liver segmentation from CT scans', 'Liver segmentation using image processing techniques' and 'Liver segmentation using deep learning techniques'. Articles were rejected if it was determined from the title and abstract that it failed to meet the inclusion criteria. The fourth stage involved assessing the quality of the chosen literature. This is important because it shows that there are no biases within the reports of the published studies and that the data extracted is of good quality. The last two stages of the SLR includes identifying the important information, analyzing the literature, performing a critical analysis and portraying the findings. This is exhibited within this section.

Table 2.1: Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Liver segmentation methods	Segmentation methods of other organs
Liver segmentation using deep learning techniques	Segmentation of other organs using deep learning techniques
Liver segmentation from 3D medical imaging modalities	Liver segmentation from 2D medical imaging modalities
Studies done in English	

2.3.2 Slice Alignment

A fully automated algorithm for the alignment of 2D serially acquired sections forming a 3D volume was presented by Krinidis et al [60]. This approach depends on the optimization of a global energy function that is based on the shape of the object. The global energy function measures the similarity between a slice and its neighbourhood within the 3D volume. The energy function used is associated with a pixel similarity metric based on the Euclidean distance transform. The

energy function formulated indicates that for a slice that is to be aligned, the similarity between the transformed slice and all other slices that have already been transformed is accumulated in the energy function. The slices are considered per triplets and are aligned with respect to other slices in its neighbourhood. Slices are chosen at random to be aligned. The Iterated Conditional Modes (ICM) algorithm, a deterministic optimization algorithm, is used to conduct the minimization of the local energy function. The algorithm converges toward a local minimum of the local energy function which corresponds to the local minimum of the global energy function. The solution obtained is improved further by a gradient descent technique and multigrid data processing is also implemented to improve the speed of the algorithm. The method was evaluated by applying the algorithm to the reconstruction of five CT volumes. The alignment errors obtained for one of the volumes were a mean of 0.37 for the horizontal translation, 0.37 for the vertical translation and 0.19 for the rotation. It was concluded that this method is efficient and achieved alignment with high accuracy. This method aligns the slicing by minimizing the distance between the contours of the images. However, if the inter-slice distance is large such that the contours cannot be matched from one slice to another or if contours are not clear and distinct, the method might not be able to align the images accurately.

Fei et al [33] created and evaluated a slice-to-volume registration algorithm and its potential application to interventional MRI-guided radio-frequency thermal ablation of prostate cancer. The presented algorithm included a multiresolutional approach, two similarity measures and automatic restarting to avoid local minima. The two similarity measures used are mutual information for higher resolutions and correlation coefficient for lower resolutions. The restarting feature restarts with randomly perturbed parameters. A rigid-body transformation and trilinear interpolation is used for registration and the downhill simplex method of Nelder and Mead or the Powell method was used for optimization. More than 800 experiments were carried out and a 100% success rate was achieved for transverse image slices. However, errors occurred when the image slices were obtained at other orientations and positions mostly because of inconsistent image content.

Collins et al [23] presented a non-linear cerebral registration technique with sulcal constraint. Here, it was demonstrated that cortical registration could be improved by combining linear registration and resampling into a stereotaxic space, cortical surface extraction, tissue classifcation, automatic sulcal extraction and non-linear registration. The registration of two volumes consists of two parts. First, the linear part of the transformation determines the correlation between Gaussian-blurred features extracted from the two volumes, the source and target volumes. Thereafter, a 3D deformation field is estimated by sequentially stepping through the target volume in a 3D grid pattern. At each node in the grid, the deformation vector required to achieve local registration is found by optimization of three translational parameters that maximize the objective function evalutaed only in the neighbourhood of the node. The algorithm is applied iteratively reducing the grid size and blurring the image after each iteration, thus refining the fit. This method

reduced the sulcal mis-registration by 20%.

Murgasova et al [78] proposed a novel method for correction of geometric distortion induced by main static magnetic (B0) field in the fetal Echo planar imaging (EPI). The method consists of two registration processes. The first is a rigid registration of EPI slices with magnetic resonance imaging (MRI) volume to correct for motion and the second is a deformable registration with Laplacian constraint to correct for geometric distortion of the EPI slices. The optimization of the objective function consists of two steps. The first step is a rigid registration of the acquired slices to a model locating the fetal head to estimate the motion parameters. The second is a deformable registration of acquired slices to the simulated slices regularized by the Laplacian constraint to estimate a smooth distortion field. The motion parameters are estimated by rigid registration and distortion by B-spline registration. It was concluded from the results obtained that this method is a viable alternative to correction using acquired B0 field maps, especially if the B0 field map is corrupted by fetal or maternal motion. This registration method is specific to fetal EPI and cannot be applied to other medical images.

Fischer et al [35] proposed a new curvature-based registration method. The implementation of this technique is based on a time marching algorithm where a time step is introduced. A stencil is generated using convolution with respect to Neumann boundary conditions. The function is solved using linear reduction which is then realised by a discrete cosine transform resulting in fast implementation. The method was evaluated on clinical 2D MR images of a female breast. The convergenec behaviour is smooth and the MSD was improved by 40%.

Brown and Bradshaw [16] implemented an improved fingercode alignment for accurate and compact fingerprint recognition which performed better than traditional texture-based fingerprint recognition. The novelty of the proposed system is the combination of the ECC algorithm with initial minutiae triplet alignment. Here, the use of the ECC alignment algorithm produced promising results as it offers a reliable solution to the texture alignment problem. Evangelidis et al [30] also presented a method using the ECC function as a performance criterion in the projective image registration problem. They investigated the use of the ECC for the eight-parameters nonlinear projective registration problem where the results achieved were superior to that of the Lucas-Kanade algorithm. In both these cases the ECC alignment algorithm is not used for the alignment of a series of medical images.

2.3.3   Liver Region of Interest Detection

A common technique used to extract a region of interest is using a bounding box. Peng et al [81] used a 3D bounding box for the external body which was extracted with thresholding and then cropped out of the image to isolate the body for liver segmentation. Campadelli et al [18] also used a body box extraction for liver segmentation using thresholding where the threshold value is the grey levels obtained

from histogram analysis. This was done to extract the patients body from the CT scan to ensure that any external artifacts do not affect further processing. Lu et al [70] also implements a method to localize the liver by deleting slices within the volume that do not contain liver tissue. However, this method is undesirable as it was manually done.

Recently, one of the most common techniques of choice is deep learning. Convolutional neural networks have been used to automatically divide CT scans to localize a search space for radiologists to work with. Sugimori [95] conducted a study aimed at classifying CT images in different slice positions using deep learning. CT images were divided into ten classes for brain, neck, chest, abdomen and pelvis with contrast-enhanced and plain imaging. The AlexNet and GoogLeNet architectures were trained and evaluated for the image classification. However, these architectures were developed for classification of natural images and not medical images.

Convolutional neural networks are also used for classification of medical images for other purposes such as detection or localization of abnormalities. Heeneman [46] proposed a method that uses a 2D convolutional neural network with a sliding window approach to localize lung nodules. After the lungs are segmented, a sliding window of 64 by 64 pixels is used with an overlay of 24 pixels. The 2D convolutional neural network classifier receives the 64 by 64 window as input and performs binary classification predicting a positive or negative classification of the window. This produces a final prediction and localization of the lung nodules.

De Vos et al [26] proposed a method for automatic localization of one or more anatomical structures in 3D medical images through detecting their presence in 2D image slices using a convolutional neural network. The convolutional neural network is used to detect the presence of the anatomical structures of interest in axial, coronal and sagittal slices extracted from the 3D medical image. Thereafter, the outputs of the neural network are combined and 3D bounding boxes are created. In the work presented in this paper, the convolutional neural network is trained to detect anatomical structures in order to classify 2D slices and remove the irrelevant slices to obtain a region of interest suitable for liver segmentation.

2.3.4   Liver Segmentation Methods

Pre-processing Techniques

The purpose of pre-processing techniques is to process the images in preparation for the next stage of image processing. CT images are not always of the same quality or may have artifacts or noise that may degrade the performance of the liver segmentation techniques due to the various CT scanners used in radiology [52].

Several methods begin with contrast enhancement using histogram equalization and normalization [108, 54, 37, 65, 21], to increase the contrast in the image making

edges more visible as well as a median filter to smooth the image [53, 105, 25, 108, 65]. Huang et al [54] implemented a noise suppression procedure which is based on Block Matching 3D Sparse Transform Domain Collaborative Filtering. A recursive Gaussian filter was used by [106] for smoothing. Another common pre-processing technique used was the non-linear mean shift filter [73, 87] and the anisotropic diffusion filter [83, 106, 21].

Since the liver is the organ of interest, many methods begin with attempting to isolate the liver tissue and a common way this was done is by a combination of histogram analysis to determine the liver intensity and thresholding [66, 108]. Christ et al [21] and Thong et al [98] windowed the CT scan between -150 to 250 Hounsfield units to isolate the body from regions composed of air and to exclude bone regions. This allows the next stage to focus on soft tissues and organs only. Another technique used by Christ et al [21] is the removal of non-liver slices and resampling the volume. A 3D bounding box was also used to isolate the body with thresholding [81, 18]to ensure that any external artifacts do not affect further processing.

Resampling and image registration are also used as a pre-processing technique for the purpose of re-aligning image data and reducing or eliminating deformations. Affine registration and resampling towards reference images of training sets were used in [88, 103]. Resampling by linear interpolation [101, 37] and bicubic interpolation were also used [25].

Learning based and deep learning methods require training of the segmentation model that will be used for the liver segmentation. Data augmentation is conducted for the training process. Christ et al. [21] uses data augmentation such as elastic deformation, translation and rotation. Using this technique allows learning models to be able to detect and segment the liver in images that are slightly rotated, translated or deformed which increases the potential of correct detection and segmentation.

The preprocessing techniques used in the proposed method is windowing the CT scan between -150 to 250 Hounsfield units to isolate the body from regions composed of air and to exclude bone regions and data augmentation to produce more training data for the training of the neural networks.

Anti-learning Based Segmentation

Massoptier and Casciaro[73] proposed an approach for automatic segmentation of liver complex geometries that focused on the automation of the graph cut technique and its application to liver segmentation. The results were compared to those obtained using the gradient-vector flow (GVF) snakes technique. First, a fully automatic method was proposed based on a hybrid approach consisting of the adaptive thresholding technique and GVF active contours were used to obtain an accurate liver surface which enabled the extraction of boundaries close to real ones in a relatively short time. The results obtained were good in terms of quality

and processing time of segmentation however, there was some limitation in some cases where a tumour is located close to the liver surface or vessels are located just under the liver surface. In these cases, the algorithm considers the tumour or blood vessels to be outside the liver and the wrong boundary is detected. Then, a new method was proposed using the graph-cut technique which was developed with the intention of eliminating those limitations encountered in their first method. An adaptive thresholding technique is applied to obtain samples of the targeted object and of the background which is used to initialize the graph-cut. The graph-cut technique is applied to segment the liver from the surrounding background using an energy function. The work based on graph-cut was compared to that of the results obtained by the first method based on the GVF snakes. The results achieved by the graph-cut technique was better as the dice similarity coefficient (DSC) increased to 0.95 from 0.93 and both the false negative ratio and the false positive ratio averages decreased under 5%. It was evident that the graph-cut technique is a more robust segmentation technique. The technique does not fail when tumours are near the liver surface.

Sangewar et al [87] proposed a novel variational model based on intensity probability distribution propagation and region appearances propagation which is used to focus on the target liver regardless of background complexity. The segmentation is based on combining a modified k-means method with a localized contouring algorithm. The first part of the segmentation is the k-means clustering which requires the user to manually pick five points of varied intensity in the scan. The points are used as centroids and the distance objects to centroids are calculated. The pixels in the image are then grouped based on minimum distances into five groups. This process is iterated through a number of times allocating all pixels to a group. A logical intersection of three independent segmentation runs were carried out in order to obtain results free of misclassifications. Thereafter, a region growing technique is applied in order to grow regions from the selected seed points to adjacent points depending on a specified criterion. Lastly, an active contours model is used to detect the edges of the liver and points are plotted for the volume calculation of the liver. It was concluded that the proposed method had the potential to provide fast and accurate liver segmentation and rendering and delineating tumour regions. However, this model requires user interaction to accurately perform segmentation.

Zayane et al [108] proposed an automatic segmentation method based on a priori knowledge of the image such as location and shape of the liver. The image was first thresholded and thereafter, the grayscale consistency constraint (GCC) algorithm is developed to detect the largest connected component which in this case is the liver. The GCC algorithm is composed of two modules; the first for labelling and the second for finding that largest component. A morphological closing filter is applied followed by the application of the nonlinear Sobel filter to detect contours of the liver. This results in the segmented image. The results obtained from this experiment were compared to that of manual segmentation of an expert. It was evident that the proposed methodology was acceptable but not one hundred percent reliable. The liver was not segmented properly in some cases because the natures

of the histogram of the images are completely different.

Gupta et al [93] presented a new classification method using fusion of the best features from different domains to classify liver in Ultra Sound images determining whether the liver contains fatty liver tissue. To achieve this goal, information related to the liver surface was extracted using various texture models and the best features were selected and combined to produce a single quantitative metric which is used to classify the liver. The first stage is pre-processing during which seven different texture models are used to extract features from the region of interest. The region of interest is selected by an expert radiologist. Using the texture models, 35 features are extracted from a training set. Thereafter, outlier values are removed from the feature dataset. The next stage is feature reduction. Linear discriminative analysis, in this case Fishers Discrimination Ratio (FDR) and Pearsons Correlation Coefficient (PCC), are used to select the best features. This process reduces the features to seven best features which forms the proposed classifier. Then the modified z-scores are calculated and the FDRs and their weightage are normalized. The information is fused into a linear model and then a single quantitative metric is calculated called the Discriminative Index. Thresholding is used to detect fatty liver tissue and normal liver tissue. The results obtained from the experiment achieved an accuracy rate of 95% at 100% sensitivity which makes this method comparable to existing classifiers. It is evident that this method is simple and has a high accuracy rate.

Evans et al [32] presented a method using a topology adaptive active contour model or snake to accurately segment the liver outline from 2D CT scans. The active contour model that was developed was a modified version of the T-snake where internal energy forces act as a smoothness constraint and an inflation force is employed to push contours towards edges within the image. As the contour moves it is reparameterized at regular intervals to a superimposed grid on the image. At each reparameterized step the set of nodes are replaced with a new node at each point where the contour intersects with the grid. Reparameterization overcame problems associated with inflationary contours allowing contours to flow into the complex shape of the liver. The advantages of this approach is that firstly, the interior of the healthy liver is more uniform than the exterior therefore there are less noisy edges that can trap nodes as they move outwards toward the edge of the liver and secondly, the snake can be initialised at any point within the liver. This avoids the issue of correctly initializing the contour which resulted in a faster and more robust segmentation. The basic snake equation is used with standard Gaussian gradient function being implemented as the external energy function and an inflationary constant to push the contour towards the edges. The T-snake function is also modified in that it has two threshold levels, an upper and lower threshold. This prevents the snake from moving into regions of other organs. The reparameterization of the snake is dependent on the changing resolution of the grid. The grid changes depending on the curvature of the snake at each node. This allows segmentation of complex and highly irregular areas of the shape of the liver. This methodology requires the user to select a seed point within the liver to initiate

the segmentation. In some cases, where a lobe of liver appears separate from the main liver body the user has to select a second seed point for segmentation of the separate lobe. Thereafter, the results are added to the main segmentation. This technique produced very encouraging results however; the parameters have to be set by the user.

Bhatt et al [40] proposed a new 3D liver segmentation method based on coupling a modified k-means segmentation method with a special localized contouring algorithm. The algorithm semi-automatically segments the liver region from the 3D CT scans and then extracts the liver region and renders the segmented liver for 3D viewing. The initialization process of the algorithm requires a user to manually select five points of varied intensity in the scan as well as a rough outline of the liver in widely spaced slices throughout the given dataset. The method consists of three parts. During the first part, the k-means clustering algorithm is applied to partition the dataset into various clusters whose means are similar. This produces five regions being identified which are the liver, surrounding organs, peripheral muscles, ribs/spinal cord and outside of the body. The points selected by the user act as seeds for the five masks. The liver mask and the mask with surrounding organs are logically ORed together to obtain a final mask. The second part of the method involves the initialization of the Contouring-Based Segmentation. This algorithm takes a user defined mask as input. The user selects the approximate boundary of the liver in the slices. The image is divided into blocks and then marked as part of the liver if at least half the pixels were found in the abdominal CT window. The first slice is always selected for manual contour initialization. Thereafter, the slices have to show a 5% change in the marked blocks to be selected for manual contour initialization. Slices not selected used the same initial contour as the slice closest to it. The third part of the method is the Localized-Region-Based Active Contouring Algorithm. The smoothened Heaviside function and the Chan-Vese energy paradigm is used to model the interior and exterior of the contour for segmentation. The localized versions of the mean intensities used to model the contour are obtained by restricting the field of view to the masked region produced in the first part of the method. The results were an average error was 1.73% and a high mean accuracy rate of 98.27%. However, this methodology relies heavily on user-interaction.

Li et al [65] also developed a fully automatic segmentation method using Fuzzy C-Means (FCM) clustering and level set method. Segmentation consists of FCM clustering and liver region selection, cluster merging, statistical rejection and applying the level set method to refine the segmentation result. During segmentation the image is clustered into groups based on predetermined equations. The resulting clusters are grouped into the liver, background, other tissue brighter than liver tissue such as bones and tissues darker than liver such as muscles. The CT slice is divided into four equal areas and the area that contains the liver is located. The selection of the liver group is done based on prior knowledge such as the liver's location and size. The liver may be clustered into several groups therefore similar groups are merged by comparing the difference between the average gray-level intensity of each group. The average intensity and standard deviation of the liver

group is used to reject tissues that don't belong to the liver. The largest connected region is then used to initialize the level set. The region is extended outwards by five pixels and then the level set is used on the image for forty iterations. There were three main problems produced in the output of the level set method; there may be tissue other than liver tissue, the contour is not smooth where the boundary isn't clear and holes appear in the liver as a result of contrast enhancement and statistical rejection. The application of morphological operators reject tissue around the liver. Thereafter, the largest connected region is selected to extract the liver and finally, the median filter is used to smooth the boundary. The experiment results obtained had a high accuracy rate of an average of 99.86% and specificity of 99.89%. When compared with the standard level set method this method is fully automatic and achieved better segmentation results. However, when there are blood vessels or other homogenous tissue on the edge of the liver under-segmentation occurs. This showed as indentations on the edge of the liver which results in a boundary that is not clear.

Rusko et al [86] presented a method based on intensity analysis and region growing where the issue of under-segmentation or over-segmentation is handled. The method consists of four steps. The first step includes determining the seed region. During this step the histogram is calculated for the right side of the body as this is where the liver is located in order to determine the mean intensity of the liver voxels. The histogram is smoothed and two significant maxima are produced. One belongs to muscles and the other to the liver. Since the intensity of the liver is higher than that of the muscles, the maximum represents the modus of the liver intensities. The minimal and maximal intensity of the liver voxels is then calculated from the histogram which enables the automatic location of the region inside the liver. From this a binary image is calculated. The image of possible liver regions is eroded so that small regions are deleted. The largest connected region is considered as the seed region for segmentation. The second step is the liver and heart separation. First the lungs are segmented based on the characteristic intensity of air. In order to do this, the largest connected air region for the left and right are determined separately. Using these regions as seeds the region growing method is used to segment the left and right lung lobes. Thereafter, two curves are determined from the bottom contour of the right and left lobes. The surface separating the liver and heart is calculated by averaging the curves located in neighbouring slices providing a smoother resolution. Then the voxels located above the surface are set to an artificial intensity value to prevent region-growing from going into this region. The next step is region-growing. This step includes calculating a histogram of the image from which the intensity range of the liver is determined. During the region-growing method a voxel is added to the region if all voxels in its neighbourhood have acceptable intensity. This modification to the region-growing method assists in the prevention of over-segmentation. The final step is post-processing during which problems associated with under-segmentation are dealt with. The MICCAI-SLiver07 dataset was used to evaluate this method. In some cases, the region of the liver near the right lung lobe is under-segmented due to the region having

lower intensities. A closed connected region between the liver and right lung lobe is defined and a new intensity interval is calculated based on this region which is used by another region-growing method. In some cases, the region-growing method leaks through the inferior vena cava due to it having a similar intensity to the liver tissue. In order to remove the inferior vena cava from the segmentation result, first the Hough transform is used to detect the cylindrical shape of the vessel. Then the liver region is thresholded to find all the voxels that are not part of the liver tissue and these voxels are erased. Sometimes the segmentation doesnt include the portal vein which is considered to be part of the liver is it is completely surrounded by liver tissue. The portal vein doesn't get included as a result of its high intensity. In order to fill this vessel, the contour of the liver is determined and the surface normal for each vessel is calculated. The surface voxel is marked if another liver voxel is in the direction of the surface normal. Finally, the liver is dilated at each marked surface voxel. It was evident in the experiment results that the presented method performs well only if the liver does not have very large lesions. In some cases, the liver was over-segmented.

A fully automatic segmentation method was proposed by Arzhaeva et al [101]. The method consists of a voxel labelling technique where the probability of each voxel is part of the liver is determined using a statistical classifier and a set of features. The core of this method consists of a voxel labelling procedure. For each voxel in the test set a feature vector is computed and a statistical classifier is trained with examples of feature vectors and their corresponding truth maps the vector to a probability that the voxel is part of the liver. The methodology consists of four stages which are pre-processing, registration, voxel classification and post-processing. First, the data was resampled in the vertical direction to eliminate the variation in the slice thickness. In some cases, this resulted in more blurring. Liver scan range determination was done using connected component analysis. A range was selected by inspection of the vertical range of the liver in relation to the lung based on the training data. Next, maximum intensity projection along the vertical direction and thresholding was used to determine the orientation of the scan and the image is rotated as needed. The registration stage of the methodology involves the registration of training scans with a volumetric registration package. It treats the registration as an optimization problem in which the similarity between the target and atlas image is maximized. Voxel classification includes computation and selection of features and classification of voxels using a k-NN classifier. Voxel classification is only applied to voxels that lie in a box that is determined from the bounding box of the atlas segmentation. The dataset used here was the MICCAI-Sliver07 dataset. The method produced results showing that performs quite robust however; two types of errors were made. The first is that occasionally a completely incorrect border is found and the second is the imprecise localization of the border in some cases.

Qi et al [83] developed an automatic segmentation method for the segmentation of liver tumours from CT scans. The method uses Bayesian Rule-Based 3D Region growing to achieve segmentation. This algorithm was proposed to cope with the

variance of tumours. The region growing algorithm is used to classify voxels into two classes: liver and lesion. The liver class is modelled by a single Gaussian and the tumour class is modelled by multiple Gaussians. The algorithm begins with the user selecting several seeds where one seed should be located inside the liver and at least one seed should be located inside a lesion. The seeds are used to set up initial configurations. It is assumed that the neighbourhood of each seed is inside the liver or the lesion. First, the parameters of the Gaussian model are estimated using the neighbouring voxels of the liver seed point as samples. Second, for each seed point in the lesion, Gaussian models are also estimated and compared with the single Gaussian. All the similar Gaussians of the lesion are combined by calculating new parameter values. The region growing algorithm begins from the seed points. It adds neighbouring voxels into regions if they are similar based on two criteria. The first criterion is that the intensity of the voxel should be accepted by one of the Gaussians of the lesion. The second criterion is that one of the intensity distribution models of lesions should be similar to that of the current voxel. The intensity distribution is determined by the intensity histogram within the 3D neighbourhood of the point. If the voxel meets both criteria then it is considered as part of a lesion and added into the region. Their intensity voxels are also used to update the parameters of the corresponding Gaussian model. The algorithm examines the connected voxels until the region fixes. By inspection of the results obtained in the implementation of this method, it was evident that the algorithm provided reasonable results for most of the lesions. However, the region growing process was not easy to control in 3D when there were unclear boundaries resulting in leakage which propagated to adjacent slices.

Peng et al [81] proposed a novel multi-region appearance based approach with graph cuts to delineate the liver surface. This approach was proposed to handle livers with multiple sub-regions. A geodesic distance based appearance selection scheme is introduced to utilize proper appearance constraint for each sub-region. The segmentation is modelled with energy functions which incorporate region and boundary information and an adaptive balancing weight is introduced and learned from training sets. Two models are proposed: Single region-appearance based model (GC-SRAP) and Multiregion-appearance based model (GC-MRAP). The GC-SRAP segmentation method requires at least initial region inside the liver. Three features are used for appearance description which is intensity, LBP and local variance. The single region-appearance based model is extended into a discrete energy optimized by graph cuts algorithm. The GC-MRAP segmentation method requires initializations on both the healthy liver tissue and tumour. A geodesic distance based selection scheme is introduced which employs location and appearance information from initializations. Initialization requires the user to select initial regions in cylindrical shape. In the cases where tumours are present, the user is required to select initial regions in these tumours as well. When the initialization contains two class regions the GC-MRAP model is used, otherwise the GC-SRAP model is employed. The segmentation module consists of the minimization of the proposed model being casted as a graph-cut problem and optimization using the

min-cut max-flow graph algorithm. The dataset used in this work was the MICCAI-Sliver07 dataset. The results achieved by the proposed algorithm agreed with that of the reference segmentation. This method achieved an average of 4.53% for VOE, 0.15 for RVD, 0.79mm for ASD, 1.49mm for RMSD, 16.75mm for MSD and 97.7% for DSC. These results can be compared with those of state-of-the-art methods. Quantitative validation has shown that the proposed method can delineate the liver surface accurately and efficiently.

A novel method was proposed by Wu et al [106] for automatic delineation of liver on CT volume images using supervoxel-based graph cuts. The segmentation consists of three steps: liver volume of interest (VOI) extraction and largest liver slice selection, Simple Linear Iterative Clustering (SLIC) supervoxel partition on liver VOI and graph cuts segmentation. The abdominal region is first extracted to remove several non-abdominal slices and voxels. A 3D abdominal bounding box was measured using maximum integral projection (MIP) and thresholding. Three sets of coordinates needed to be calculated which are the lower bounding, upper bounding and in-plane bounding coordinates. Coronal and axial MIP images are obtained by applying the MIP algorithm. The segmentation of bones was performed on the coronal MIP image by use of the Otsu algorithm to obtain the binary bone mask. The mask was further processed with morphological operators and median filtering. Analysis was done to determine the maximum number of connected bone pixels and locate the lower bounding coordinate. The axial MIP image was used to segment the abdomen by applying the Otsu algorithm and a series of thresholding. After applying morphological operators and selecting the largest connected component a mask of the abdomen was obtained. From this mask the in-plane bounding coordinates were obtained and a 2D bounding box was determined. Thereafter, a binary lung mask was obtained as a result of thresholding based on the fact that the air in lungs has very low intensity values. The maximum number of lung pixels in the axial slice was used to determine the upper bounding coordinate. A 3D bounding box is created and based on this the image is cropped to obtain the region of abdomen. The liver VOI is then extracted from the region of abdomen by applying a histogram based adaptive thresholding method and morphological operations. The liver intensity range is calculated from the histogram for adaptive thresholding. By applying binary thresholding the initial binary liver mask was obtained. The largest connected component on the right side of the abdomen was selected on each axial slice and morphological opening and largest connected component was applied to each coronal slice to obtain binary masks of the liver. The largest liver slice corresponding to the largest cross section of liver contains the liver as a whole object which is where seed points are extracted from. The SLIC is a k-means clustering algorithm that was used to generate supervoxels that have approximately uniform size while their boundaries match closely to the true image boundaries. The intensity of each supervoxel was computed as the average intensity of the voxels in the supervoxel. Graph cuts segmentation is then applied. It achieves an optimal solution by minimizing the energy function using the max-flow min-cut algorithm. Hard constraints or seed points in this case and intensity distributions are needed

for the segmentation. Foreground and background seed points were automatically sampled from regions in the largest liver slice corresponding to the liver and background binary masks respectively. Background seeds were also selected from the heart and kidney regions to prevent over-segmentation. Thereafter, all the seed points were converted to corresponding supervoxels in the liver VOI because the graph cuts algorithm is applied to supervoxels. Then the intensity distributions of foreground and background were represented as Gaussian Mixture models (GMM). The expectation maximization algorithm was applied to estimate the parameters of GMMs. The MICCAI-SLiver 2007 dataset was used in this work.The quantitative evaluation of segmentation results showed that the proposed method performed accurately and efficiently. The average results achieved are 7.87% for VOE, 1.31% for RVD, 1.29mm for ASD, 2.50mm for RMSD and 23.56mm for MSD. The proposed segmentation algorithm succeeded in separating the liver from the heart and kidney to avoid over-segmentation and handled the presence of tumours very well. However, some small over-segmentation and under-segmentation occurred near the liver boundaries due to it having a lower contrast, at the tip of the liver and high intensity regions as well as at the place of the vena cava.

The segmentation approach introduced by Goryawala et al [41] requires minimal user interaction and the algorithm integrates a localized contouring algorithm with a modified k-means method and localized region growing. The modified k-means algorithm is the first step. A central slice of the liver is presented to the user who selects five points belonging to the liver, surrounding organs, peripheral muscles, ribs/spinal cord and outside of the body. The user-selected seed points are used as initial centroids for clustering. The modified k-means algorithm is applied minimizing within-cluster sum of square in the online update mode. Clusters from mask 1 (liver) and mask 2 (surrounding organs) are combined to obtain a rough liver region. The next step is single slice contour initialization. The user selects a point from the liver mass in the first and last slice. Linear interpolation is used to generate the intersection point of the seed line which is used for each slice. Region growing is performed based on a threshold devised in terms of the standard deviation of the image. This algorithm absorbs portions of neighbouring regions like the spleen and stomach. Next a volume of interest (VOI) based correction method is used. This algorithm is based on the idea that a voxel belongs to the liver in 3D space only if it's marked by region growing in at least two of the three views (axial, transverse and coronal). A 3D corrected mask is generated. An algorithm was developed to automatically select the largest slice which is the slice in which the liver seen in its largest extent. The selected slice is presented to the user to set the liver boundaries in that slice. A mask is produced showing the largest slice of the dataset which is applied to all slices in the dataset in order to exclude portions not belonging to the liver region. Lastly, a localized contouring algorithm is employed. The corrected mask is used for initialisation of the localized contouring algorithm. The contour is modelled as a smoothed approximation of the Heaviside function and then fitted to reduce the Chan-Vase energy approximations of the interior and exterior of the contour. A region of a specified radius is selected about the point of interest

to restrict energy calculations to a localized region. The interior of the contour is defined as the final segmented liver and the volume is calculated. The results achieved by the proposed algorithm were an average of 72.38ml for volumetric error and 2.78% for absolute percentage error. An average Dice coefficient of 0.92 was obtained for the proposed algorithm. The segmentation was done correctly by this method but there were some cases where the surrounding organs were of similar intensity which resulted in incorrect segmentation.

A fully automatic algorithm for fast segmentation of liver tissue and tumours from CT scans was presented by Massoptier and Casciaro [74]. A statistical model-based approach was combined with an active contour technique and an automatic classification technique. First, each slice of the volume is divided into 64 squared regions and the mean intensity and standard deviation of each region is calculated. The internal abdominal regions which have minimum standard deviation are separated from all volume slices and sorted in ascending order of mean values. Since the liver is the largest organ and has homogenous intensity most of these regions belong to the liver. Then, the liver ROI is selected. The statistical information allowed for the implementation of an automatic adaptive histogram-based thresholding technique to partition images. The threshold was evaluated using the mean and standard deviation of the ROI. A gate function is then applied to the volume to select the voxels that belong to liver tissue in function of their image intensity value. Thereafter, morphological techniques were applied to isolate the liver. This was done by connected components labelling and largest component selection. Hole-filling concluded the liver surface initialization. The third stage is the refinement of the liver surface segmentation by use of the GVF snake model. First, the edge map and GV field were computed for each volume. Then, initial snakes are produced from the previous stage. Lastly, the GVF snake deformation is performed producing an accurate and smoother live surface. The last stage of this methodology is automatic tumour segmentation. The results of the segmentation showed that the statistical model discrimination was fast, robust and reliable in obtaining the initial segmentation of the liver surface. However, after this step, the outer rim regions of liver were not segmented completely. This is due to the image intensities of these regions decreasing and enlarging its variance. Since the intensities decreased smoothly, the GVF snake was able to exploit this and find the true boundaries resulting in more accurate results. The proposed algorithm achieved an average DSC of 94.2%, false-negative ratio of 6.3% and false-positive ratio of 3.9% for liver segmentation.

Huang et al [53] proposed an automatic liver segmentation approach using a single-block linear detection algorithm (SBLDA). The SBLD segmentation consists of seed initialization in traditional SBLDA. During the liver edge extraction part of the methodology, histogram analysis is used to roughly define the maximum and minimum liver intensity ranges. The seed image is thresholded using the determined range and is used as the initialized liver area. Then the SLBD process is implemented using a ratio parameter, confidence matrix and a predefined threshold value to determine whether each block of pixels belong to the liver or not. If the

threshold value is too low then the liver area with noise is misclassified as liver edge and if its too large then some liver edge may be missed and non-liver areas would be misclassified as liver. This method was evaluated on the 3Dircadb1 dataset produced a high accuracy rate of 99.03% and specificity of 98.65% which proves that this is a promising approach. However, some small segmentation error existed in CT images with small liver regions and when tumours occur at the edge of the liver.

Campadelli et al [18] proposed a grey-level based liver segmentation technique which avoid problems due to inter- and intra-patient grey-level variability. The methodology consists coarse liver volume extraction and liver volume refinement. First, the biggest 3D connected component is extracted. The liver position is exploited to automatically localize the liver. Since the liver occurs below the heart and at its left side, a coarse heart region is determined in each successive slice to form an initial 3D heart volume which is further refined using a 3D region growing algorithm. Next is the coarse liver volume extraction. A 3D bounding box is defined below the heart volume. Histogram analysis shows a unique peak in the grey levels which correspond to the liver. A liver range is determined from the histogram. The segmentation is then interpreted as a 3D labelling problem where the body box volume is partitioned into five disjoint classes: liver, bones and kidneys, spleen, stomach and organs with similar grey levels and background. The labels are assigned according to grey levels and spatial relationships between neighbouring voxels and is achieved by the minimization of an energy function using the -expansion algorithm which employs a graph cut algorithm. The biggest volume corresponding to the lowest label is the liver. The final stage is liver volume refinement. In this stage, voxels not belonging to the liver are removed from the liver volume. Three 2D images are created by projecting the liver volume onto the Y-Z, X-Y and X-Z planes. The voxels not belonging to the liver are identified as their projections have the lowest values in either one of the three images. Next, an iterative process is applied to consider the relationship among neighbouring slices. Consecutive slices are compared and voxel are deleted where they do not intersect with liver regions of neighbouring slices. This procedure is applied along the X, Y and Z axes. To recover some liver regions that may have been removed by the previous process, a 3D region growing algorithm is applied which produces the final segmentation result. The segmentation method was evaluated and results were compared to ground truth data which were manually traced by experts. The results achieved a symmetric volume overlap of 95.29% and a standard deviation of 1.22%. In some cases, over-segmentation occurred due to voxels belonging either to the stomach or muscles while under-segmentation occurred due to fatty liver near the liver edges.

Stawiaski et al [94] presented an application of minimal surfaces and Markov random fields to segment liver tumours. An interactive method is implemented to delineate tumours in 3D CT images by applying these models to the region adjacency graph of a watershed transform using a region graph instead of a pixel graph. The first step of the methodology is the definition of a sub-volume contain-

ing one or more tumours that need to be segmented. This was done manually to reduce computation time of the segmented algorithm. Thereafter, the methodology is based on a region adjacency graph of an unsupervised watershed segmentation. From this initial segmentation result, a region adjacency graph is extracted and optimized. The classification of the liver pixels in either the tumour class or healthy tissue class provides the tumour boundaries. In order to achieve this the liver pixels are modelled as a Markov random field and classification is performed through the maximum posteriori estimation. Classification is supervised by user defined markers that specify tumour and healthy tissues. The liver boundaries are extracted by computing the minimal surface based on user defined markers. The user has to specify normal liver tissues, tumour tissues and external tissues surrounding the liver. An unsupervised watershed transform of the morphological gradient of the original CT volume is used to produce a region adjacency graph. The watershed transform produces a partition of the image composed of small and numerous homogenous regions. The quality of the first unsupervised segmentation is important to guarantee a minimal loss of information. An approximate surface using the region adjacency graph is used to extract the liver boundaries. A combinatorial problem was proposed to achieve this: finding a surface composed of a finite union of watershed contours such that the surface minimizes a given geometric functional. This combinatorial problem was then solved using graph-cuts optimization on a region adjacency graph to produce the extracted liver boundaries. Then another segmentation method is used to detect the tumours in the liver. The image to be segmented is modelled as a Markov random field where each random variable corresponds to the mean value inside a region of the watershed transform. The prior function implemented is contrast sensitive and therefore takes into account the contrast between two regions and is equivalent to a constant in the areas where the contrast is low. This allows for efficient boundary detection of high contrast tumours, while low contrast boundaries are smoothed so that the surface of the object is minimized. Finally, the energy function is minimized by computing a minimal graph cut. The method achieved an average volumetric overlap error of 29.49%. Some evaluation results showed that some structures were misunderstood and were required to be extracted but were not. The method proved to exhibit promising results, however many problems exist. The segmentation of multiple tumours occurring in one liver required the user to mark the tumours separately. This is because the method merged tumours in a single object if they occur too close to each other. The user interaction required in this method increases processing time of the segmentation. Another problem is that the method doesn't include any preprocessing of the images which could simplify the segmentation and classification steps in the methodology.

A novel system for automatic detection and segmentation of focal liver lesions in contrast enhanced CT images was presented by Militzer et al [76]. The system uses a probabilistic boosting tree (PBT) to classify points in the liver as either lesion or parenchyma. It also integrates an iterative classification scheme that incorporates knowledge gained from earlier iterations into later decisions. The

presented algorithm consists of four steps: liver segmentation, intensity standardization, voxel classification and post processing. The method adopted for liver segmentation is based on modelling the liver by a hierarchical mesh-based shape representation. The liver is detected estimating its pose and location on the coarsest level using a marginal space learning scheme. Then the model is refined by applying a learning-based boundary localization. This aids the system to become reliable to heterogeneous intensity patterns. The liver surface is then decomposed into patches depending on surrounding anatomical structures and patch dependent classifiers are employed to cope with different texture patterns. The second step consists of normalizing the intensities to make the images more comparable. A non-rigid matching of the histogram of each target image to the histogram of one reference data is conducted. The estimated mapping is applied to the intensities of the target image. The third step is voxel classification where each voxel of the liver is classified as either lesion or parenchyma by previously trained classifier. Here, a PBT is used in combination with AdaBoost and decision stumps. The final step includes smoothing of the probability map generated in the previous step using a median filter. Then, morphological opening is used to eliminate small and isolated false positive detections. Lastly, thresholding is used to convert the image into a lesion candidate mask.

Fernandez-de-Manuel et al [25] proposed an active surface method to segment liver volumes from pre-operative CT images using a level-sets active surface method. The method finds the surface that minimizes an energy function combining intensity inside and outside the surface, gradient information and curvature restrictions. The implementation is based on a level set technique and requires a single seed point inside the liver to initialise the active surface. The method is implemented following a multi-resolution strategy in three steps. The algorithm starts from a small surface region from a seed point within the liver. The first step includes reducing the resolution of the images by a quarter using bicubic interpolation. The initial segmentation approximates the liver region. In the second step, the resolution is reduced by half. The growing surface is initialized by the previous step and the result iterates further to the actual liver surface. Lastly, the resolution is set to the original resolution to perform a final growing of the surface resulting in the segmented liver. An active contour method based on the minimization of an energy function that depends on the intensity values inside and outside the curve at each iterative step is used to segment the liver. Additional image gradient information is integrated into the function to improve the robustness of the algorithm. The method performed well however, areas of the heart and vena cava were misclassified and the method requires user input for an initial seed point.

Lim et al [24] proposed a new approach to automatic segmentation of the liver for volume measurement in CT images. The proposed segmentation scheme consists of three stages. The first stage is image simplification, the second stage is search range detection with initial liver contour by using morphological filtering and the last stage is contour-based segmentation using a labelling-based search algorithm that refines the initial liver boundary obtained in the second stage. During image

simplification the abdominal CT image is divided into non-overlapping blocks and the right bottom region, where the liver doesn't occur, is discarded. Then, the intensity distribution is analysed to obtain the liver grey-level range which is used for multi-level thresholding. This removes most non-liver tissue. The second stage is search range detection which includes four steps. Here, multi-scale morphological operations are used to find the first and second search regions. The multi-scale morphological filter uses a combination of erosion and dilation with a flat structuring element that combines shape, size and scale. Multi-scale morphology is performed recursively based on region-labelling and clustering. However, if the previous liver boundary is currently being used for segmentation, region-labelling is omitted. The second step is region-labelling where a 4-connected region-labelling algorithm is implemented that uses a breadth-first approach. This reduces noise created by the previous step and detects the coarse liver region. The largest labelled region is marked as the candidate region of the liver. The third step includes the implementation of a modified K-means algorithm. The algorithm uses three centroids where the middle centroid corresponding to the mean intensity of the liver is calculated with each iteration but the other two centroids remain fixed to the minimum and maximum value in the labelled liver region. This produces a finer initial liver region. The final step in the second stage is the second morphological filtering to reverse the possible reduction of any liver tissue from the previous step. The second search region is constructed based on the result of the morphological operation, region-labelling and modified K-means clustering. The final search region is determined by excluding the second search region from the first search region. The last stage is contour-based search algorithm that deforms the initial liver boundary within the search range to find a final liver contour. First a gradient-label map is created. The isolabel-contour map is enhanced by using gradient magnitude. The spatial gradient of the search range image is approximated by use of a morphological gradient operator and the gradient image within the search range is reversed and normalized producing the gradient-labelled map to be used in the next step. The deformable contouring starts at the lowest located pixel of the initial contour toward the clockwise direction on the gradient-label map. A local cost function is used to determine the optimal path from each pixel. This stage produces the final liver boundary. The error rate of volume measurement is 3%. However, this method required the CT images to be oriented in a supine position.

Furukawa et al [37] developed a liver segmentation method that extracts liver regions based on two stages. The first stage is a rough extraction based on a maximum a posterior probability (MAP) estimation and the second is a precise segmentation using a level set method. The first stage is a rough extraction of the liver based on MAP estimation and includes four steps. The first step is MAP based segmentation. Liver regions are extracted by calculating posterior probabilities of tissues at each voxel and assigning labels corresponding to the class with the maximum a posterior probability. It is assumed that there are four kinds of tissues which are liver, heart, right kidney and other tissue. The region being processed is limited to the region surrounding the liver where prior probability is greater than

zero. Step two is MAP based segmentation with different normalization. Here, the input image is normalized based on the top position of the right kidney, centres of mass and the width of the right lung along the y-axis. The liver regions are extracted again by the same way described in the first step. The third step is the removal of the heart region using a depth map of the lungs. A two-dimensional depth map of the lungs is generated and the depth values are interpolated to approximate the bottom surface of the heart. Two reference pixels are determined in each row of the depth map and linear interpolation is performed at pixels between the two reference pixels. The voxels with a z-coordinate equal to or less than the corresponding depth value are removed from the extracted regions in step two. Lastly, step four is the extraction of metastatic carcinoma and cyst using likelihood images and fusion of the extracted regions. The step is performed only when the average probability value of the liver region is higher than the liver threshold value. The likelihood image of abnormal regions is generated. The likelihood of lesions is multiplied by a prior probability to eliminate false positives extracted in the outside of the liver. Candidate regions are determined by simple thresholding followed by morphological opening and closing operations. The second stage of segmentation includes the level set method where the level set function is calculated using the gradient of probabilities convolved by the gaussian filter. A new term is introduced to the level set function to supress possible over extraction and prevent muscles adjacent to the liver from being incorrectly extracted as liver. It was evident that the method performed well producing an average volumetric overlap of 87.2%. However, the portal vein and inferior vena cava was over extracted as part of the liver and in some cases, parts of the kidney and heart were incorrectly segmented. The detection of lesions was inaccurate as only part of the lesions was detected.

Classical Learning Based Segmentation

Zhang et al.[109] and Erdt et al. [29] presented approaches for liver segmentation based on a statistical shape model (SSM). Zhang [109] uses an SSM with an optimal surface detection strategy. The liver shape model was built from various ground truth segmentation results. Liver localization by 3D Generalized Hough Transform (GHT) during which the triangulated mesh of the average shape model is used as a template shape. Points correspondence for the SSM and alignment of all training triangulated meshes based on the similarity transform which determines which regions are part of the liver and which regions are not. Thereafter, the principal component analysis (PCA) is applied to represent variation modes of the training sets. Lastly, the shape model is deformed to adapt to the liver contour through an optimal-surface detection technique based on graph theory. This work was evaluated on the MICCAI-SLiver dataset. However, segmentation errors occurred at some long and narrow regions of the liver. The approach presented by Erdt et al [29] is based on a multi-tiered framework that is robust against model initialization errors. The segmentation uses a locally constrained shape model. The SSM is built on non-public reference shapes of the liver. Local shape variation is incorporated

into the deformation term and since the model is explicit, correspondence is given through landmarks. The Gaussian curvature is computed at each point of all training shapes. The mean curvature is stored as a weight to define regions of differing elasticity. To prevent the model from leaking into structures surrounding the liver a local gradient threshold and an adaptation weight is applied which represents the influence of a found boundary point in the deformation. In the case of deviation from trained samples, image specific constraints are derived from the current position of the shape model in the image. The adaptation is restricted to those regions which show a high curvature in the training and the similar high curvature in the model shape after adapting to the SSM space. Image specific intensity constraints are calculated by using thresholding and histogram analysis techniques. The image specific shape and intensity constraints are updated in every deformation of the model i.e. with every iteration. Rule based boundary detection is implemented based on sampling intensity patterns and assigning a fitness value to decide whether a point lies on the boundary or not. The rule-based system deactivates boundary points in the adaptation to remove outliers and the remaining points are interpolated by the optimization. This approach was tested on four different datasets including the MICCAI-Sliver dataset and proved to have high accuracy rates for segmentation. This also shows broad capability of the approach and it is more robust than standard SSM segmentation.

A 3D medical image segmentation algorithm was presented by Seghers et al [88]. The algorithm models an object as a set of landmarks augmented with local appearance in which multiple shape models are used. A global cost function incorporating local intensity and local shape is optimised iteratively. The method has two parts which are model building and segmentation. Model building consists of training shapes, intensity model and shape model. First, one image with an average liver was randomly chosen as a reference image from the training set. The algorithm requires that all training shapes have corresponding landmarks which was obtained by first taking the reference intensity image and non-rigidly registering it to all other intensity images using a B-spline transformation model and mutual information. This produces a set of transformation fields that warp the reference image to the other training images. These transformation fields are applied to the surface mesh of the reference image to align the reference shape to the shapes of all the training images. Lastly, a second point-registration algorithm is applied to find accurate correspondences. Then the intensity and shape models are created. A set of feature images are computed from the intensity image. A spherical intensity profile is collected at the landmark point. For each landmark a statistical intensity model is obtained by estimating the mean and covariance of each feature profile from all images in the training images. A statistical model is built for every edge vector of the triangulation. A shape cost of an edge vector is used to define training shapes. The second part is segmentation which implements a search algorithm that consists of two steps. During the first step, for every landmark a set of candidate locations is extracted by scanning the image. During the second step, one candidate has to be selected for each landmark taking both intensity and shape knowledge

into account. Finally, post-processing is done to compensate for the initial affine registration which includes projecting the surface mesh in the original image space. It was evident in the results obtained from the segmentation method that this method does not perform very well in most of the test images. Two problems were observed. One was that the incorrect training meshes were used to build the model and the second was that the search regions in the algorithm missed the lower part of the liver in some cases.

A semi-automatic method was developed for liver tumour segmentation from 3D CT images by Zhou et al [110]. The method first trains a Support Vector Machine (SVM) classifier to extract tumour region from a single 2D slice by voxel classification. Then the extracted tumour contour is projected to its neighbouring slices. This propagation procedure continues until all slices containing tumour are processed. This method of segmentation was tested and evaluated on ten tumours. In many cases, the segmented volume was almost double the size of the reference volume and it only matched one third of the reference segmentation such that two thirds of the segmented volume belonged to false positives. This method proposes a limitation in that although 3D information was used in this investigation, the voxel classification was done on the basis of 2D slices. Hence, the segmentation result may not keep good smoothness and continuity in 3D.

Li et al [71] proposed an approach for liver segmentation that combines high order statistical texture features and anatomical structural features. The methodology consists of three major processes including texture analysis, liver distribution image calculation and liver organ localization. This methodology begins with texture analysis where the statistical approach and transform approach were utilised. The Gray-Level Co-occurrence Matrix and the nine Haralick texture descriptors were calculated. The Haralick texture descriptors include entropy, energy, contrast, sum average, variance, correlation, maximum probability, inverse difference moment and cluster tendency. The wavelet transform is used to calculate the wavelet coefficients. The wavelet transform was chosen over the Gabor transform and the Fourier transforms because it can represent textures at a suitable scale by varying the spatial resolution. The Haralick texture descriptors and wavelet coefficients are used as inputs to the next stage of the methodology. The second stage of the methodology is the liver distribution image calculation. The pixels in the image are classified by Support Vector Machines (SVM) to be either a pixel that has the most possibility of being liver or not. This results in the liver distribution image is a binary image which can indicate most of the liver correctly. The last stage is liver region localization where two problems are dealt with. The first being misclassified pixels and the second being that shape and spatial information is not considered. During this stage, binary morphological operators are implemented to obtain the delineation of the liver out of the distribution image. First dilation and erosion are applied to get connected regions and then the largest object is retained, pixels that are misclassified as non-liver pixels are connected and smoothing operators are applied. The experiments carried out produced accurate results achieving as high as 97.3% on true positive volume fraction. However, the results obtained

on the false positive volume and the false negative fractions were 11.1% and 5.1% respectively.

A hierarchical, learning-based approach was proposed by Ling et al [67]for automatic liver segmentation. Three techniques are combined to solve the segmentation problem. First, a hierarchical framework is used to monitor the accuracy propagation in a coarse-to-fine fashion. Second, marginal space learning (MSL) and steerable features are used for robust boundary inference. Third, a shape space initialisation is used to improve traditional methods that are limited to similarity transformation. The shape of the liver is represented by a closed triangle mesh and a conical mapping from the liver surface to a unit sphere is built. The mesh serves as a reference mesh that provides point correspondence and spherical coordinates. A mesh pyramid is built by recursively applying a down sampling process. A volume pyramid is built such that meshes at different layers correspond to volumes at different resolutions. Using the hierarchical shape representation, statistical shape models are built for each layer. Liver detection requires finding the best pose of a liver given the volume which is determined by the nine-parameter space to three marginal spaces which learned using the probabilistic boosting tree (PBT). 3D Haar features are used for location detection and steerable features are used for orientation and scale inferences. The mean shape is placed in the detected box. Learning techniques are used for inference and shape initialization. The training samples are extracted from the 3D parameter spaces in the range for a coefficient determined by the corresponding eigenvalues. The feature set is intensity based features which are sampled along the surface of the mesh. The next step is boundary refinement which consists of three key components: hierarchical scheme, probabilistic boundary response and patch-based boundary classifiers. Starting from the coarsest layer, the mesh at the current layer is first refined and then up sampled to a finer layer. A thin plate spline (TPS) warping is used for the up sampling between layers. This procedure is iterated until the finest layer is reached. The learning based boundary detection technique uses PBT and steerable features to learn the maximum boundary probability response. Spherical coordinates of mesh points are included as features and are computed through the mapping from the mesh to a unit sphere. Patch based boundary refinement is implemented as a result of the heterogeneity of texture pattern along liver boundaries. The liver surface is decomposed to five patches: liver-lung, liver-kidney, liver-heart, liver-tissue and liver-misc. Each point in the base mesh is allocated to a patch which is dependent on the estimated patch frequency of the point. A hard-patch scheme is used as it takes only the response from the patch with maximum prior probability. This approach produced promising results as it achieved a mean error measured in the average symmetric surface difference of 1.76mm 0.99mm and a median of 1.45mm. This approach produced excellent performance in accuracy.

Huang et al [54] proposed a method that uses a kernel-based Extreme Learning Machine (ELM) to detect and segment liver tumour voxels in CT scans. The ELM is designed as a single hidden layer of feed forward networks which has a learning speed that is faster than that of other learning algorithms such as an SVM. The

kernel ELM includes a positive regularization coefficient in the learning system making it more stable. The feature vector is calculated for each voxel. Texture features are generated in terms of the voxels neighbourhood mean and variance, intensity, intensity power, entropy, intensity co-occurrence, Laws texture and Sum and Difference Histogram. Tumour detection consists of two techniques. The first is two-class ELM training and testing in which the ELM is trained with samples from two classes i.e. tumour and non-tumour which are ground truth data. The second is a one-class ELM for tumour detection where only healthy liver voxels are selected and characterized as healthy liver voxels and lesion voxels are detected simultaneously. The last stage is tumour segmentation during which the boundary of the tumour is extracted. The one-class ELM outperformed the conventional ELM as was evident in the results.

Deep Learning Based Segmentation

Hu et al [52] proposed an automatic segmentation framework based on 3D convolutional neural network (CNN) and globally optimized surface evolution. The deep 3D CNN is trained to learn a subject-specific probability map of the liver which gives the initial surface and acts as a shape prior. Then global and local appearance information from the prior segmentation are adaptively incorporated into a segmentation model which is globally optimized in a surface evolution way. The methodology includes deep 3D CNN training, initial liver segmentation by the 3D CNN, development of the energy model using the shape prior and global optimization-based surface evolution. The 3D CNN predicts a probability map as a subject-specific prior which assigns each voxel a probability of being a part of the liver for the target image. Thereafter, the image is thresholded to produce an initial liver segmentation which accurately locates the liver and predicts a rough shape of the liver. Due to severe over- or under-segmentation especially around the liver surface, a novel segmentation algorithm was proposed to accurately delineate the liver surface. An energy function based on global and local statistics from prior information is used to refine the initial segmentation. The energy function is formulated in a hybrid model that integrates region statistics, shape prior constraints and gradient-edge map. The model selectively acts as a region-based or edge-based model in different regions. In order to constrain the model with the shape prior the prior probability map is represented by a negative log-likelihood which determines whether a voxel belongs to the background or the liver. The data term incorporates two kinds of image statistics which are intensity distribution and region appearance. To obtain this the probability density functions of a voxel is determined and a region appearance distance potential is introduced in order to capture edges weak in gradient. In abnormal liver regions the intensity distribution and region appearance are often different from global statistics of healthy liver tissues. Therefore, spatial location is considered. The data term adaptively utilizes global and local prior without separating the liver region. In order to determine the global-prior-based data term, firstly, the PDF of foreground and background from the shape

prior is globally estimated. Intensity, local binary pattern (LBP) and local variance are combined for region appearance description. The global estimation of region appearance distance potential of a voxel is calculated with the application of the Wasserstein distance. A local-prior-based data term is developed by locally by estimating the intensity distribution which is done by employing nonparametric model via a kernel density estimator. The local prior is only employed in abnormal liver regions to estimate the PDF of the foreground and region appearance difference. By integration and rearrangement of the equations obtained, a model is produced which propagates a contour or surface to its global optimum in each iteration by solving a sequence of convex optimization problems by use of a continuous max-flow algorithm. The segmentation was evaluated using five metrics: volumetric overlap error (VOE), relative volume difference (RVD), average symmetric surface distance (ASD), root mean square surface distance (RMSD) and maximum symmetric surface distance (MSD). The average results obtained in this experiment were a VOE of 5.36%, RVD of 0.03%, ASD of 0.96mm, RMSD of 1.84mm, MSD of 19.20mm and an average Dice similarity coefficient of 97.24%. It was evident that the proposed method is accurate and may be efficient for clinical practice.

An automatic 3D liver location and segmentation method was presented by Lu et al [70]. The method consists of two main techniques: simultaneous liver detection and probabilistic segmentation using 3D convolutional neural networks (CNNs) and accuracy refinement of initial segmentation with graph cut and the learned probability map. The architecture of the proposed 3D CNNs contains one input feature map corresponding to a CT image block. It then stacks eleven convolutional layers by some filters. Each layer is followed by a Rectified Linear Unit (ReLu). The network uses pooling and softmax layers as well. The first convolutional layer contains 96 feature maps which is then followed by the application of a local response normalization scheme and a mean pooling layer. The second convolutional layer is then applied followed by another mean pooling layer which generates 256 feature maps. The next five convolutional layers have 512 feature maps. The next three convolutional layers are followed by rearranging layers which can contain unambiguous boundaries. The convolutional layers occurring after each rearranging layer can eliminate blocking artefacts. The last rearranging layer produces 16 feature maps. The last layer is a log-regression layer that produces an output that range from 0 to 1 which represents the probability of each voxel in the output image block being classified as a liver voxel. The initial liver region location consists of training the CNNs for 53 iterations to generate the liver likelihood map. The parameters of the 3D CNNs are trained by gradient-based optimization. Then, the initial liver shape is located by thresholding that probability map produced by the trained 3D CNNs. The final stage, segmentation refinement, consists of the liver probability map being used to initialize the graph cut and incorporate it into the energy function to achieve an accurate segmentation result. The main advantage of this method is that it does not require any user interaction. The results achieved by this method are a VOE of 5.90%, RVD of 2.70%, ASD of 0.91mm, RSD of 1.88mm and MSD of 18.94mm. This segmentation method produced more accurate segmentation re-

sults when compared with those of state-of-the-art automatic liver segmentation methods.

Vivanti et al [103] presented a new automatic algorithm for liver tumour segmentation in follow-up CT studies. This method combines a follow-up based detection with CNN-based segmentation where the inputs are a baseline CT scan and a delineation of the tumours in it and a follow-up scan and the outputs are delineated tumours in the follow-up scan. The presented method consists of registration, deep learning and segmentation. The registration stage begins with automatically computing a liver mask using a liver segmentation method that relies on Bayesian classification, adaptive morphological operations and active contours. This segmentation is performed for baseline and follow-up scans. Next, a ROI is defined that contains the follow-up tumour with high probability. The baseline delineation is transferred to the follow-up scan and the follow-up tumour ROI is doubled in each direction to account for possible tumour growth. A CNN is used to classify each voxel as being healthy liver or tumour. The classification is based on voxel intensities in an axis-aligned square centred at the voxel. The CNN used has seven hidden layers. The input layer has one node for each pixel. The first layer is convolutional with 48 kernels. Each kernel computes a convolution of the input followed by a ReLu function and a pooling layer. The third and fifth layers are convolutional with 48 kernels followed by a ReLu function and a pooling layer. Layer seven is fully connected with 200 nodes followed by a ReLu function. The output layer is the classification layer which is based on a softmax function where the probability of a voxel belonging to a class is computed. The segmentation of the ROI is done by classifying all its voxels in four steps. First, the trained CNN is run in feed-forward to classify each patch. Then, the non-liver voxels are classified as healthy tissue using the liver mask of the follow-up scan. The experiments carried out by two metrics: VOE and ASD. The results obtained were 16.75% for VOE and 2.05mm for ASD. The overall success rate for this method was 90.47%.

An automatic segmentation method was presented by Christ et al [21] to segment liver and lesions in CT and MRI abdominal images using cascaded fully convolutional neural networks (CFCNs). This method uses two FCNs which are cascaded. The first FCN is used for segmentation of the liver to obtain the ROI for the second FCN which is used to segment lesions within the ROI. The CT and MRI volumes are trained in a CFCN. The first FCN segments the liver from the abdomen producing the ROI. This is then used as the input for the second FCN which segments lesions from the liver ROI. Post-processing includes the application of a 3D conditional random field (CRF) in order to capitalize on the location information across slices within the dataset. Thereafter, a final segmentation is produced. The presented method achieved a mean Dice of 93.1% to 94.3% for liver segmentation and 56% for lesion segmentation using the 3DIRCAD dataset. Using a clinical CT dataset, The Dice score achieved was 88% to 91% for liver segmentation and 61% for lesion segmentation. Using a clinical MRI dataset, the method achieved a Dice score of 87% for liver segmentation and 69.7% for lesion segmentation. It was evident that the method produced a higher accuracy rate than that of a method that

uses a single FCN for segmentation.

A novel 3D deeply supervised network (DSN) for automatic liver segmentation was presented by Dou et al [28]. This method employs a fully convolutional architecture to produce a high-quality score map which is processed further by the employment of a conditional random field (CRF) to obtain a refined segmentation. The network consists of eleven layers i.e. six convolutional layers, two max-pooling layers, two deconvolutional layers and one softmax layer. The deep supervision mechanism is involved in the third and sixth layers. The 3D convolutional layers and 3D max-pooling layers are alternatively stacked to abstract the intermediate features. Large kernel sizes are used to form a proper receptive field for liver recognition. After many stages of down sampling, the dimensions of feature volumes are gradually reduced and are smaller than the ground-truth mask. Therefore, 3D deconvolutional layers are developed to bridge those feature volumes to dense probability predictions. A per-pixel-wise binary classification error minimization problem with respect to the ground-truth mask is formulated for the learning of the 3D network. An additional supervision is injected into some hidden layers to counteract the effects of gradient vanishing which makes the loss back-propagation ineffective in the first few layers. Some lower-level and middle-level feature volumes are up- scaled using additional deconvolutional layers and then the softmax layer is employed to obtain dense predictions for calculating classification errors. Gradients are derived from these branch predictions and the last output layer which alleviate the effects of gradient vanishing. Although the 3D DSN generates high-quality probability maps, a graphical model is employed to refine segmentation results due to imprecise segmentation in ambiguous regions. A fully connected CRF model on the transverse plane is exploited which solves an energy function to refine the segmentation. The MICCAI-SLiver07 dataset was used. This method was trained on twenty CT scans and test on ten CT scans. The learning process of the proposed 3D DSN showed that it converges much faster and achieves lower training/validation errors than that of a pure 3D CNN. The segmentation results achieved were a VOE of 5.42% and an average symmetric surface distance of 0.79mm. It can be concluded that the presented framework is effective and efficient and achieves competitive segmentation results to state-of-the-art approaches.

Hoogi et al [51] proposed a fully automated pipeline for detection and segmentation of liver lesions and pathological lymph nodes. The pipeline includes three steps which are organ detection, lesion detection and lesion segmentation. And applies machine learning techniques such as marginal space learning and CNNs as well as active contour models. The organ detection step is based on marginal space learning where detection is performed by using a sequence of learned classifiers, starting with a few classifiers and ending with a classifier that models all the desired organ parameters. Each learned classifier is a Probabilistic Boosting Tree (PBT) with 3D Haar and steerable features and is trained. The architecture has two layers: A Discriminative Anatomical Network (DAN) and a database-guided segmentation module. The DAN supplies an estimate regarding the scale of the patient and portion of the visible volume and it detects a set of landmarks for navigating the

volume. The database-guided segmentation module uses the DANs output for detection of position, orientation and scale of the organs visible in the volume. Next is lesion detection based on cascaded classifications which includes three steps. 3D Haar-like features are exploited using an integral image of the organ sub-volume of interest. Then a classifier is used to perform feature selection and classifier training simultaneously. Lastly, self-aligned image features are used to train another classifier to prune candidates. The final step is lesion segmentation which incorporates a CNN and active contours. An adaptive approach was used to calculate the size of the local window. The adaptive window is calculated using the lesion scale and its texture. Texture analysis includes the extraction of Haralick features. The CNN estimates the location of the Zero Level Set contour relative to the lesion. Three possibilities are considered which are outside the lesion, near the lesion boundaries or inside the lesion boundary. The CNN is made up of two convolutional layers followed by two fully connected layers. Each convolutional block is composed of a filter bank layer, a nonlinearity layer and a max-pooling layer. The CNN outputs a probability for each of the three classes. These probabilities are used to set two weighting parameters: one that contracts the contour and one that expands it. The proposed methodology was evaluated using two datasets: the first comprising of 42 3D CT volumes containing liver lesions and 86 3D CT volumes containing pathological lymph nodes. The results achieved by this methodology was detection (sensitivity) of 0.51 for lymph nodes and 0.85 for the liver and segmentation (Dice) of 0.71 for lymph nodes and 0.67 for liver lesions. It was evident from the results obtained that this methodology proved to be robust in its handling of high lesion diversity.

Ben-Cohen et al [14] explored the use of a fully convolutional network (FCN) for liver segmentation and liver metastases detection in CT examinations. They proposed a network architecture that has 16 layers where all fully connected layers are converted to convolutions. A 1x1 convolution with channel dimension 2 is appended to predict scores for lesion or liver at each of the coarse output locations which are followed by a deconvolution layer to upsample the coarse outputs to pixel-dense outputs. Upsampling is performed in-network for end-to-end learning by backpropagation from the pixelwise loss. The initial network used is a FCN-8s DAG network which learned to combine coarse, high information with fine, low layer information. The addition of a lower level linking layer was also explored creating a FCN-4s DAG network. The input image was modified to have three slices, the relevant slice and two neighbouring slices, before and after, which were interpolated using linear interpolation. Input images and their corresponding segmentation maps are used for training with a stochastic gradient descent and GPU acceleration. Two networks are trained here, one for liver segmentation and the other for lesion detection. The softmax log-loss function was calculated pixelwise with different weights for each pixel class. Two framework variations were used for evaluation as well as the initial framework. The segmentation performance was evaluated using the Dice index, sensitivity and positive predicate values (PPV). The FCN-8s with 3 slices produced a Dice of 0.89, the FCN-8s produced a Dice of 0.88 and the FCN-

4s with 3 slices produced a Dice of 0.87. The FCN-8s with 3 slices produced the best results with an average sensitivity 0f 0.86 and an average PPV of 0.95. The lesion detection performance was evaluated using two metrics: true positive rate (TPR) and false positive per case (FPC). The dataset used is the MICCAI-SLiver07 challenge dataset. Here, the FCN-4s with 3 slices performed the best producing a TPR of 0.88 and a FPC of 0.74. The method presented here is promising for both lesion detection and liver segmentation. However, it should be noted that no significant pre-processing or post processing was implemented which could possibly have improved the results achieved by the presented method.

Post-processing Techniques

The aim of post-processing techniques is to improve the segmentation result produced by the primary segmentation techniques. Segmentation refinement techniques are initialized by the product of the previous technique in the segmentation pipeline.

Resampling was implemented in [101, 106, 53] in order to return the segmentation results back to the original spatial resolution. This is basically implementing the pre-processing techniques that were used in the reverse order. Zayane et al [108] used a Sobel filter for contour detection to delineate the liver boundaries. The most common post-processing technique used is morphological operators which are used in [52, 81, 94, 25, 37, 108, 103] for cavity filling, to smooth extracted surfaces and eliminate unconnected regions. A combination of blurring, thresholding, largest connected component selection and morphological techniques are also used in [101, 106, 65].

Lu et al [70] implemented a graph cut based segmentation refinement which is initialized by the rough liver region generated by CNNs. The aim of the graph cut based model is to label each voxel such that the general energy function is minimized where a coefficient controls the balance between the data fitting term and the boundary penalty. A regional cost term describes the similarity between a voxel and the background or foreground while the boundary cost term describes the discontinuity between two neighbouring voxels. A thresholding map is determined with the use of a roughly estimated intensity range of the liver which is obtained from the probability map. A local appearance term is introduced which is formulated from the image intensity, a modified LBP and the local variance of intensity. The data term is then computed by combining the probability map, thresholding map and local appearance map. A graph is defined to minimize the total energy function by the graph cut algorithm.

Christ et al [21] used a dense conditional random field (CRF) as a segmentation refinement. The dense CRF is specified using the probability map produced by a FCN and a complete graph formulated using an energy function that includes a Gibbs distribution and incorporates the Potts function. This allows consideration of both spatial coherence and appearance. The best labelling is estimated using a

mean field approximation algorithm and the weights and kernels of the CRF are chosen using a random search algorithm.

A novel segmentation scheme was presented by Beichel et al [13] based on a true 3D segmentation refinement concept that utilizes a hybrid desktop/ virtual reality user interface. The scheme consists of an interactive segmentation refinement that is initialized by the product of a graph cut based segmentation technique. The refinement technique allows the user to fix arbitrary segmentation errors and consists of two steps. The first step is chunk-based refinement where the segmentation result is subdivided into disjunct subregions or chunks. By thresholding, a binary boundary volume representing boundary/surfaces parts is generated and merged with the boundary from the segmentation using a logical "or" operation. The distance transform is calculated and the distance map results are inverted which is interpreted as a height map. Then all small local minima are eliminated to avoid over-segmentation. Watershed segmentation is applied to the distance map resulting in volume chunks. Refinement is done by the user selecting/deselecting predefined chunks using tools for interactive data inspection and selection methods. The second step includes the selected chunks being converted to a simplex mesh representation. Different tools allow for deformation of the mesh representation. The method was tested on ten datasets. After the initial graph-cut segmentation, five cases have an overlap error larger than 10% and the overall average overlap error is 14.3%. After using the chunk-based refinement method the average overlap error is reduced to 6.5% and after the use of the simplex mesh refinement it reduced further to 5.2%. This proves that the method delivers good segmentation results; however, the method relies heavily on user intervention.

2.3.5 Conclusion

Despite the vast amount of research available on the techniques explored to achieve accurate liver segmentation, it seems challenging to develop methods that will achieve higher accuracy rates. Many successful methods have been proposed however; they have their advantages as well as disadvantages.

Different types of techniques exist such as anti-learning-based methods, which include region growing, graph-cut and level set methods, and learning-based methods which include CNNs. The drawbacks of anti-learning-based methods stems from the fact that they are mainly based on gray-level information. They pose the challenge in preventing over-segmentation from surrounding tissues or organs with similar intensities and avoiding under-segmentation of inhomogeneous liver regions.

Examining the learning-based methods, either a series of 2D slices of the CT scan or a while 3D CT scan volume are fed as input into the neural network models in current methods. The methods discussed in this chapter do not narrow down or limit the search space for the liver segmentation. Reducing the search space reduces computational cost. However, the impressive success rate of learning-based techniques inspired the proposed method presented in this dissertation.

The next chapter, details a slice alignment method that can be used for the alignment of 2D slices of a CT volume.

Chapter 3

Slice Alignment

3.1 Introduction

Image alignment is a fundamental task in 3D medical image processing that serves to align two or more 3D scan slices. Medical imaging tasks such as those related to image fusion for image guided interventions and those related to motion correction and volume reconstruction have pushed the community towards developing more accurate and efficient strategies for image alignment [34].

Most methods do not simultaneously consider two significant difficulties regarding medical and CT scanned data registration or alignment[60]. Sometimes, consecutive slices may differ due to discontinuities in anatomical structures, distortions that are possibly caused by movement or cuts and tears. Inter-slice distance contributes to the degree of these effects. The larger the distance between the slices, the more pronounced the discontinuity in anatomical structures. Therefore, alignment and registration methods are required to be robust with regards to missing data or slices that are outliers. Another aspect that should be taken into consideration is that when slices are aligned sequentially, it may lead to different types of mis-alignment. In this case, each slice is aligned with respect to the previously aligned slice and if an error occurs the error is propagated through the rest of the volume. Thus, if the volume contains a large number of slices it will result in an observable global offset of the volume.

The accuracy of image alignment influences further processing that the image volume may undergo. Ensuring that the alignment of a series of images are appropriately aligned promises more accurate results if an object in a volume is to be segmented and/or reconstructed. This also allows the reconstruction of slices to produce smoother surfaces making the object appear more real.

3.2   Methodology

The proposed methodology consists of four stages. In the pre-processing stage, the 3D CT volume undergoes smoothing using median filtering. The second stage includes the serially acquired slices of the volume being processed in order to determine the optimal liver threshold by applying a Laplacian weighted histogram to the largest liver slice of the volume. The threshold value obtained is used in the third stage where the entire volume is skeletonized to remove unwanted tissue. The last stage is image alignment where each slice of the skeletonized volume is aligned with respect to neighbouring slices using the ECC image alignment algorithm.

Figure 3.1 depicts an overview of the methodology for slice alignment.
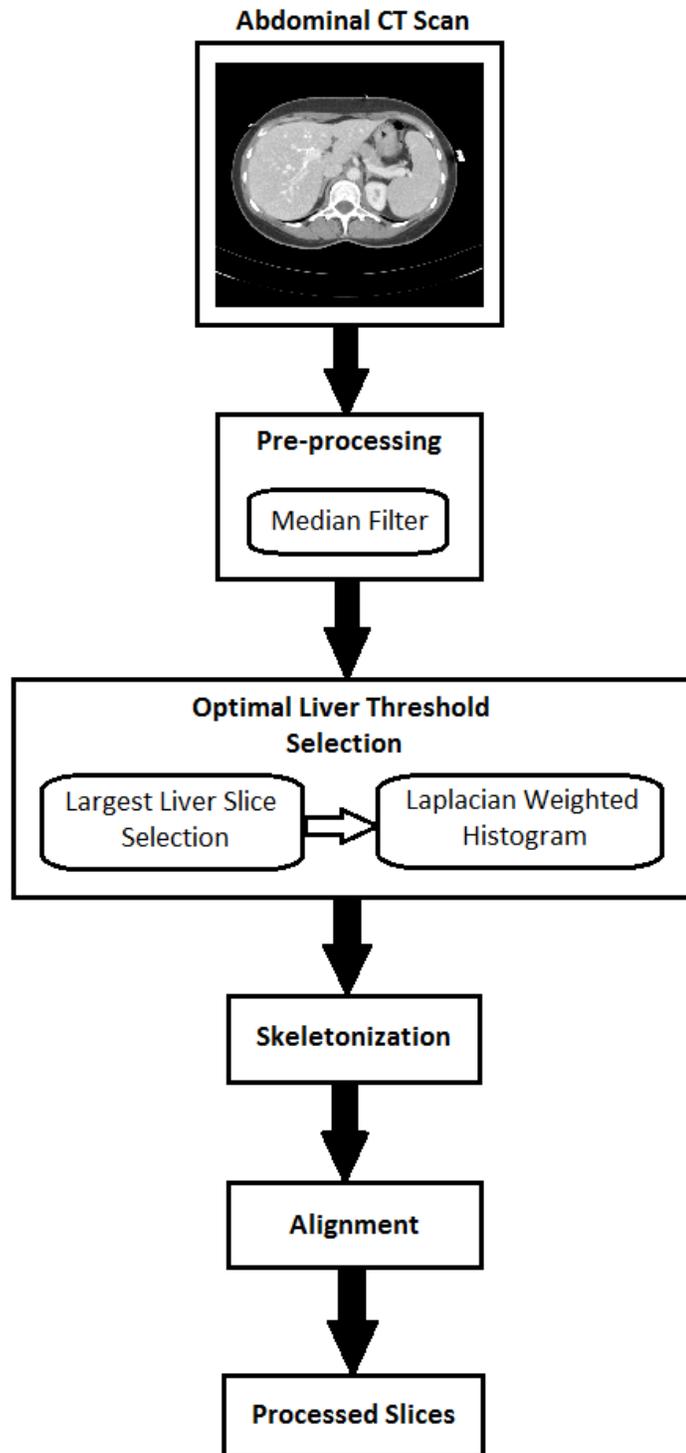
Figure 3.1: Overview of Liver Segmentation Framework

3.2.1 Pre-processing

The pre-processing technique used is median filtering. All smoothing techniques are effective at removing noise in smooth patches or smooth regions, but adversely affect edges. In the case of this particular application it is important to preserve the edges. For small to moderate levels of noise, the median filter is demonstrably better than a Gaussian blur at removing noise whilst preserving edges for a given, fixed window size. The whole 3D volume is smoothed using a 3D median filter. The median filter used is defined in Equation 3.1.

$$y[i,j,k] = median\{x[i,j,k],(i,j,k)\varepsilon w\} \tag{3.1}$$

where $y[i,j,k]$ is the new voxel value, $x[i,j,k]$ is the original voxel value and w is the 3D neighbourhood centered around location $[i,j,k]$.

3.2.2 Optimal Liver Threshold Selection

Two dimensional slices are serially acquired from the median filtered 3D CT volume. Since the liver occurs on the right side of the body and this stage determines the liver intensity range, the right half of the slices are extracted. The mean intensity is calculated for each half slice and the slice with the highest mean is declared the largest liver slice. This is the slice where the liver can be seen in its largest extent. The Laplacian weighted histogram analysis is applied to the largest liver slice.

Li et al. [65] implemented histogram analysis to determine the liver range for their work done in liver segmentation using fuzzy C-means clustering and level-set techniques. Huang et al. [54] also implemented histogram analysis to determine the liver range. In this pre-processing method a Laplacian weighted histogram is calculated for the largest liver slice. This method leads to better results and is much more efficient than convoluting the neighbouring voxels with a generalized filter.

The method presented by Gasser [38] computes a Laplacian weighted histogram to determine the optimal threshold of a volume for extracting isosurfaces and for volume representation. The methodology presented in this paper uses this method to determine the optimal threshold value for the liver tissue in a slice of the volume. The method was adapted to work with Hounsfield Units as the input is a slice of CT volume. Furthermore, it was adapted to work with the maximum and the minimum of the current CT volume due to the fact that the range of the intensity values differ from volume to volume.

After the histogram is calculated, it is normalized and the global minimum of the histogram is taken as the liver threshold value for the CT volume that the slice belongs to. The algorithm used to determine the threshold value is presented in Algorithm 1. The threshold value obtained from this stage is used in the skeletonization stage.

Algorithm 1 Laplacian Histogram Algorithm
___

 Input: largest liver slice

 Output: optimal threshold value

$lowerbound = lowestIntensityValue$

$upperbound = highestIntensityValue$

$numBins = (upperbound - lowerbound) + 1$

for $k = 0$ to $D, k = k + 1$ **do**

    for $j = 0$ to $H, j = j + 1$ **do**

        for $i = 0$ to $W, i = i + 1$ **do**

            $value \leftarrow x(i, j, k) + lowerbound$

            $L(i, j, k) \leftarrow (x(i, j, k) - x(i, j, k) - x(i, j + 1, k) - x(i, j, k) + x(i, j, k) - x(i, j + 1, k) - x(i + 1, j - 6, k + 1) - x(i, j + 1, k) + x(i, j, k) - x(i, j, k) - x(i, j + 1, k) - x(i, j, k))$

            $bins[value] \leftarrow bins[value] + L(i, j, k)$

        end for

    end for

end for

$highest = 0$

for $i = numBins - 2, i \geq 0, i = i - 1$ **do**

    if $bins[i] > highest$ **then**

        $highest \leftarrow bins[i]$

    end if

end for

$normalizedBins = array[numBins]$

for $i = 0$ to $numBins, i = i + 1$ **do**

    $normalizedBins[i] = bins[i]/highest$

end for

$G_{min} = normalizedBins[0]$

$binIndex = -1$

for $i = 0$ to $numBins, i = i + 1$ **do**

    if $lowest \leq normalizedBins[i]$ **then**

        $G_{min} \leftarrow normalizedBins[i]$

        $binIndex \leftarrow i$

    end if

end for

$G_{min} \leftarrow binIndex + lowerbound$
___

The laplace value $L(x)$ is computed at each voxel location. A histogram is built from the filtered dataset incrementing a histogram bin $G(I(x))$ where $I(x)$ is the image intensity at voxel location x by the Laplace value $L(x)$ at voxel location x.

$$G(I(x)) \Leftarrow G(I(x)) + L(x) \tag{3.2}$$

The laplacian weighted histogram is normalized. The global minimum G_{min} of the normalized histogram corresponds to the most dominant material transition which in this case is the liver tissue. The intensity value corresponding to G_{min} is the optimal threshold value for the liver tissue in that particular slice.

3.2.3 Skeletonization

3D skeletal forms refer to visually perceivable line segments achieved after removing closed surfaces in a 3D image until there is no closed surface left out for removal [59]. This iterated closed surface removal is called 3D skeletonization. The skeletal forms are also called superficial features. An empty widow is used to detect superficial features. The purpose of the skeletonization stage is to remove unwanted tissue that are insignificant for the alignment of the slices. Since the images are of the abdominal area, it is appropriate that the images are aligned with respect to the larger organs and the bones in the images. This ensures that when the images undergo further processing such as segmentation and then reconstruction of a 3D volume it produces a smoother surface which is more accurate.

The skeletonization of each 2D slice proved to be computationally expensive and time consuming. Therefore, the whole CT volume is processed at once. A skeletonization method is presented by Khan et al [59] where a 7-neighbourhood window is traversed throughout an entire volume but this method also proved to be very time consuming. So, the method is adapted to traverse the volume in less time, producing a result slightly different but achieving a skeletonization that is usable in this particular application. The algorithm for skeletonization is presented in Algorithm 2.

The algorithm consists of the iteration of the 3D blank window throughout the entire image volume of extent WxHxD where W is the width, H is the height and D is the depth of the image with the 3D blank window. At each iteration the values of the specified seven neighbourhood voxels are extracted and stored in an array. The seven neighbourhood voxels can be seen in Figure 3.2(b). The maximum intensity and the minimum intensity values of the array are determined. The difference is calculated and compared to the threshold value determined in the previous stage. If the difference is less than or equal to the threshold value, six out of the seven voxel values are changed to zero. The only voxel value that remains its original value is the center voxel. Do not force a value 0 to a voxel if it happens to be the corner voxel of the subimage under the window. If the difference is greater than the threshold value, then the blank window is moved to the next window position.

The iteration of the blank window is such that it moves three voxels horizontally with each iteration. The original algorithm moves the blank window horizontally one voxel with each iteration. The change made to the algorithm improved the performance speed of the skeletonization stage by approximately half the original speed. These steps are repeated iteratively until there is no 3D surface left out in the image to be removed. During the iteration process it is ensured that already skeletonized objects are not erased. The resulting images after skeletonization can be seen in Figure 3.3, where 3.3(a), 3.3(b), 3.3(c) and 3.3(d) are the slices of different volumes before skeletonization and 3.3(e), 3.3(f), 3.3(g) and 3.3(h) are the slices after skeletonization.

i-1,j-1,k-1	i-1,j,k-1	i-1,j+1,k-1
i,j-1,k-1	i,j,k-1	i,j+1,k-1
i+1,j-1,k-1	i+1,j,k-1	i+1,j+1,k-1

i-1,j-1,k	i-1,j,k	i-1,j+1,k
i,j-1,k	i,j,k	i,j+1,k
i+1,j-1,k	i+1,j,k	i+1,j+1,k

i-1,j-1k+1	i-1,j,k+1	i-1,j+1,k+1
i,j-1,k+1	i,j,k+1	i,j+1,k+1
i+1,j-1,k+1	i+1,j,k+1	i+1,j+1,k+1

(a) Blank 3x3x3 window

Front Plane   Middle Plane   Rear Plane
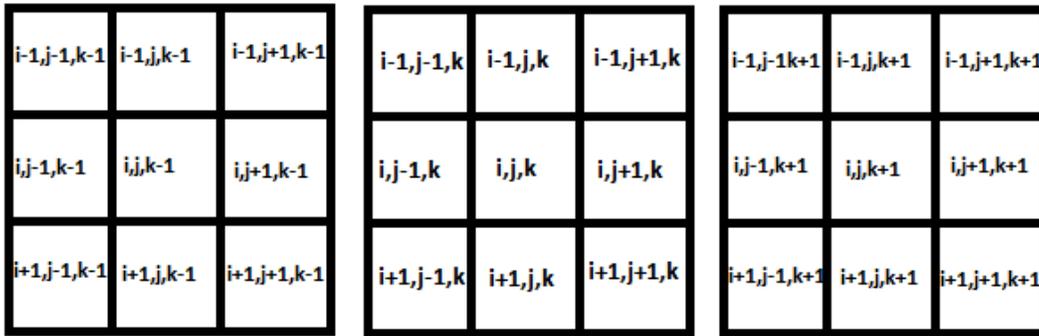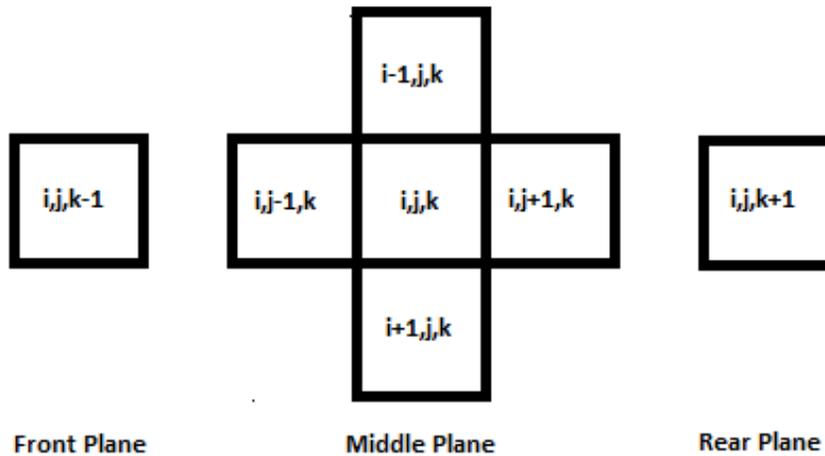
(b) 7-neighbourhood window

Figure 3.2: Blank window and 7-neighbourhood windows

Algorithm 2 Skeletonization Algorithm

Input: CT volume
Output: resulting volume after skeletonization

for $k = 0$ to $D, k = k + 1$ **do**
  for $j = 0$ to $H, j = j + 1$ **do**
    for $i = 0$ to $W, i = i + 3$ **do**
      $array \leftarrow E(i, j, k - 1), E(i - 1, j, k), E(i, j - 1, k), E(i, j, k), E(i, j + 1, k), E(i + 1, j, k), E(i, j, k + 1)$
      $Max(E(i, j, k - 1), E(i - 1, j, k), E(i, j - 1, k), E(i, j, k), E(i, j + 1, k), E(i + 1, j, k), E(i, j, k + 1))$
      $Min(E(i, j, k - 1), E(i - 1, j, k), E(i, j - 1, k), E(i, j, k), E(i, j + 1, k), E(i + 1, j, k), E(i, j, k + 1))$
      $difference \leftarrow Max(E(i, j, k - 1), E(i - 1, j, k), E(i, j - 1, k), E(i, j, k), E(i, j + 1, k), E(i + 1, j, k), E(i, j, k + 1))Min(E(i, j, k - 1), E(i - 1, j, k), E(i, j - 1, k), E(i, j, k), E(i, j + 1, k), E(i + 1, j, k), E(i, j, k + 1))$
      $T \leftarrow threshold$
      if $difference \leq T$ **then**
        $E(i, j, k - 1) \leftarrow 0$
        $E(i - 1, j, k) \leftarrow 0$
        $E(i, j - 1, k) \leftarrow 0$
        $E(i, j + 1, k) \leftarrow 0$
        $E(i + 1, j, k) \leftarrow 0$
        $E(i, j, k + 1) \leftarrow 0$
      end if
    end for
  end for
end for
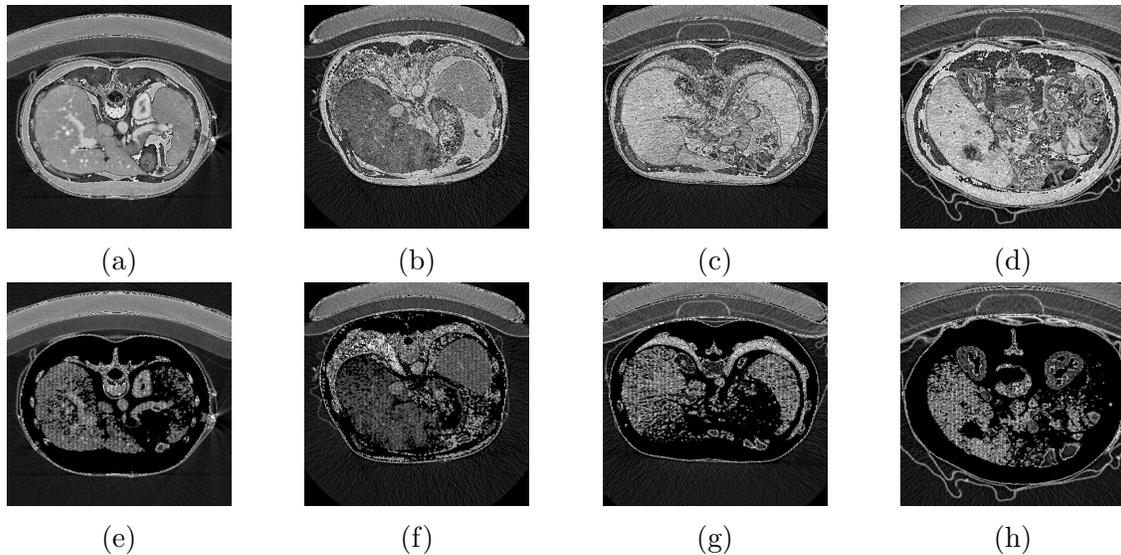
Figure 3.3: Skeletonized Slices

3.2.4 Alignment Algorithm

The Enhanced Correlation Coefficient (ECC) alignment algorithm is implemented to align the serially acquired slices of the 3D CT volume. This alignment algorithm was presented by Evangelidis et al [30]. They proposed the ECC similarity measure for estimating the parameters of a motion model. The motion model used in this implementation is the Euclidean motion model where the first image a rotated and shifted version of the second image. Using this approach has two advantages. The first is that the ECC similarity measure is invariant to photometric distortions in contrast and brightness which is unlike the traditional similarity measure of difference in pixel intensities. The second advantage is that even though the objective function is a nonlinear function, the iterative scheme used to solve the optimization problem is linear. Therefore, a problem that appears to be a computationally expensive isn't that complex due to the simplicity of the iterative solution.

The alignment algorithm presented here is a variation of the algorithm to cater for a large number of slices whereas, the algorithm above is for a simple alignment of two images. Since the number of slices that make up the CT volume is so large, the probability of an error being propagated through the volume is larger. Therefore, the consideration of such an aspect is significant. The algorithm doesn't iterate through the slices of the volume from one end to the other. Slices are randomly selected for alignment to reduce the chance of an error propagation. Another aspect considered is the variance of anatomical structures between the slices. The inter-slice distance is also a contributing factor to the variance. Furthermore, neighbouring slices do not overlap. Therefore, the slices are aligned with the neighbouring slices and not slices that are too far away within the volume because the anatomical structures in those slices compared to the chosen slice are different i.e. slice five presents different anatomical structures compared slice ninety.

48

Three different implementations of the ECC alignment algorithm is presented and their performance are compared. All implementations take in a list of 2D slices that have been skeletonized. The first implementation iterates through the list aligning all images to the first image in the list. The algorithm is shown in Algorithm 3. The first slice is set as the reference image and then the list is iterated selecting each image and setting it as the floating image. The iteration starts from the second slice and a warped matrix is estimated using the ECC algorithm to align it to the reference image. Lastly, the warped matrix is applied to the floating image aligning it to the reference image. The second implementation is similar to that of the first one. The algorithm iterates through the list of slices and estimates a warped matrix to be applied to each slice. However, the reference image is the slice preceding the floating image. This algorithm is presented in Algorithm 4.

Lastly, the third implementation can be seen in Algorithm 5. In the last algorithm, all slices are declared unvisited. A slice is randomly selected and if it's visited then another random slice is selected. The selected slice becomes the floating image to be aligned and its neighbouring slices, one slice before and one slice after, are put into a list. If the selected slice is the last slice, the the neighbours used are the two slices preceding it and if the first slice is selected then the neighbours used are the two slices after it. The list of neighbours are used as a series of reference images to estimate a warped matrix for the floating image. the warped matrix is updated each time after each neighbour is processed and thereafter, the floating image is aligned to its neighbours by applying the warped matrix to it. Lastly, the slice is declared visited. This process continues until all slices are declared visited.

Algorithm 3 Alignment Algorithm 1

   Input: list of 2D slices
   Output: list of aligned 2D slices
  $reference \leftarrow slice[0]$
  for each slice in $list$ **do**
    $floatingImage \leftarrow slice[i]$
    $motionModel \leftarrow euclideanMotion$
    $warpMatrix \leftarrow newMatrix$
    $terminationCriteria \leftarrow maxIterations$
    estimate $warpMatrix$ using ECC algorithm
    apply $warpMatrix$ to $floatingImage$
  end for

3.3 Conclusion

In this chapter, a slice alignment method for alignment of serially acquired 2D slices is presented and discussed. The method uses techniques such as a Laplacian weighted histogram to determine the optimal liver threshold, skeletonization to remove unwanted anatomical structures and an alignment algorithm that implements

Algorithm 4 Alignment Algorithm 2

Input: list of 2D slices
Output: list of aligned 2D slices
for each slice in *list* **do**
    $reference \leftarrow slice[i]$
    $floatingImage \leftarrow slice[i+1]$
    $motionModel \leftarrow euclideanMotion$
    $warpMatrix \leftarrow newMatrix$
    $terminationCriteria \leftarrow maxIterations$
    estimate $warpMatrix$ using ECC algorithm
    apply $warpMatrix$ to $floatingImage$
end for

Algorithm 5 Alignment Algorithm 3

Input: list of 2D slices
Output: list of aligned 2D slices
$slices = unvisited$
$floatingImage \leftarrow slices[random]$
$neighbours = newlist$
while $slices \leftarrow visited = false$ **do**
    $random \leftarrow unvisitedSlice$
    if $random = lastSlice$ **then**
        $neighbours \leftarrow slice[random - 1]$
        $neighbours \leftarrow slice[random - 2]$
    else if $random = firstSlice$ **then**
        $neighbours \leftarrow slice[random + 1]$
        $neighbours \leftarrow slice[random + 2]$
    else
        $neighbours \leftarrow slice[random - 1]$
        $neighbours \leftarrow slice[random + 1]$
    end if
    for each slice in $neighbours$ **do**
        $reference \leftarrow slices[0]$
        $motionModel \leftarrow euclideanMotion$
        $warpMatrix \leftarrow newMatrix$
        $terminationCriteria \leftarrow maxIterations$
        estimate $warpMatrix$ using ECC algorithm
        update $warpMatrix$
    end for
    apply $warpMatrix$ to $floatingImage$
    $floatingImage = visited$
end while

the ECC algorithm. The method takes into consideration the limitations of sequential slice alignment by implementing an algorithm that conducts the alignment of slices randomly. This method also ensures that no particular direction is privileged in the method avoiding any global offsets and error propagation.

The following chapter describes the liver region of interest detection method which implements CNNs to classify 2D slices in order to determine a region of interest for liver segmentation.

Chapter 4

Liver Region of Interest Detection

4.1 Introduction

The liver is located in the abdominal area of the human body and in order to aid in the segmentation of the liver from 3D CT scans, it is important to obtain a region of interest. Abdominal CT scans usually contain the chest and pelvis along with the abdominal area as the abdomen occurs between these regions. Obtaining a region of interest of the liver, in this case the abdomen, focuses further processing within the region of interest. This prevents further processing, which includes liver segmentation, from being affected by the regions outside the region of interest. Localizing further processing increases the chance of higher accuracy as well as reduces the computation time due to the reduced search space for the liver.

Taking these factors into consideration, a method that utilizes convolutional neural networks is used to obtain a region of interest for liver segmentation. This method uses minimal pre-processing as CNNs have the ability to learn features from scratch. The CNN is first trained to classify slices into pelvic slices and abdominal slices. Then it is trained to classify the abdominal slices from the previous classification results into abdomen and chest slices. Minimal post-processing is used where the pelvic and chest slices are discarded. The remaining slices are the abdominal slices which is the resulting region of interest.

Furthermore, two different CNNs are implemented and trained to perform the slice classifications as described above. One CNN is a small network with a few layers and the other is a deeper network with many layers where a filter size is repeated for more than one layer. These networks are trained and tested on the same dataset in order to determine which network performs better and can be used prior to the liver segmentation step. In this chapter the region of interest detection method is presented.

The CNNs used as well as their training and data augmentation are discussed in Section 4.2. The methodology of the region of interest detection process is described in Section 4.3 which includes the pre-processing, pelvic girdle slices detection, chest slices detection and post-processing.

4.2   Convolutional Neural Networks

For this binary classification task, two convolutional neural networks are implemented. The capacity of convolutional neural networks varies depending on the number of layers. The higher the number of layers the network consists of the higher the level of features it will be able to capture. Heeneman [46] proposed that deeper networks perform better and the repeating filters of the same size also improves the accuracy of detection. Two networks are tested to determine whether the deeper network performs better in this case of classification. The model that performed the best and produced the most accurate results was used to obtain the region of interest for liver segmentation.

4.2.1   *Small* Network

The first network has one input feature map of size 128 x 128, two convolution layers, two max-pooling layers and two dense layers as well as one softmax layer. Each convolution layer is followed by a rectified linear unit to expedite the training of the network.

   The first convolutional layer contains 16 feature maps and each feature map is linked to the input feature maps through filters of size 5 x 5. Then a max-pooling layer with a stride size of two pixels are used to generate 32 feature maps. This is then followed by three convolutional layers that contain 32 feature maps each and each feature map is linked to all of the feature maps in the previous layer by filters of size 3 x 3. Then a max-pooling layer with a stride of two pixels are used to generate 64 feature maps. This is then followed by a dense layer with 500 units of neurons and a dropout of 0.5 is used. The output layer is a dense layer with 2 units of neurons and a softmax activation function. The architecture is described in Table 4.1.

Table 4.1: *Small* Network

Layer	Input	Filter	Stride	Output
Conv	128x128x1	5x5, 16	1x1	128x128, 16
Pooling	128x128,16	2x2	2x2	64x64, 16
Conv	64x64, 16	3x3, 32	1x1	64x64, 32
Pooling	64x64, 32	2x2	2x2	32x32, 32
Dense	32x32, 32			500
Dense	500			2

4.2.2   *Deep* Network

The second network has one input feature map of size 128 x 128, seven convolution layers, four max-pooling layers and two dense layers as well as one softmax layer.

Each convolution layer is followed by a rectified linear unit to expedite the training of the network.

The first convolutional layer contains 16 feature maps and each feature map is linked to the input feature maps through filters of size 5 x 5. Then a max-pooling layer with a stride size of two pixels are used to generate 32 feature maps. This is then followed by three convolutional layers that contain 32 feature maps each and each feature map is linked to all of the feature maps in the previous layer by filters of size 3 x 3. Thereafter, a max-pooling layer with a stride of two pixels are used to generate 64 feature maps. This is followed by two convolutional layers with 64 feature maps each where each feature map is linked to the feature maps in the previous layer by 3 x 3 filters. This is followed by a max-pooling layer with a stride size of two pixels are used to generate 128 feature maps. The last convolutional layer has 128 feature maps where each feature map is linked to the features in the previous layer by 1 x 1 filters. This is then followed by a dense layer with 500 units of neurons and a dropout of 0.5 is used. The output layer is a dense layer with 2 units of neurons and a softmax activation function. The architecture is described in Table 4.2.

Table 4.2: *Deep* Network

Layer	Input	Filter	Stride	Output
Conv	128x128, 1	5x5, 16	1x1	128x128, 16
Pooling	128x128,16	2x2	2x2	64x64, 16
Conv	64x64, 16	3x3, 32	1x1	64x64, 32
Conv	64x64, 16	3x3, 32	1x1	64x64, 32
Conv	64x64, 16	3x3, 32	1x1	64x64, 32
Pooling	64x64, 32	2x2	2x2	32x32, 32
Conv	64x64, 16	3x3, 64	1x1	64x64, 64
Conv	64x64, 16	3x3, 64	1x1	64x64, 64
Conv	64x64, 16	3x3, 128	1x1	64x64, 128
Pooling	64x64, 128	2x2	2x2	32x32, 128
Dense	32x32, 32			500
Dense	500			2

4.3   Data Augmentation

Data augmentation is a useful way to increase the volume of training images using only information within the training data, thereby reducing over-fitting on a model [104]. In this research, affine transformations of the original images are represented by Equation 4.1.

$$y = Wx + b \qquad (4.1)$$

Where y is the transformed image, x is the original image, b represents shift in height, rotation angle or shear factor and W represents the zoom range.

The transformations used are described as follows:

- Width shift = 0.1

- Height shift = 0.1

- Shear range = 0.2

- Zoom range = 0.2

- Rotation = 45 degrees

4.4 Training

Before using the network to classify the 2D slices, the network is trained using the cases in the training set. The training of the convolutional neural network was done in Python by using the Keras package with Tensorflow as back-end. Updating the weights was done after each epoch using batch-wise training with a batch size of 32 examples. For the loss function binary cross-entropy is used since it is a binary classification problem. The network parameters are trained using the adaptive gradient descent algorithm Adam optimization was used to update the weights. The initial learning rate is set to $x = 1e - 3$, a learning rate scheduler was used to adjust the learning rate. Metrics to evaluate the model during the training process are accuracy and error rate. After training the probability of the image being classified as either of the two classes is learned by the network and can be used to classify 2D slices.

4.5 Methodology

4.5.1 Pre-processing

Convolutional neural networks have the ability to learn features from scratch therefore only minimal amount of pre-processing is applied [52]. The image intensity of all CT scans was truncated to the range [-200, 250] HU to remove irrelevant details. The image slices are then resampled to size 128 x 128. Examples of slices before and after pre-processing are depicted in Figure 4.1 where Figures 4.1(a), 4.1(b), 4.1(c) and 4.1(d) show the images before pre-processing and 4.1(e), 4.1(f), 4.1(g) and 4.1(h) are their respective images after pre-processing.
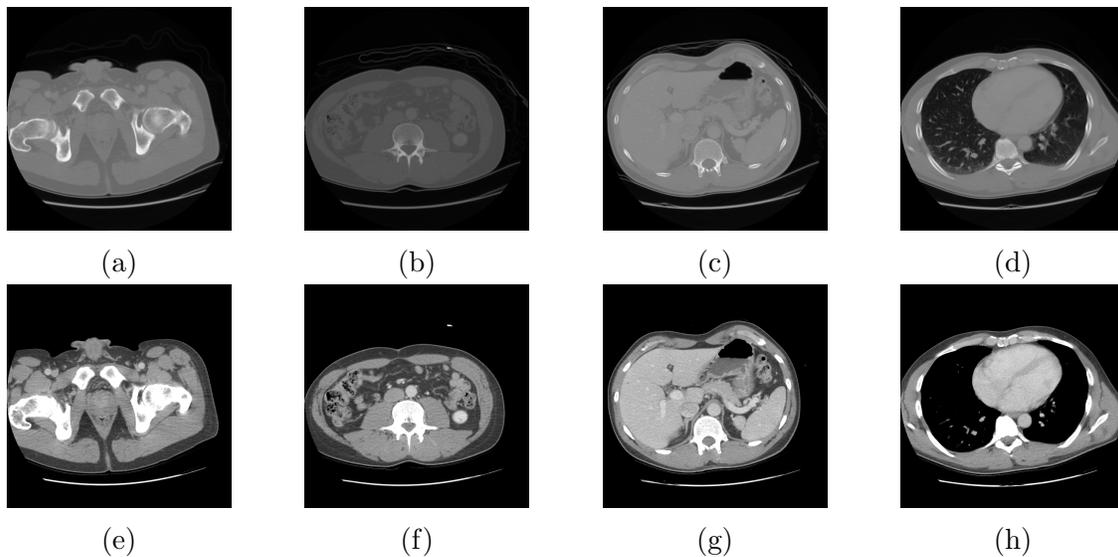
Figure 4.1: Slices before and after pre-processing

4.5.2 Classification of Abdominal and Pelvic Slices

The convolutional neural network was first trained to produce a different model that classifies 2D slices as either an abdominal slice or a slice containing the pelvic gridle. For this classification the network is trained for 20 epochs which took approximately 4.7 hours. The number of iterations for each epoch was 500. Once the training is complete, the network takes a 2D CT slice in the form of a 128 x 128 image and produces the probability of the image being an abdominal slice and the probability of the image being a slice of the pelvis. This classification model takes in an image that has undergone thresholding so that the classification is based on the bone structure that appears in the image. Thresholding is represented by Equation 4.2.

$$g(x,y) = \begin{cases} 0 & if f(x,y) \leq T \\ 1 & if f(x,y) > T \end{cases} \tag{4.2}$$

Where T is the threshold value, which in this case is 200 as this was found to be the optimal value to remove most soft tissue allowing the bone structure to be visible for detection.

Examples of the thresholding results can be seen in Figure 4.2.

(a)	(b)	(c)

(d)	(e)	(f)
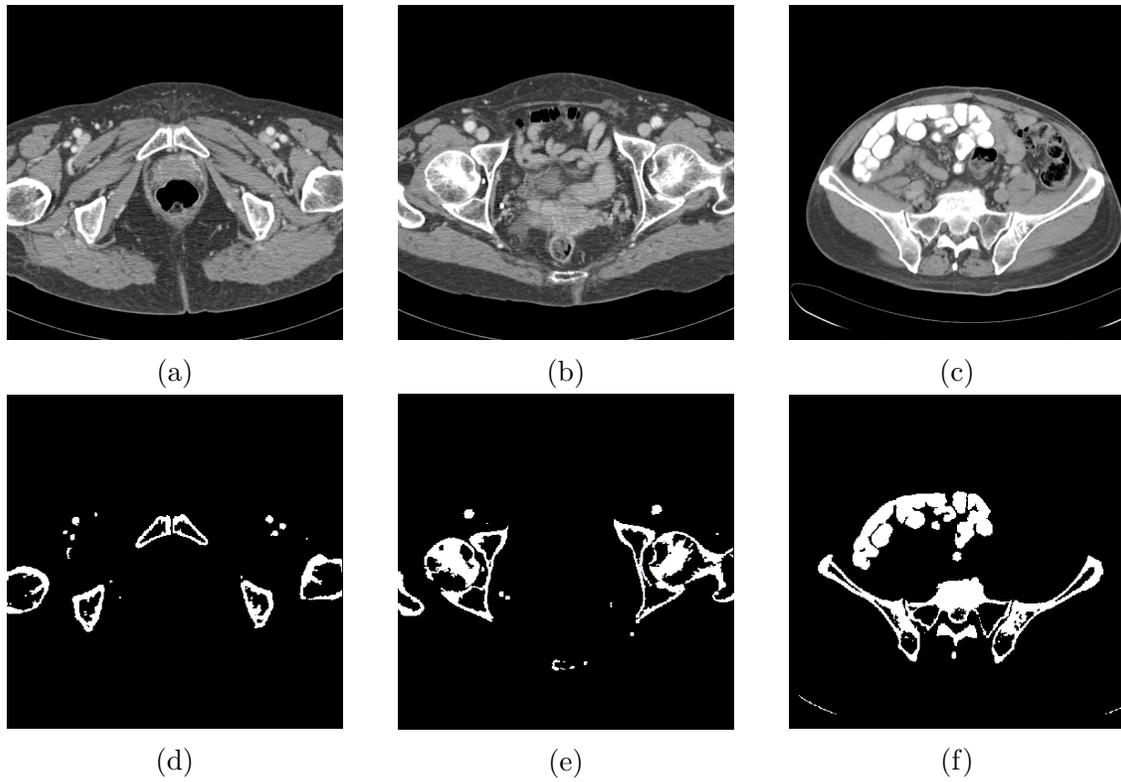
Figure 4.2: Results after thresholding of pelvis slices

The images containing the evidence of the pelvic girdle will be classified as a pelvic slice and those that don't are classified as an abdominal slice. The slice is categorized as the class with the higher probability. Examples of pelvis slices can be seen in Figure 4.3.

(a)

(b)

(c)

(d)
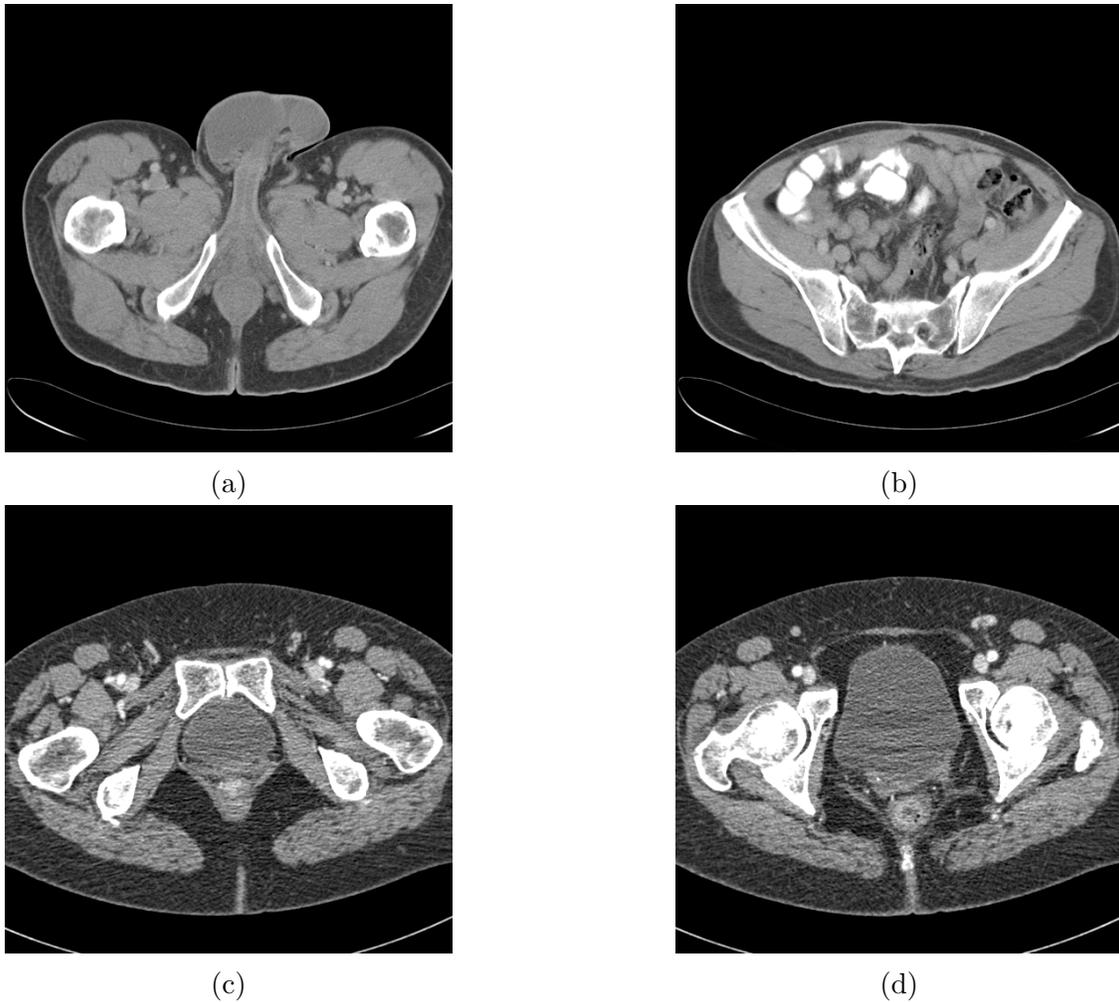
Figure 4.3: Pelvis Slices from CT Scan

4.5.3   Classification of Abdominal and Chest Slices

The network was trained again to produce a model that classifies 2D slices as either an abdominal slice or a chest slice. The network was trained for 25 epochs which took a duration of approximately 5 hours. After the training, the network takes a 2D CT slice in the form of a 128 x 128 image and produces the probability of the image being an abdominal slice as well as the probability of the image being a slice containing the chest cavity. This classification is based on the anatomical structures that are soft tissues and organs. The images containing the chest cavity and shoulders are classified as chest slices and the slices containing the organs and soft tissue below are classified as abdominal slices. The slice is categorized as the class with the higher probability. Examples of chest slices can be seen in Figure 4.4.
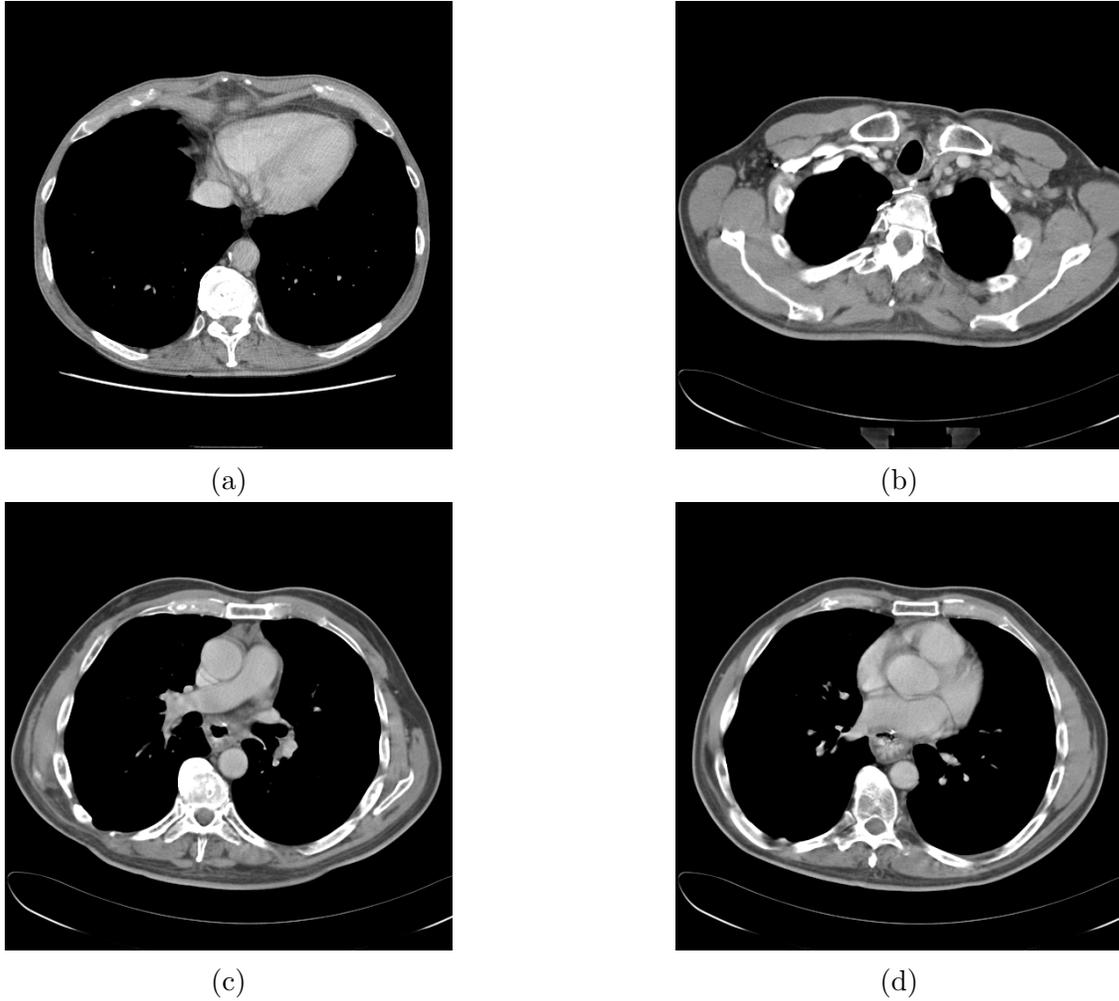
Figure 4.4: Chest Slices from CT Scan

4.5.4 Post-processing

The class of each slice predicted by the convolutional neural networks is stored in an array. The 2D slices that have been classified as pelvic slices and chest slices are removed and the abdominal slices make up the region of interest for further processing which is liver segmentation as the liver occurs within the abdomen. Algorithm 6 briefly describes the entire region of interest process. Examples of abdominal slices can be seen in Figure 4.5.

Algorithm 6 Region of Interest Detection

  Input: list of 2D slices
  Output: list of slices containing region of interest

$regionOfInterest \leftarrow volume$
for each $slice$ in $volume$ **do**
  if $slice = pelvisslice$ **then**
    remove $slice$ from $regionOfInterest$
  else if $slice = chestslice$ **then**
    remove $slice$ from $regionOfInterest$
  else
    $slice \leftarrow abdominalslice$
  end if
end for

(a)

(b)
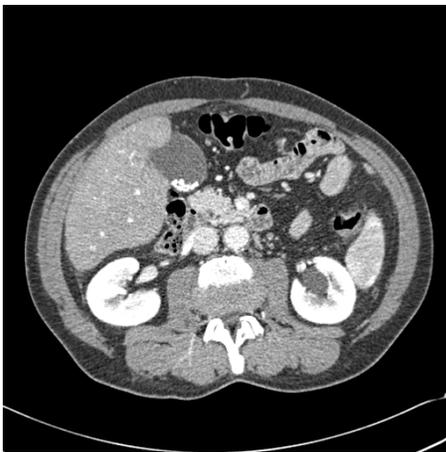
(c)
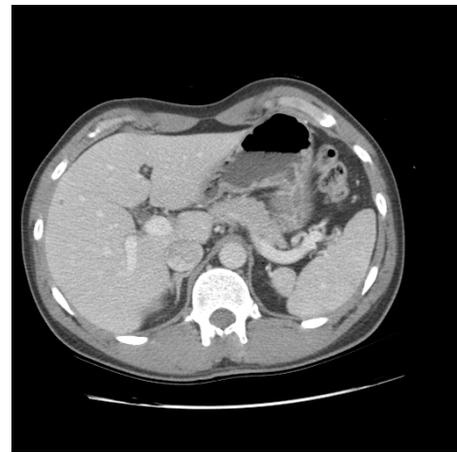
(d)

Figure 4.5: Abdominal Slices from CT Scan

4.6   Conclusion

In this chapter the implementation of a method to obtain a region of interest has been discussed. Due to the performance of the deeper network with a filter size of 3 x 3 being repeated through a few consecutive layers, this network was chosen to be used as the CNN used to obtain a region of interest. Furthermore, this is a new method used for obtaining a region of interest for liver segmentation. Previous methods use bounding box extraction, thresholding or manually deleting slices that do not contain liver tissue. This method is independent of grayscale image analysis and features are automatically learned from scratch so that unnecessary slices are removed from further processing.

In the next chapter, a liver segmentation method is presented using the results of this section as input. The various steps of locating and segmenting the liver are discussed in detail.

Chapter 5

Liver Segmentation

5.1 Introduction

The liver location and segmentation is the most important step of the method. If the liver is correctly located then subsequent steps are far more likely to yield good results. The main difficulty is deciding on the correct technique from those discussed previously. It is important to take into consideration that there is low contrast between the liver and its surrounding organs, the great differences in liver shape and appearance, the heterogeneity of background anatomical structures as well as inhomogeneous appearances due to tumours and other pathologies [72]. Another vital aspect to be considered is that since this methodology caters for 2D slices of the CT, the liver shape will be different on each slice, for example, the slices that occur at the center of the liver will show a large area of liver whereas the slices occurring at the ends of the liver will show a smaller area of the liver.

After taking these factors into account, a convolutional neural network is implemented to locate and segment the liver from each 2D slice. This is due to the ability of convolutional neural networks to learn where the liver is located and segment liver tissue [52]. The slice classification technique that precedes this step removes the slices that have a low probability of containing liver tissue which makes the task of the technique implemented in this chapter easier. In this chapter the liver location and segmentation method is presented as well as the pre-processing and post processing steps implemented to maximize its efficiency.

The use of concatenate layers in the network and its effect on liver segmentation is also investigated in this chapter. Two concatenate layers are added among the convolution layers in order to test its effectiveness on the segmentation of the liver to determine whether this addition improves the results produced by the convolutional neural network.

Figure 5.1 depicts an overview of the methodology for liver segmentation. The pre-processing techniques used are described in Section 5.2. The convolutional neural network is detailed in Section 5.3 as well as the training and data augmentation involved. The post-processing technique used is described in Section 5.4.
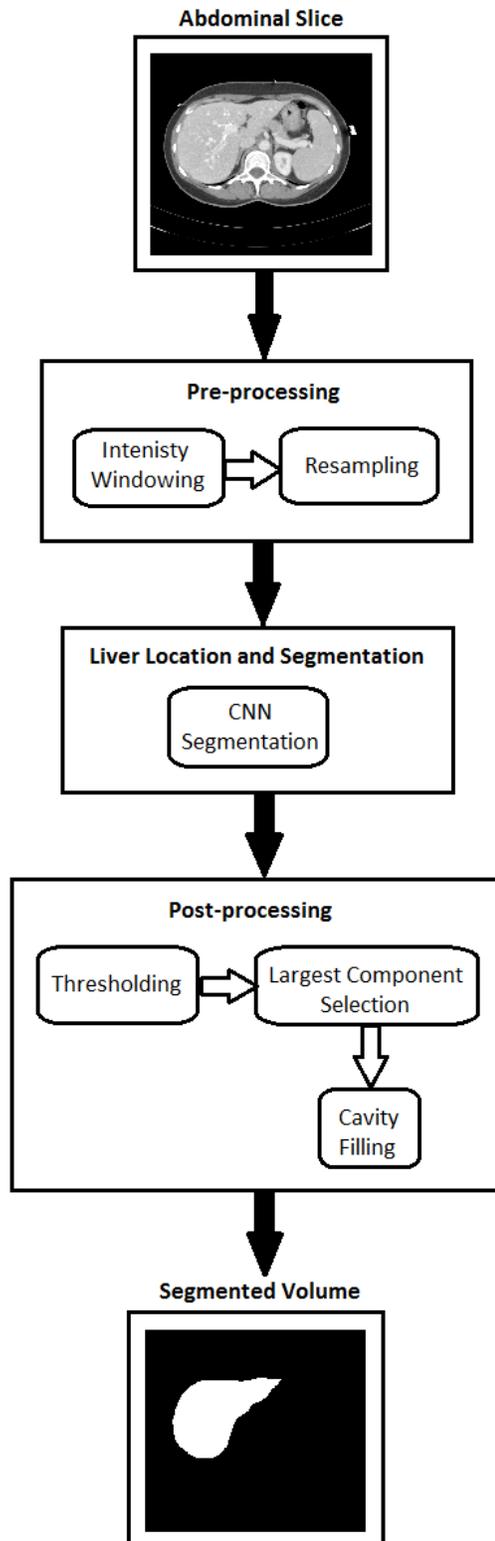
Figure 5.1: Overview of Liver Segmentation Method

5.2   Pre-processing

Convolutional neural networks have the ability to learn features from scratch therefore only minimal amount of pre-processing is applied [52]. The image intensity of all CT scans was truncated to the range [-200, 250] HU to remove irrelevant details. The image slices are then resampled to size 128 x 128. Examples of slices before and after pre-processing are depicted in Figure 5.2 where Figures 5.2(a), 5.2(b), 5.2(c) and 5.2(d) show the images before pre-processing and 5.2(e), 5.2(f), 5.2(g) and 5.2(h) are their respective images after pre-processing.

| (a) | (b) | (c) | (d) |

| (e) | (f) | (g) | (h) |
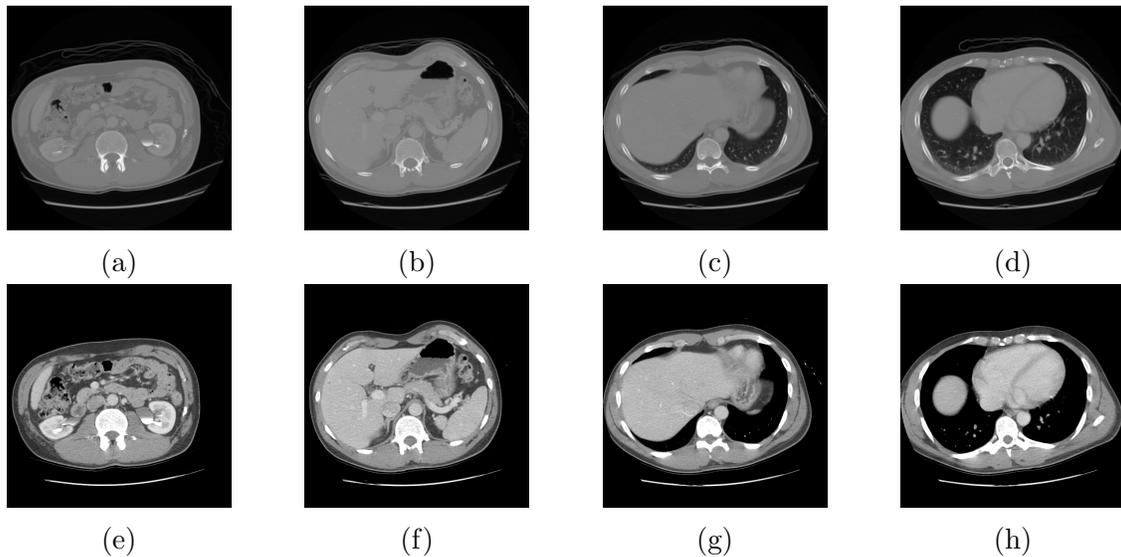
Figure 5.2: Slices before and after pre-processing

5.3   Liver Location and Segmentation

5.3.1   2D Convolutional Neural Networks

In this research, the 2D convolutional neural network is designed and trained to automatically detect the liver. The network takes a 2D slice as input and produces a probability map as a subject specific prior, which assigns each pixel the likelihood of being the liver for the target image. A CNN is a variation of multi-layer perceptron where several convolutional layers are stacked on top of each other. Each convolutional layer extracts feature maps from its preceding layer which is connected by some filters [69].

5.3.2 Network Architecture of Model 1 (CNN without concatenate layers)

Inspired by the work done using the U-Net architecture [84], a similar CNN structure is adopted. It consists of a contracting path and an expansive path where the contracting path follows the architecture of a typical convolutional network which consists of convolutional layers followed by max-pooling layers for down sampling [61]. The expansive path consists of upsampling followed by convolutions. To ensure seamless production of the output segmentation map it is vital that the input size is such that all max-pooling operations are applied to a layer with an even x- and y-size. The architecture of the CNN takes an input image of size 128 x 128 which corresponds to feature maps of a 2D slice of the CT volume. The components of this network include twelve convolution layers, two max-pooling layers, two up-sampling layers and a softmax layer. Each convolutional layer is followed by a rectified linear unit to expedite the training on the CNNs. The architecture is described in Table 5.1.

Table 5.1: Network architecture of Model 1

Layer	Input	Filter	Stride	Output
Conv	128x128x1	7x7, 32	1x1	128x128, 32
Pooling	128x128, 32	2x2	2x2	64x64, 32
Conv	64x64, 32	5x5, 64	1x1	64x64, 64
Pooling	64x64, 32	2x2	2x2	32x32, 32
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Upsampling	32x32, 128			64x64, 128
Conv	64x64, 128	3x3, 64	1x1	64x64, 64
Upsampling	64x64, 64			128x128, 64
Conv	128x128, 64	3x3, 32	1x1	128x128, 32
Conv	128x128, 32	1x1, 16	1x1	128x128, 16
Conv	128x128, 16	1x1, 1	1x1	128x128, 1
Logistic				

The first convolutional layer contains 32 feature maps and each feature map is linked to the input feature maps through filters of size 7 x 7. Then a max-pooling layer with a stride size of two pixels are used to generate 64 feature maps. The second convolutional layer contains 64 feature maps and each feature map is linked to all of the feature maps in the previous layer by filters of size 5 x 5. Following this a max-pooling layer with a stride of two pixels are used to generate 128 feature maps. This is followed by six convolutional layers with 128 feature maps where

each feature map is linked to the feature maps in the previous layer by 3 x 3 filters. Thereafter, an up-sampling layer with a stride size of two pixels generates 64 feature maps. This is followed by an upsampling layer which expands the size of the feature map to 128 feature maps. The ninth convolutional layer contains 64 feature maps and each feature map is linked to the previous layers feature maps by 3 x 3 filters. Then an up-sampling layer with a stride size of two pixels generates 64 feature maps. This is then followed by two more convolutional layers where the first one contains 32 feature maps and each feature is linked by 3 x 3 filters to the previous layers feature maps. The last convolutional layer has 16 feature maps where each feature map is linked to the features in the previous layer by 1 x 1 filters and a dropout of 0.5 is used. The output layer contains 1 feature map and is linked to the previous layers feature maps by 1 x 1 filters and a soft-max activation function is used here.

5.3.3 Network Architecture of Model 2 (With concatenate layers)

To investigate the effect of concatenation layers within the convolutional network, two concatenate layers were added among a series of convolution layers. Since there may be a loss of important features during the application of convolution filters, the concatenation layers allow for lost features to be reintroduced by concatenation. Concatenating layers that have different precision helps retrieving fine-grained spatial information, as well as coarse contextual information [6].

The architecture of the proposed CNN takes an input image of size 128 x 128 which corresponds to feature maps of a 2D slice of the CT volume. The components of this network include twelve convolution layers, two max-pooling layers, two up-sampling layers, two concatenate layers as well as a softmax layer. Each convolutional layer is followed by a rectified linear unit to expedite the training on the CNNs. The first concatenate layer is placed after the fourth convolution layer which concatenates the feature maps of the third and fourth convolution layers. The second concatenate layer is placed after the sixth convolution layer which concatenates the feature maps of the fifth and sixth convolution layers. The architecture is described in Table 5.2. The segmentation results obtained from the CNNs are presented in Figure 5.3 where Figure 5.3(a) shows the original image, Figure 5.3(b) shows the liver probability map produced by Model 1 and Figure 5.3(c) shows the liver probability map produced by Model 2.

Table 5.2: Network architecture of Model 2

Layer	Input	Filter	Stride	Output
Conv	128x128x1	7x7, 32	1x1	128x128, 32
Pooling	128x128, 32	2x2	2x2	64x64, 32
Conv	64x64, 32	5x5, 64	1x1	64x64, 64
Pooling	64x64, 32	2x2	2x2	32x32, 32
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Concat				
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Concat				
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Conv	32x32, 32	3x3, 128	1x1	32x32, 128
Upsampling	32x32, 128			64x64, 128
Conv	64x64, 128	3x3, 64	1x1	64x64, 64
Upsampling	64x64, 64			128x128, 64
Conv	128x128, 64	3x3, 32	1x1	128x128, 32
Conv	128x128, 32	1x1, 16	1x1	128x128, 16
Conv	128x128, 16	1x1, 1	1x1	128x128, 1
Logistic				

5.3.4   Data Augmentation

Data augmentation is a useful way to increase the volume of training images using only information within the training data, thereby reducing over-fitting on a model [104]. In this research, affine transformations of the original images are represented by Equation 5.1.

$$y = Wx + b \tag{5.1}$$

Where y is the transformed image, x is the original image, b represents shift in height or rotation angle and W represents the zoom range.

The transformations used are described as follows:

- Width shift = 0.1

- Height shift = 0.1

- Zoom range = 0.1

- Rotation = 10 degrees

(a) Original	(b) Result of Model1	(c) Result of Model2
(d) Original	(e) Result of Model1	(f) Result of Model2
(g) Original	(h) Result of Model1	(i) Result of Model2
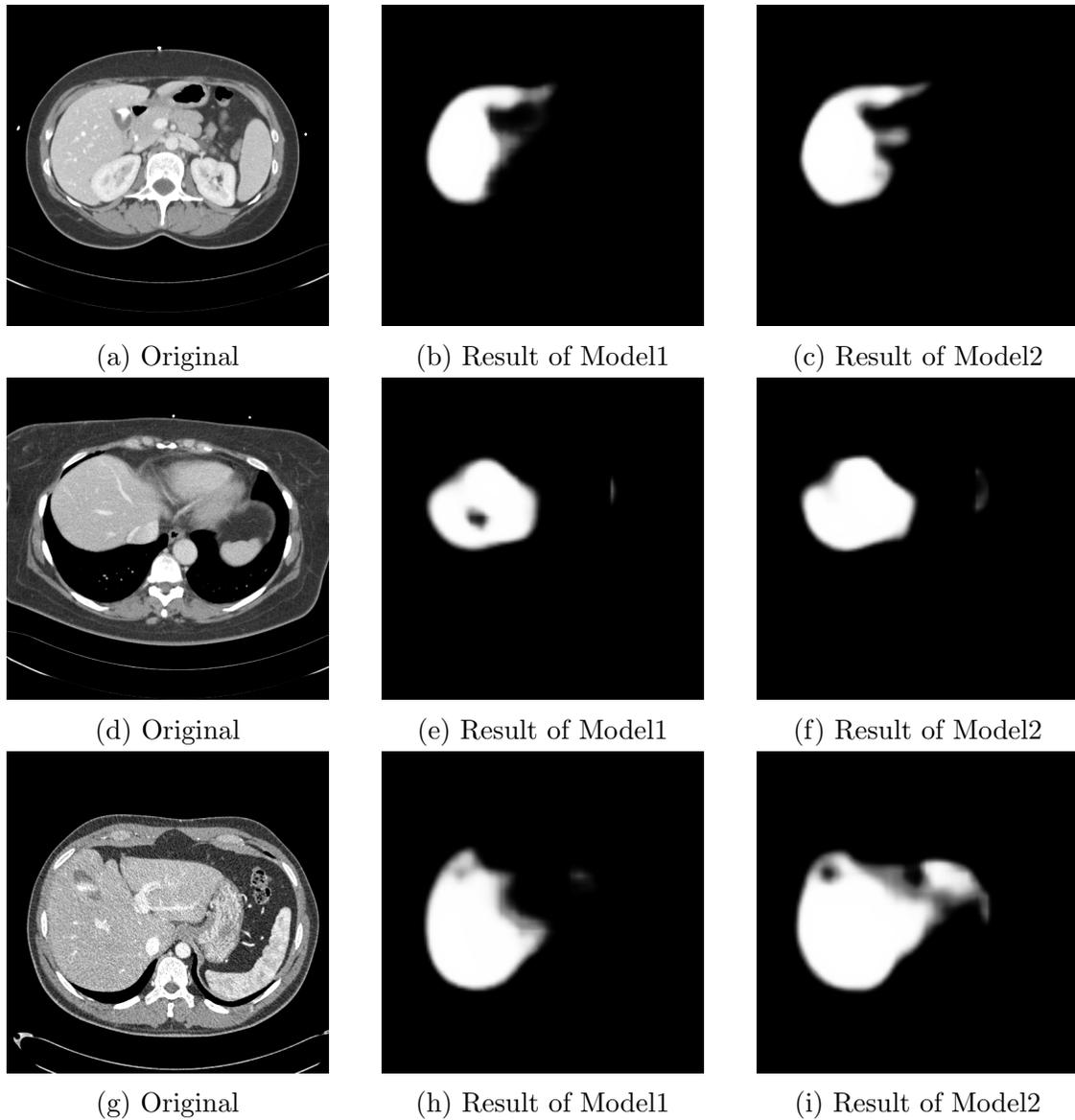
Figure 5.3: Liver location and segmentation results

5.3.5 Training

Prior to using the convolutional neural network to locate and segment the liver from the 2D CT slices, it needs to be trained using the cases in the training set. The networks were trained for 25 epochs with 1000 iterations per epoch to generate the liver likelihood map. During each iteration a 128 x 128 slice is randomly chosen from the training data and a 128 x 128 ground truth slice as the output. For the loss function binary cross-entropy is used since it is a binary classification problem [46]. The network parameters are trained using the adaptive gradient descent algorithm Adam optimization was used to update the weights. The learning rate is set to $x = 1e - 3 \times 0.8^x$ and reduced using a learning rate scheduler. The initial weights

are set using a random seed and the momentum and decay are adopted from the U-Net used in [84]. Updating the weights was done after each epoch using batch-wise training with a batch size of 16 examples. Assuming that the training dataset is made up n labeled samples:

$$\{(x^1, y^1), (x^2, y^2), ..., (x^n, y^n)\}, \; where \; y^i = 0 \; or \; 1, \; i = 1, 2, ..., n.$$

Denote θ be the set of all the parameters including the filters and softmax parameters of the CNNs. For logistic regression, the following cost function is minimized with respect to θ:

$$E(\theta) = \frac{1}{n} \left\{ \sum_{i=1}^{i=n} y^i log F_\theta(x^i) + (1 - y^i) log(1 - F_\theta(x^i)) \right\}. \qquad (5.2)$$

Training each network takes approximately 45 hours per network using a CPU. After the training is complete, the trained convolutional neural network is used to predict a liver likelihood map of a 2D slice thereby, locating and segmenting the liver.

5.4 Post-processing

Once the network produces a probability map of the liver likelihood, the result undergoes some simple post-processing. Post-processing is used to refine the segmentation results produced by the primary segmentation techniques. As discussed in Chapter 2, there are many techniques that have been used for post-processing. The refinement techniques used in this work include thresholding, largest connected component selection and cavity filling using morphological operations.

5.4.1 Thresholding

The technique used is binary thresholding to eliminate the pixels with a probability below 50 percent in the probability map. This produces an initial liver segmentation. The result of the thresholding can be seen in Figure 5.4 where 5.4(a) and 5.4(c) are images before thresholding and 5.4(b) and 5.4(d) are their respective results after thresholding. Equation 5.3 defines the thresholding function used.

$$g(x, y) = \begin{cases} 0 & if f(x, y) \le T \\ 1 & if f(x, y) > T \end{cases} \qquad (5.3)$$

Where T is the threshold value, which in this case is a pixel value 127 as this is 50% of the maximum pixel value of 255.

<center>(a)</center> <center>(b)</center>

<center>(c)</center> <center>(d)</center>
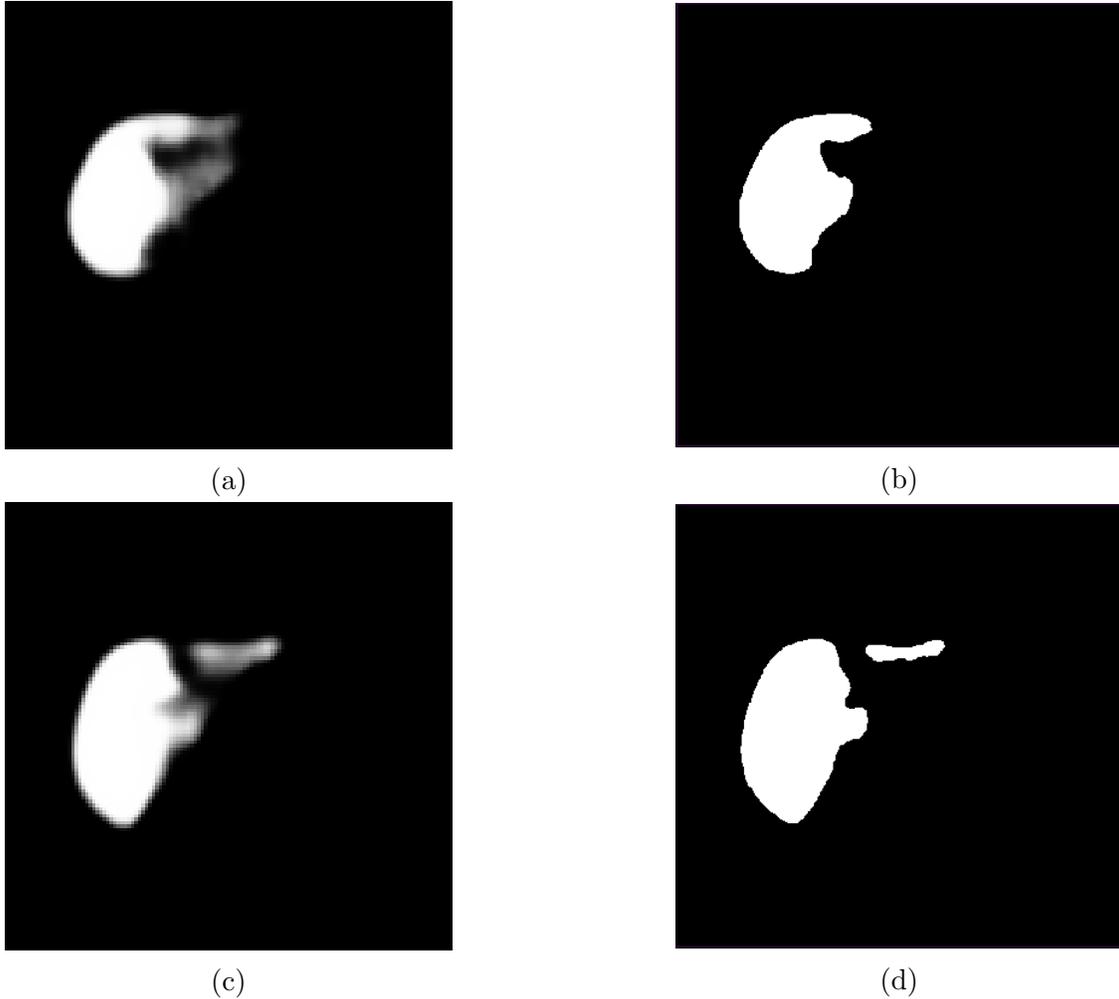
<center>Figure 5.4: Results after thresholding the outputs of the CNN</center>

5.4.2 Largest Component Selection

Since the liver is the largest organ present in the abdomen, largest connected component selection is a technique that has been used as a primary segmentation technique in anti-learning-based methodologies. Zayane et al [108] used a GCC algorithm to detect the largest connected component after performing thresholding. Li et al [65] used largest component detection after using a clustering method to initialize the level set method as well as after the level set produced a segmentation result. Rusko et al [86] and Campadelli et al [18] used this technique to limit a search space for seed selection to initialize a region growing method. Similar to the methodologies presented in [101] and [106], this technique is used for segmentation refinement or post-processing. In some cases, islands of misclassified pixels may occur in the segmentation result produced by the primary segmentation technique which in this case is the convolutional neural network described in the previous chapter.

The algorithm used accepts the resulting segmentation volume from the liver

<center>70</center>

location and segmentation method as input. The background pixel intensity is set to zero and each component that is not zero within the volume is labelled. The volume of each component is calculated and the component with the largest volume is selected. This volume is then set as the threshold volume and all components with a volume less than the threshold volume are removed using morphological operators. The algorithm used in this process is described in Algorithm 7. Morphological opening is used to remove the smaller components [102]. The result of the largest component selection can be seen in Figure 5.5 where 5.5(a) and 5.5(c) are images before largest component selection and 5.5(b) and 5.5(d) are their respective results after being processed.

Algorithm 7 Largest Component Selection

Input: segmentation volume produced by segmentation
Output: segmentation volume containing largest component

for each *component* in *components* **do**
   $area[i] \leftarrow calculateArea(component[i])$
end for
$largestComponent \leftarrow 0$
for each *area* in *areas* **do**
   if $area[i] > area[largestComponent]$ **then**
     $largestComponent \leftarrow i$
   end if
end for
$threshold \leftarrow area[largestComponent] - 1$
if length of $area > 1$ **then**
   for each *component* in *components* **do**
     if $area[component] < threshold$ **then**
       remove *component*
     end if
   end for
end if

(a)

(b)

(c)

(d)
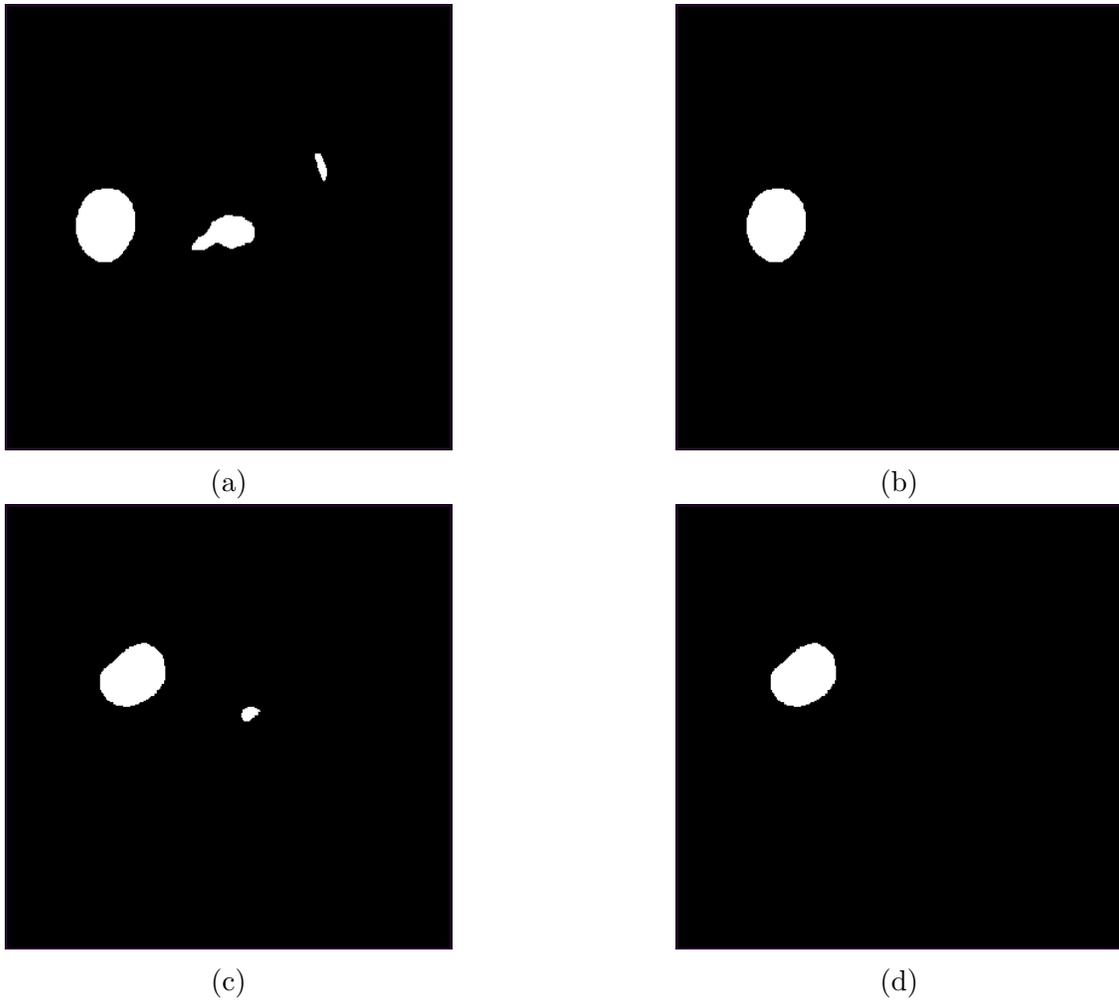
Figure 5.5: Results after largest component selection

5.4.3 Cavity Filling

In some cases, small cavities are present in the segmented liver. This may be as a result of inhomogeneous liver tissue where tumours or lesions are present. The most common post-processing technique used is morphological operators which are used in [52, 81, 94, 25, 37, 108, 103] for cavity filling, to smooth extracted surfaces and eliminate unconnected regions.

After the largest connected component is detected and the smaller components are removed, the resulting volume is processed using morphological operators to remove the cavities from the segmented volume. Morphological closing is used to fill the cavities [102]. The resulting volume is the final segmentation volume. The result of cavity filling can be seen in Figure 5.6 where 5.6(a) and 5.6(c) are images before cavity filling selection and 5.6(b) and 5.6(d) are their respective results after being processed.
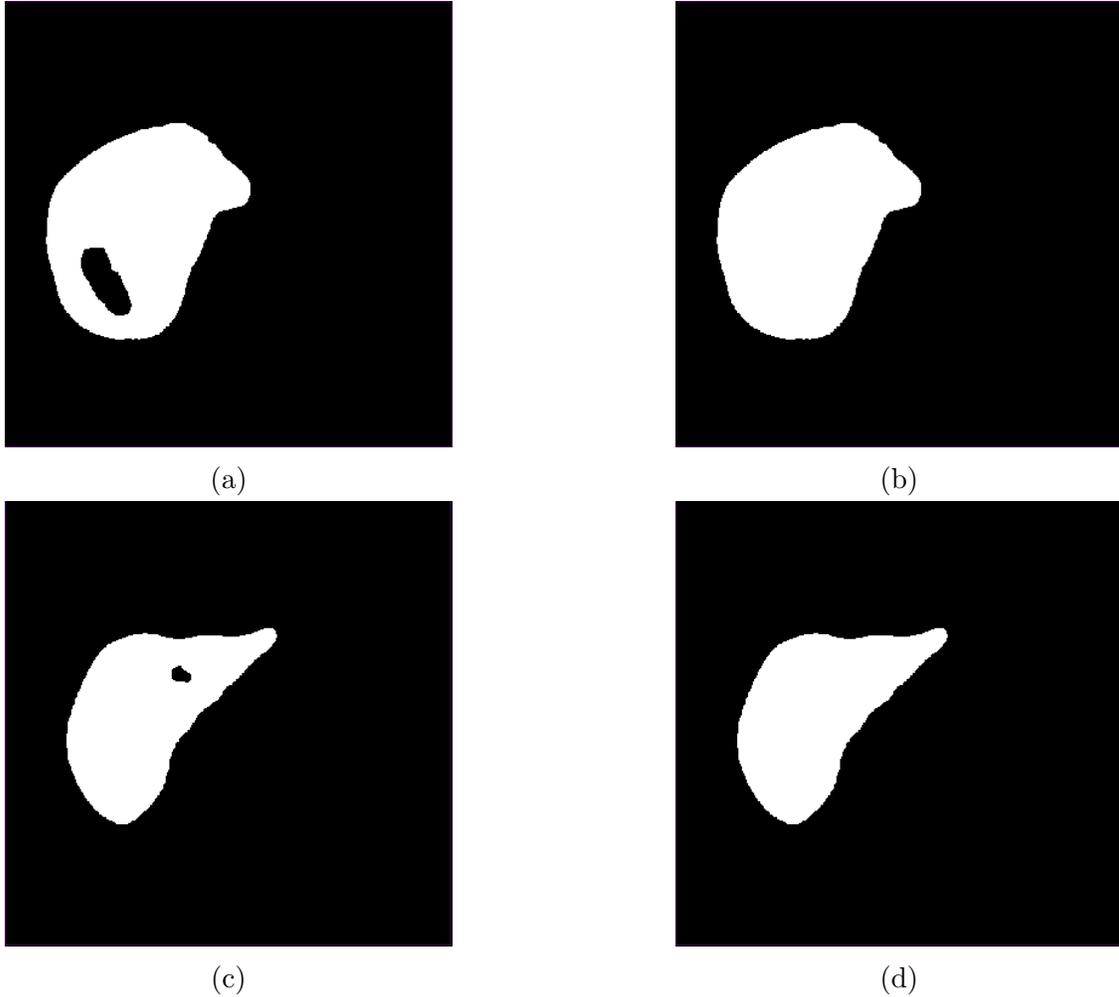
Figure 5.6: Results after cavity filling

5.5 Conclusion

In this chapter, the implementation of the liver location and segmentation is discussed. Minimal pre-processing allows for the neural network to learn features from scratch making it more efficient in detecting and segmenting the liver. The addition of concatenation layers to the convolutional network is also investigated.

Two networks are described; one without concatenate layers and one with concatenate layers. Both networks are trained with the same hyper-parameters and with the same training data and are tested with the same test data. The experiments and results of both networks are detailed in the next chapter. The use of concatenate layers may be useful in the detection and segmentation of the liver boundaries because they help in retrieving fine-grained spatial information, as well as coarse contextual information.

The use of the post-processing thresholding allows for the elimination of the

pixels that are less likely to be liver tissue which prevents the probability of over-segmentation around the liver boundaries. In order to deal with cases where small islands of misclassified pixels exist in the segmentation results produced by the liver location and segmentation method, the resulting volume undergoes segmentation refinement which includes largest component detection and morphological opening. Furthermore, the volume undergoes morphological closing to remove cavities within the liver volume which may have been misclassified by the convolutional neural network due to the inhomogeneous liver tissue. The final segmentation result is produced after this segmentation post-processing stage.

In the following chapter, the methods presented in this dissertation are evaluated and the results are presented and discussed.

Chapter 6

Results and Discussion

6.1 Introduction

In order to determine whether the proposed methods presented in this dissertation has performed successfully, the methods must be evaluated and analyzed. It is also important to understand the nature of the dataset and the programming environment used.

The results obtained by other proposed liver segmentation methods and techniques that were discussed in Chapter 2 are readily available, therefore comparisons between those results and the results obtained by the proposed method were conducted. In order to determine the performance of the proposed method, comparisons needed to be made against the ground truth segmentation obtained by manual segmentation conducted by a radiologist.

In this chapter, information pertaining to the dataset is presented in Section 6.2 and a brief description of the programming environment is given in Section 6.3. Section 6.4 gives an overview of the framework and Section 6.5 details the experiments done, evaluation metrics, results and discussion of the slice alignment method. Section 6.6 describes the experiments done, evaluation metrics, results and discussion of the liver region of interest detection method and Section 6.6 describes the experiments done, evaluation metrics, results and discussion of the liver location and segmentation method. Finally, Section 6.8 discusses the results and a comparative analysis of the method presented in this dissertation.

6.2 Dataset

The datasets used are the Medical Image Computing and Computer Assisted Intervention (MICCAI) 2007 grand challenge datasets [100]. All CT images are enhanced with contrast agent and scanned in the central venous phase on a variety of scanners (different manufacturers, 4, 16 and 64 detector rows). As it is CT, all datasets were

acquired in transversal direction. The pixel spacing varied between 0.55 and 0.80 mm and the inter-slice distance varied from 1 to 3 mm. This means that the distance between consecutive slices differ between scans. There is no overlap between neighboring slices. Examples can seen in the Figure 6.1, which shows consecutive slices with a larger inter-slice distance where the difference in anatomical structures are very conspicious, and Figure 6.2 which shows consecutive slices with a smaller inter-slice distance where the differences in anatomical structures are more subtle.
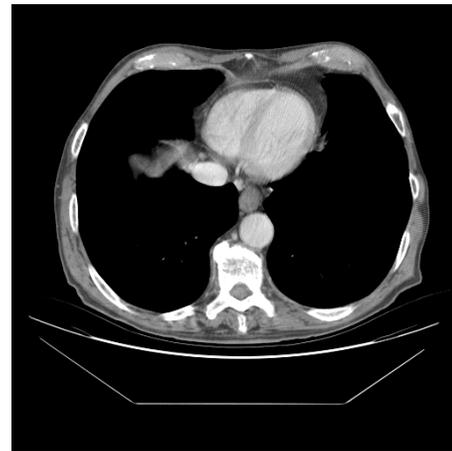
(a) (b)

Figure 6.1: Consecutive slices with a larger distance

(a) (b)

Figure 6.2: Consecutive slices with a smaller distance

Most of the images are pathologic and include tumors, metastasis and cysts of different sizes. Examples of these pathological abnormalities can be seen Figure 6.3. The dataset used for training is the MICCAI 2007 grand challenge training set (MICCAI-Training) which consists of 10 volume images with corresponding ground

truth segmentations. The dataset used for testing is the MICCAI 2007 grand challenge testing data (MICCAI-Testing) which consists of 10 volume images.

<div align="center">

(a)   (b)   (c)

</div>

Figure 6.3: Images with different pathological abnormalities

6.3 Programming Environment

The proposed framework was implemented and run with Intel Core i7-2640M CPU @ 2.80GHz with 16.0GB RAM. The methods were implemented in Python using different forms of plugins such as Tensorflow, Keras for the implementation of deep learning techniques presented in this dissertation as well as OpenCV.

6.4 Overview of Liver Segmentation Framework

In order to process the dataset, a three stage liver segmentation framework is proposed. The individual stages have been thoroughly discussed in the previous chapters but the overview is as follows:

- Region of Interest Detection

- Liver Location and Segmentation

Figure 6.4 depicts an overview of the framework and how these stages relate to each other and the processes that each stage involves.
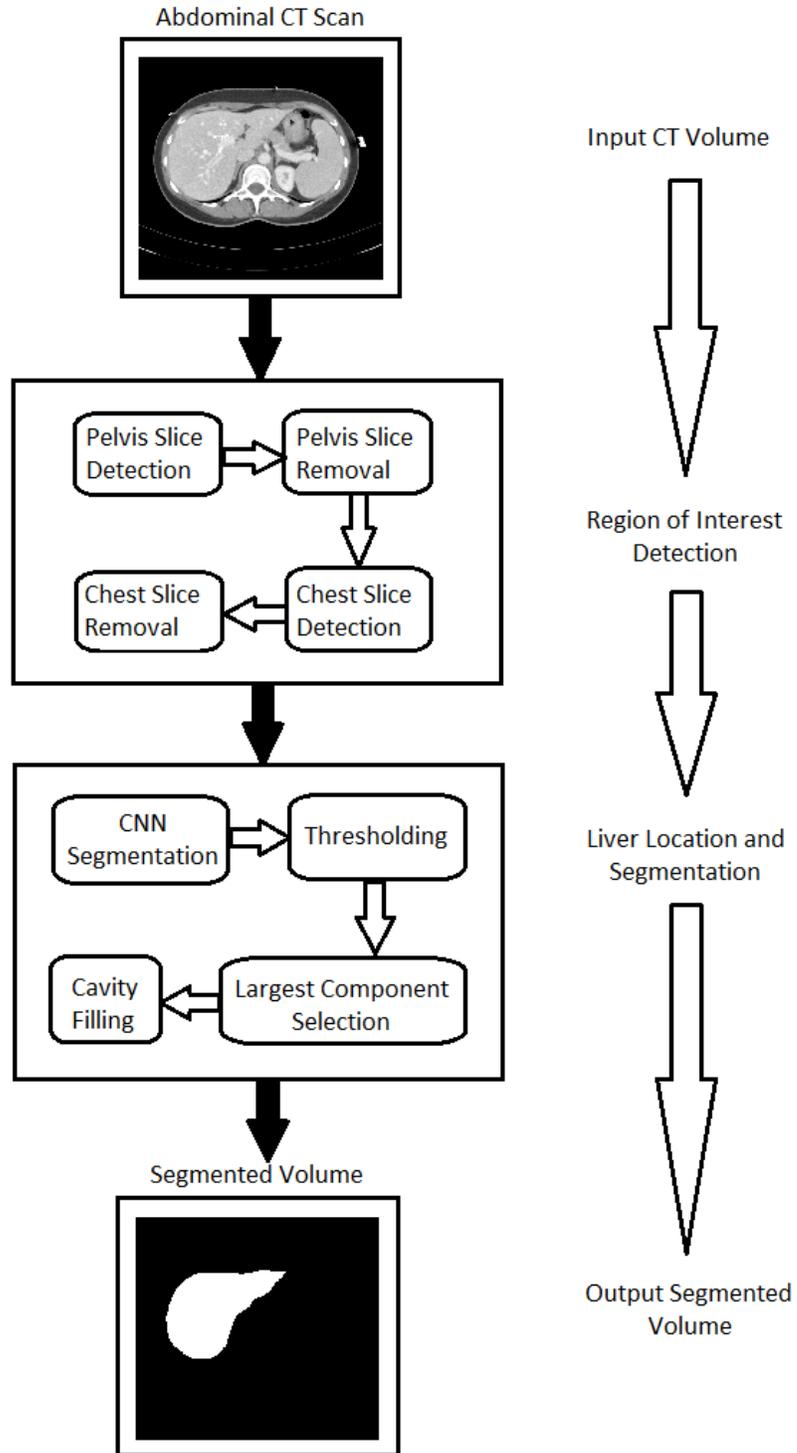
Figure 6.4: Overview of Liver Segmentation Framework

6.5 Results of Slice Alignment

6.5.1 Experimental Setup

Three different experiments were carried out and the results were compared. The first experiment implemented the ECC alignment algorithm such that each image was aligned with the first image in the series of 2D images. The second experiment implemented the ECC alignment algorithm such that each image in the series was aligned with the image preceding it and the images were aligned in order from the first image to the last. The final experiment included using the ECC algorithm to align each image with its neigbouring slices in both directions. The images were selected randomly to be aligned until all images were aligned. The results obtained from these experiments were compared with each other to determine which is the most appropriate for this dataset.

6.5.2 Evaluation

After the experiments were carried out the results were evaluated by calculating the mean square difference(MSD) of the entire volume after each slice was aligned [63]. The MSD obtained for the three experiments were compared for each 3D CT volume. MSD is defined in equation 6.1.

$$MSD = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} (I_{org}(i,j) - I_{int}(i,j))^2 \tag{6.1}$$

where m is the width of the image, n is the height of the image, $I_{org}(i,j)$ is the intensity value of the voxel at location (i,j) of the original image and $I_{alg}(i,j)$ is the intensity value of the voxel at location (i,j) of the aligned image.

6.5.3 Results and Discussion

The results obtained in the three experiments are shown in Table 6.1. The convergence of the ECC algorithm for experiment 1 is slower due to the difference in anatomical structures as the inter-slice distance grows with each iteration. The closer the floating slice to the reference slice the faster the convergence. Since each slice is being aligned to the first slice in the volume, the interslice distance increases with each iteration resulting in the convergence speed increasing with each iteration. In many cases, convergence is not reached. This is due to the algorithm not being able to estimate a warped matrix to align the floating image to the reference image which in this case is the first slice in the series. This occurs because of the difference between the reference image and the floating image. In some cases the slices were incorrectly aligned. An example of an original slice and an incorrectly aligned slice can be seen in Figure 6.5(c) and Figure 6.5(d) respectively. With the

Table 6.1: Results of Alignment Experiments

Image	Before Alignment	Experiment 1	Experiment 2	Experiment 3
001	280542.04	157067.35	164103.03	152353.00
002	703386.83	434681.84	569200.62	520375.20
003	791906.94	477509.61	658236.09	617086.95
004	470235.34	no convergence	371647.04	340319.72
005	543330.58	no convergence	366675.97	334012.21
006	480735.20	no convergence	348286.23	322766.54
007	505588.42	no convergence	356710.61	324468.96
008	202188.97	203107.31	134744.77	123075.58
009	426046.29	no convergence	280013.54	285622.30
010	1466966.96	no convergence	1121529.40	1058978.80
Total	5870927.57	1272366.10	4371174.30	4079059.20
Average	587092.76	127236.61	437117.43	407905.92

second experiment, the convergence occurs faster than in experiment 1 and 3 because each slice is being aligned with its preceding slice. However, the probability of errors being propagated through the entire volume is high as this is a sequential formulation. This method is not suitable for this particular dataset because the number of slices per volume is large. As can be seen in the results obtained, the MSD of the volume obtained by experiment 2 is significantly less than that of the volume before alignment. This proves that the alignment used in this experiment performs adequately and better than the alignment in experiment 1. The convergence is faster than in experiment 1 because the slices are being aligned with neighbouring slices in both directions i.e. each slice is aligned with the slice before it and after it. The algorithm also aligns slices in a random fashion. Due to this, no particular direction is privileged in the method avoiding any global offsets, biases in the estimation and error propagation. It is evident in the results obtained that this alignment algorithm preforms better than those in experiments 1 and 2 as the MSD is less than those obtained in the previous experiments. An example of a correctly aligned slice, before and after alignment, can be seen in Figure 6.5(a) and Figure 6.5(b) respectively.

(a)   (b)   (c)   (d)
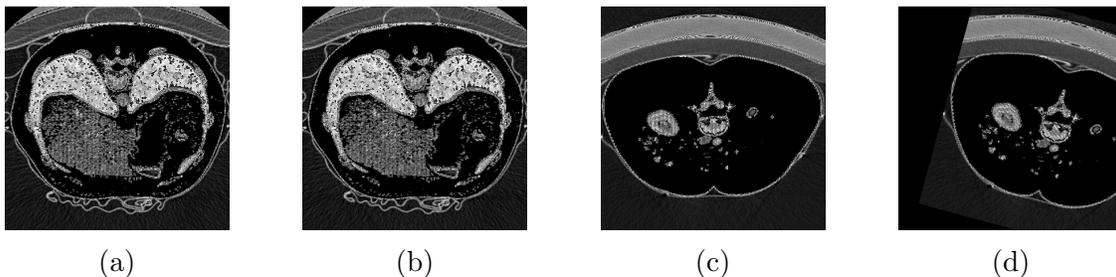
Figure 6.5: Aligned Slices

The change in MSD achieved in each experiment for each image volume can

Table 6.2: Change in MSD

Image	Experiment 1	Experiment 2	Experiment 3
001	44.01	41.51	45.69
002	38.20	19.07	26.01
003	39.70	16.88	22.08
004	-	20.97	27.63
005	-	32.51	38.53
006	-	27.55	32.86
007	-	29.45	35.82
008	-	33.36	39.13
009	-	34.28	32.96
010	-	23.55	27.81
Average	12.15	27.91	32.85

Table 6.3: Comparison of Methods

Method	Registration improvement
Fischer et al [35]	40%
Collins et al [23]	20%
Proposed	**32.85**%

be seen in Table 6.2. The average change in MSD obtained by the three image volumes that did reach convergence is 12.15%. It is evident that this algorithm is not efficient or approriate for such large volumes. An average change in MSD of 27.91% is achieved by the algorithm in experiment two and the randomized algorithm in experiment 3 achieved an average of 32.85%. This proves that the randomized algorithm performed the best.

The results obtained in experiment 3 is compared to those achieved by Collins et al [23] and Fischer et al [35]. This is presented in Table 6.3. The proposed algorithm achieved a higher improvement in alignment compared to that of Collins but lower than that of Fischer. However, the evaluation of the work presented by Fischer was evaluated on a very small sample and only two images are registered at a time. The algorithm here aligns a series of images. The method presented by Collins is tailored specifically to align or register 3D volumes of the brain with respect to sulcal tissue. The advantage of the method proposed in this research can be adjusted for use for any medical image registration or alignment simply by adjusting the parameters of the ECC alignment algorithm and the threshold value used by the skeletonization stage.

6.6 Results of Region of Interest Detection

6.6.1 Experimental Setup

An experiment was set up to test the methodology in a scientific manner. The experiment consists of two parts; evaluation and comparison of the small and deep networks for chest slice detection and the evaluation and comparison of the small and deep networks for the pelvis slice detection. This was done to determine which convolutional network would perform better to be used in this methodology. Thereafter, the methods are compared to related literature.

From the dataset described above, the training set was split up so that 80% of the images were used to train the networks and 20% was used for validation during the training. The test set was kept completely out of the training process. The training and validation ratio of 80:20 was based on the small size of the dataset. The model needs enough data to be able to generalize well as well as an adequate amount of data to validate. Through data analysis it was evident that there was variance throughout the training, validation and test sets therefore, the choice of split of the dataset is reasonable.

6.6.2 Evaluation

The four trained models are evaluated during the training and after training. During the training of a convolutional neural network classifier, the two types of errors that are found are training error and validation error. The training error occurs during training and is also known as loss which is desired to be minimized. The second type of error is validation error which is made during validation while testing the model on the validation set.

Other than error rates, accuracy, recall and precision are also used as evaluation metrics [46]. Accuracy is described as the amount of true positive (TP) classified examples and true negative (TN) classified examples relative to the total number of cases. Accuracy is defined in equation 6.2

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6.2}$$

Recall is the number of TPs relative to all true positive cases, including missed examples, which are called false negatives (FN). Recall is defined in equation 6.3

$$Recall = \frac{TP}{TP + FN} \tag{6.3}$$

Precision is defined as the number of true positives divided by the number of true positives plus the number of false positives. Precision is defined in equation

6.4

$$Precisionl = \frac{TP}{TP + FP} \tag{6.4}$$

The neural network is tested on 2D slices of 3D CT volumes. The number of slices per CT volume varies from scan to scan and the number of slices in each class per scan also varies. Each slice can either be a TP, FP, TN, or FN depending on its classification and its true class.

After the slices are classified by the model the accuracy, recall and precision are calculated. When the slices of all the CT scans have been evaluated the average scores are calculated and serve as the final scores for the models.

6.6.3 Results and Discussion

Small Network

The results are presented in the order of evaluation. The results during training are shown in Figure 6.6(a) for the chest slice detection model and Figure 6.6(b) for the pelvis slice detection model. Figure 6.6(a) shows that the model converged around epoch 20 where the purple line representing validation loss and the black line representing training loss converges and the red line representing training accuracy and the blue line represenitng validation accuracy converges. Similarly, in Figure 6.6(b)the convergence of the model occurs around epoch 11. The end results of the training of both models are shown in Table 6.4 and Table 6.5. For the chest slice detection model training, the loss is 0.0321 and the accuracy is 0.9791. The loss of validation is 0.0274 and accuracy is 0.9840. The training loss for the pelvis slice detection model is 0.0300 and the accuracy is 0.9845. The validation loss is 0.0190 and the accuracy is 0.9922.

Figure 6.6: Training loss and accuracy of *small* networks during training

Table 6.4: Results during training of *Small* Network for Chest Slice Detection

Metric	Training		Validation	
	Loss	**Accuracy**	**Loss**	**Accuracy**
Score	0.0321	0.9791	0.0274	0.9840

Table 6.5: Results during training of *Small* Network for Pelvis Slice Detection

Metric	Training		Validation	
	Loss	**Accuracy**	**Loss**	**Accuracy**
Score	0.0300	0.9845	0.0190	0.9922

The results after training are shown in Table 6.6 and Table 6.7 for the pelvis slice detection model and Table 6.8 and Table 6.9 for the chest slice detection model. These results show the recall, precision and accuracy after testing the models on the CT scans from the test dataset. For the chest slice detection model, the accuracy achieved is 0.95, recall is 1.00 and precision of 0.52. For the pelvis slice detection model, the accuracy achieved is 0.98, recall is 0.88 and precision of 1.00. The performance of these detection models is good as they achieved a high accuracy rate however; the recall and precision are low for the model trained for chest slice detection due to some misclassifications. The misclassified slices were the ones where the liver extends into the chest cavity. This resulted in some false positives and false negatives which further resulted in the low recall and precision.

Table 6.6: Results of *Small* Network for Pelvis Slice Detection

Image	No. of Slices	Pelvis (P)	Abdomen (N)	TP	TN	FP	FN
1	111	-	111	-	111	-	-
2	121	-	121	-	121	-	-
3	245	53	192	33	192	-	20
4	335	13	322	10	322	-	3
5	183	71	112	70	112	-	1
6	165	75	90	74	90	-	1
7	91	7	84	6	84	-	1
8	258	-	258	-	258	-	-
9	97	-	97	-	97	-	-
10	73	5	68	3	68	-	2

Table 6.7: Accuracy of *Small* Network for Pelvis Slice Detection

Image	No. of Slices	Recall	Precision	Accuracy
1	111	1	1	1
2	121	1	1	1
3	245	0.62	1	0.92
4	335	0.77	1	0.99
5	183	0.99	1	0.99
6	165	0.99	1	0.99
7	91	0.86	1	0.99
8	258	1	1	1
9	97	1	1	1
10	73	0.6	1	0.97
Average		0.88	1	0.98

Table 6.8: Results of *Small* Network for Chest Slice Detection

Image	No. of Slices	Chest (P)	Abdomen (N)	TP	TN	FP	FN
1	111	14	97	14	90	7	-
2	121	14	107	14	99	8	-
3	192	5	187	5	180	7	-
4	322	18	304	18	294	10	-
5	112	5	107	5	103	4	-
6	90	8	82	8	77	5	-
7	84	1	83	1	79	3	-
8	258	18	240	18	223	17	-
9	97	2	95	2	94	1	-
10	68	1	67	1	63	4	-

Table 6.9: Accuracy of *Small* Network for Chest Slice Detection

Image	No. of Slices	Recall	Precision	Accuracy
1	111	1	0.67	0.94
2	121	1	0.64	0.93
3	245	1	0.42	0.96
4	335	1	0.64	0.97
5	183	1	0.56	0.96
6	165	1	0.62	0.94
7	91	1	0.25	0.95
8	258	1	0.51	0.93
9	97	1	0.67	0.99
10	73	1	0.25	0.94
Average		1	0.52	0.95

Deep Network

The deep networks are evaluated in the same manner as the small networks. The results during training are shown in Figure 6.7(a) for the chest slice detection model and Figure 6.7(b) for the pelvis slice detection model. Figure 6.7(a) shows that the model converged around epoch 12 where the purple line representing validation loss and the black line representing training loss converges and the red line representing training accuracy and the blue line represenitng validation accuracy converges. Similarly, in Figure 6.7(b)the convergence of the model occurs around epoch 15. The end results of the training of both models are shown in Table 6.10 and Table 6.11. For the chest slice detection model training, the loss is 0.0063 and the accuracy is 0.9959. The loss of validation is 0.0183 and accuracy is 0.9920. The training loss for the pelvis slice detection model is 0.0098 and the accuracy is 0.9942. The validation loss is 0.0137 and the accuracy is 0.9932.

Figure 6.7: Training loss and accuracy of *deep* networks during training

Table 6.10: Results of *Deep* Network for Chest Slice Detection

	Training		Validation	
Metric	**Loss**	**Accuracy**	**Loss**	**Accuracy**
Score	0.0063	0.9959	0.0183	0.9920

Table 6.11: Results of *Deep* Network for Pelvis Slice Detection

	Training		Validation	
Metric	**Loss**	**Accuracy**	**Loss**	**Accuracy**
Score	0.0098	0.9942	0.0137	0.9932

The results after training are shown in Table 6.12 and Table 6.13 for the chest slice detection model and Table 6.14 and Table 6.15 for the pelvis slice detection model. These results show the recall, precision and accuracy after testing the models on the CT scans from the test dataset. For the chest slice detection model, the accuracy achieved is 0.97, recall is 0.86 and precision of 0.92. For the pelvis slice detection model, the accuracy achieved is 0.99, recall is 0.95 and precision of 1.00. The deep networks performed more successfully as is evident in the results. The reason for a low recall value is the misclassifications of a small number of chest slices where the liver extends into the chest cavity.

Table 6.12: Results of *Deep* Network for Chest Slice Detection

Image	No. of Slices	Chest (P)	Abdomen (N)	TP	TN	FP	FN
1	111	14	97	5	97	-	9
2	121	14	107	6	107	-	8
3	192	5	187	5	187	-	-
4	322	18	304	18	297	7	-
5	112	5	107	4	107	-	1
6	90	8	82	8	76	6	-
7	84	1	83	-	83	-	1
8	258	18	240	18	239	1	-
9	97	2	95	-	95	-	2
10	68	1	67	-	67	-	1

Table 6.13: Accuracy of *Deep* Network for Chest Slice Detection

Image	No. of Slices	Recall	Precision	Accuracy
1	111	0.36	1	0.92
2	121	0.43	1	0.93
3	245	1	1	1
4	335	1	0.72	0.98
5	183	0.80	1	0.99
6	165	1	0.57	0.93
7	91	0.99	1	0.99
8	258	1	0.95	0.99
9	97	0.98	1	0.98
10	73	0.99	1	0.99
Average		0.86	0.92	0.97

Table 6.14: Results of *Deep* Network for Pelvis Slice Detection

Image	No. of Slices	Pelvis (P)	Abdomen (N)	TP	TN	FP	FN
1	111	-	111	-	111	-	-
2	121	-	121	-	121	-	-
3	245	53	192	49	192	-	4
4	335	13	322	12	322	-	1
5	183	71	112	70	112	-	1
6	165	75	90	74	90	-	1
7	91	7	84	6	84	-	1
8	258	-	258	-	258	-	-
9	97	-	97	-	97	-	-
10	73	5	68	4	68	-	1

Table 6.15: Accuracy of *Deep* Network for Pelvis Slice Detection

Image	No. of Slices	Recall	Precision	Accuracy
1	111	1	1	1
2	121	1	1	1
3	245	0.92	1	0.98
4	335	0.92	1	0.99
5	183	0.99	1	0.99
6	165	0.99	1	0.99
7	91	0.86	1	0.99
8	258	1	1	1
9	97	1	1	1
10	73	0.80	1	0.99
Average		0.95	1	0.99

Comparing the results achieved by the small networks to those of the deep networks, it is evident that the deep networks were more successful in their respective tasks. The deep detection models achieved higher accuracy rates than those of the small detection models. This was due to the deep detection models having a lower number of misclassified slices.

Table 6.16: Comparison of Chest Slice Detection Methods

Method	Recall	Precision	Accuracy
AlexNet [95]	0.71	0.63	0.56
GoogLeNet [95]	0.67	0.62	0.58
Proposed (*Small* network)	**0.52**	**1.0**	**0.95**
Proposed (*Deep* network)	**0.86**	**0.92**	**0.97**

Table 6.17: Comparison of Pelvis Slice Detection Methods

Method	Recall	Precision	Accuracy
AlexNet [95]	0.42	0.52	0.56
GoogLeNet [95]	0.40	0.60	0.58
Proposed (*Small* network)	**0.88**	**1.0**	**0.98**
Proposed (*Deep* network)	**0.95**	**1.0**	**0.99**

The results obtained in this research, are compared to the results obtained using AlexNet and GoogLeNet in a study done by Sugimori [95]. It is evident in Table 6.16 and Table 6.17 that the accuracy rate achieved by all network models in this research are significantly higher than the accuracy rates achieved by AlexNet

and GoogLeNet. For the chest slice detection models, this research achieved 0.95 and 0.97 using the small and deep network models respectively whereas, AlexNet achieved 0.56 and GoogLeNet achieved 0.58. For the pelvis slice detection models, this research achieved 0.98 and 0.99 using the small and deep network models respectively whereas, AlexNet achieved 0.56 and GoogLeNet achieved 0.58. This may be because of the very large number of datasets that were used in Sugimori's study compared to the small number used in this research. Also, these classification models were designed for medical images whereas AlexNet and GoogLeNet were designed specifically to classify natural images. Natural images are optical images formed by focusing light onto some sort of 2D sensor array through an optical system whereas medical images are 2D arrays of data where the values for each pixel valus is converted to an intensity or colour for display [5]. AlexNet and GoogLeNet were pretrained for the classification of natural images which may influence its performance in classification of medical images.

6.7 Results of Liver Location and Segmentation

6.7.1 Experimental Setup

An experiment was set up to test the liver location and segmentation methods in a scientific manner. The experiment consists of two parts. First, the two convolutional neural network models; the model without the concatenate layers and the model with the concatenate layers, are evaluated. This was done to determine which convolutional network would perform better to be used in this methodology. The second part evaluates the whole methodology including the region of interest detection method as well as the post-processing method. Thereafter, the methodology is compared to related literature.

From the dataset described above, the training set was split up so that 80% of the images were used to train the networks and 20% was used for validation during the training. The test set was kept completely out of the training process and the training and validation ratio of 80:20 was based on the small size of the dataset. Through data analysis it was evident that there was variance throughout the training, validation and test sets therefore, the choice of split of the dataset is reasonable.

6.7.2 Evaluation

The two trained models are evaluated during training and after training. During the training of a convolutional neural network, the training and validation errors or loss are used to evaluate the network. After training the model, the networks are used to segment the liver in the 2D slices which are reconstructed into a 3D volume. Thereafter, the resulting segmented volumes are compared to the ground

truth volumes. The volumes are evaluated using volume overlap error, relative volume difference, maximum symmetric surface distance, average symmetric surface distance and root mean square symmetric surface distance [106]. The best performing CNN results are also represented in Dice similarity coefficient in order to compare its performance with some related work [107].

Volume Overlap Error (VOE): The segmentation volume is compared to the ground truth volume and the overlap error is calculated in percentage. This calculates the percentage of regions where the two volumes do not overlap with each other. A perfect segmentation produces an overlap error percentage of 0. Equation 6.5 defines VOE.

$$VOE = 100\left(1 - \frac{A \cap B}{A \cup B}\right) \tag{6.5}$$

where A denotes the segmented volume produced by the proposed method and B denotes the segmentated volume done by experts.

Relative Volume Difference (RVD): The segmentation volume is compared to the ground truth volume and the relative volume difference is calculated in percentage. Negative values suggest that the segmentation result is smaller than the ground truth which denotes under-segmentation whereas positive values suggest that the segmentation result is larger than the ground truth denoting over-segmentation. A perfect segmentation produces a relative volume difference of 0 percent. Equation 6.6 defines RVD.

$$RVD = 100\left(\frac{\mid A \mid - \mid B \mid}{\mid B \mid}\right) \tag{6.6}$$

where $|A|$ is the absolute value of the segmented volume produced by the proposed method and $|B|$ is the absolute value of the ground-truth volume.

Average Symmetric Surface Distance (ASD): The average symmetric surface distance is calculated in millimeters. For a perfect segmentation this distance is 0. Equation 6.7 defines ASD.

$$ASD = \frac{1}{|S(A)| + |S(B)|} \left(\sum_{S_A \in S(A)} d(S_A, S(B)) + \sum_{S_B \in S(B)} d(S_B, S(A)) \right) \tag{6.7}$$

where $S(A)$ is the set of surface voxels of A, $S(B)$ is the set of surface voxels of B. The shortest distance of a voxel in $S(B)$ to $S(A)$ is defined as $d(S_B, S(A))$ and a voxel in $S(A)$ to $S(B)$ is defined as $d(S_A, S(B))$.

Root Mean Square Symmetric Surface Distance (RMSSD): The root mean square symmetric surface distance is calculated between the segmentation volume and the ground truth volume in millimeters. The result for a perfect segmentation

is 0. This metric punishes large deviations from the contour stronger than the average symmetric surface distance. Equation 6.8 defines RMSD.

$$RMSD = \sqrt{\frac{1}{|S(A)| + |S(B)|} \left(\sum_{S_A \in S(A)} d^2(S_A, S(B)) + \sum_{S_B \in S(B)} d^2(S_B, S(A)) \right)}$$
(6.8)

where $S(A)$ is the set of surface voxels of A, $S(B)$ is the set of surface voxels of B. The square of the shortest distance of a voxel in $S(B)$ to $S(A)$ is defined as $d^2(S_B, S(A))$ and the square of the shortest distance of a voxel in $S(A)$ to $S(B)$ is defined as $d^2(S_A, S(B))$.

Maximum Symmetric Surface Distance (MSSD): The maximum symmetric surface distance is calculated between the segmentation volume and the ground truth volume in millimeters. For a perfect segmentation this distance is 0. This metric is sensitive to outliers and returns the true maximum error. Equation 6.9 defines MaxASD.

$$MaxSD = max \left(max_{S_A \in S(A)} d(S_A, S(B)), \ max_{S_B \in S(B)} d(S_B, S(A)) \right)$$
(6.9)

where A is the segmented volume, B is the ground truth volume, $S(A)$ is the set of surface voxels of A, $S(B)$ is the set of surface voxels of B. The distance of a voxel in $S(B)$ to $S(A)$ is defined as $d(S_B, S(A))$ and the distance of a voxel in $S(A)$ to $S(B)$ is defined as $d(S_A, S(B))$.

$$d(v, S(A)) = min_{S_A \in S(A)} ||v - S_A||$$
(6.10)

where $|| \cdot ||$ is Euclidean distance between voxel v and $S(A)$.

Dice Similary Coefficient (DSC): The dice similarity coefficient measures the similarity and overlap between the ground truth B segmentation and the segmentation results A achieved. The performance index ranges between zero to one with an index zero signifying no overlap between B and A while index one signifies a perfect overlap between them.

$$DSC = 2\frac{B \cap A}{|B| + |A|}$$
(6.11)

where A denotes the segmented volume produced by the proposed method and B denotes the segmentated volume done by experts. $|A|$ is the absolute value of the segmented volume produced by the proposed method and $|B|$ is the absolute value of the ground-truth volume.

6.7.3 Results and Discussion

The results of CNN without concatenate layers while training is represented in the Figure 6.8. Figure 6.8(a) shows that the model converged around epoch 20 where the red line representing validation loss and the blue line representing training loss converges. Similarly, in Figure 6.8(b) shows the convergence of the model with respect to validation accuracy represented by the red line and training accuracy represented by the blue line. The results during training of the network model without concatenate layers are shown in Table 6.18. The training loss is 0.0101 and the training accuracy is 0.9601. The validation loss is 0.0083 and validation accuracy is 0.9717. The results while training of CNN with concatenate layers is represented in the Figure 6.9. Figure 6.9(a) shows that the model converged around epoch 21 where the red line representing validation loss and the blue line representing training loss converges. Similarly, in Figure 6.9(b) shows the convergence of the model with respect to validation accuracy represented by the red line and training accuracy represented by the blue line. The results during training of the network model with concatenate layers are shown in Table 6.19. The training loss is 0.0109 and the training accuracy is 0.9585. The validation loss is 0.0061 and validation accuracy is 0.9754.
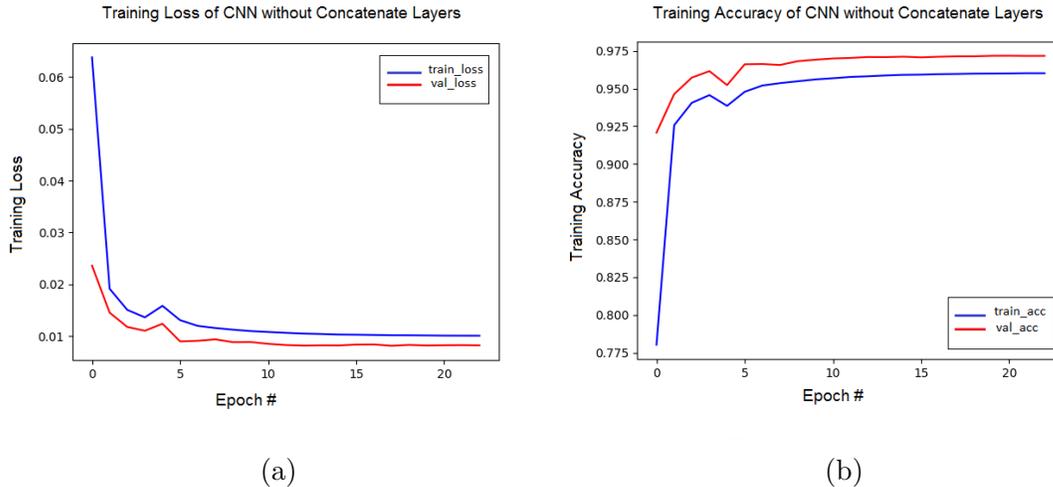

(a) (b)

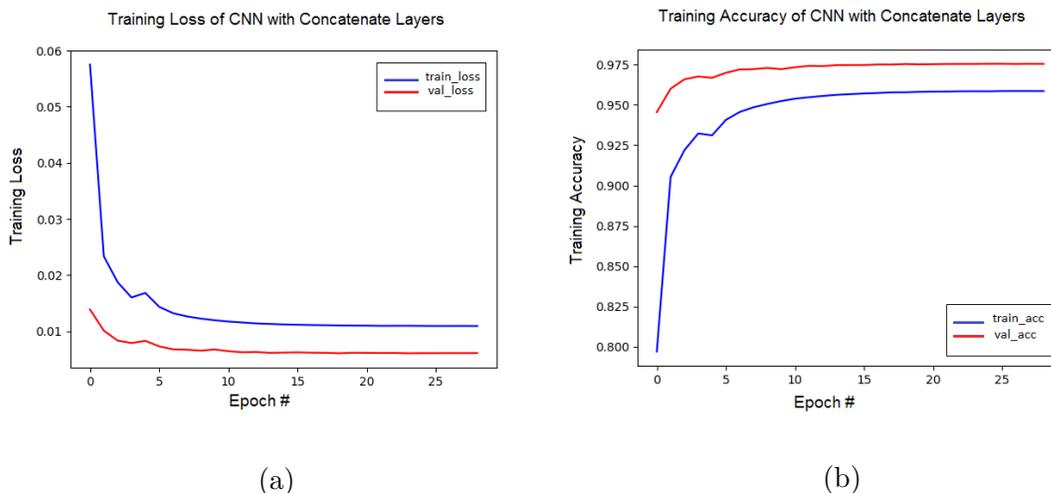Figure 6.8: Training loss and accuracy of CNN without concatenate layers

Figure 6.9: Training loss and accuracy of CNN with concatenate layers

Table 6.18: Results of CNN without Concatenate Layers

Metric	Training		Validation	
	Loss	Accuracy	Loss	Accuracy
Score	0.0101	0.9601	0.0083	0.9717

Table 6.19: Results of CNN with Concatenate Layers

Metric	Training		Validation	
	Loss	Accuracy	Loss	Accuracy
Score	0.0109	0.9585	0.0061	0.9754

The results after training are shown in Table 6.20 for the convolutional network model without concatenate layers. The average results obtained are 18.89% for VOE, -4.58% for RVD, 6.36 for ASD, 4.99mm for RMSD and 81.63mm for MaxASD. The results after training are shown in Table 6.21 for the convolutional network model with concatenate layers. The average results obtained are 16.39% for VOE, -8.77% for RVD, 3.32mm for ASD, 2.98mm for RMSD and 120.97mm for MaxASD. As is evident in the two sets of results obtained the addition of concatenate layers has an effect on the results. Even though the results for RVD, ASD, RMSD and MaxASD increased, the addition of concatenate layers in the network decreases the VOE. In most cases in the test dataset the VOE decreased with concatenate layers. There were only two cases, image 1 and 5 in Table6.20 and Table6.21, where the VOE increased. In 8 out of the ten cases, the VOE decreased and for image 8 the VOE decreased significantly by 16.50%.

Table 6.20: Results of CNN without Concatenate Layers

Image	VOE[%]	RVD[%]	ASD[mm]	RMSD[mm]	MaxASD[mm]
1	15.59	2.50	3.09	3.49	44.87
2	14.82	3.09	4.89	5.19	65.22
3	22.59	-12.60	6.67	4.37	68.86
4	12.79	-0.03	7.04	6.44	129.02
5	20.07	12.16	14.00	9.02	125.72
6	24.85	4.87	10.96	7.92	127.49
7	11.48	-4.58	2.44	3.77	48.75
8	36.05	-33.70	9.92	5.08	152.13
9	14.25	-6.03	2.42	2.52	30.02
10	16.37	-11.52	2.04	2.08	24.19
Average	18.89	-4.58	6.36	4.99	81.63

Table 6.21: Results of CNN with Concatenate Layers

Image	VOE[%]	RVD[%]	ASD[mm]	RMSD[mm]	MaxASD[mm]
1	16.22	9.16	4.64	5.13	73.99
2	12.79	6.41	3.35	4.479	68.53
3	15.17	-3.55	4.84	4.59	76.35
4	12.57	6.98	7.76	6.94	134.72
5	32.51	39.44	20.16	9.38	169.90
6	24.43	14.69	14.03	8.57	167.94
7	9.99	-2.41	2.35	4.07	48.87
8	19.55	-13.14	7.18	5.80	145.88
9	12.85	-0.94	2.60	3.71	101.99
10	12.09	-3.08	1.70	2.52	33.76
Average	16.82	5.36	6.87	5.52	102.19

The final results achieved by the framework including the region of interest detection, liver location and segmentation by the CNN without concatenate layers as well as the post-processing is presented in Table 6.22. The average results are 16.33% for VOE, -8.45% for RVD, 3.26mm for ASD, 2.74mm for RMSD and 43.59mm for MaxASD. The final results achieved by the framework including the region of interest detection, liver location and segmentation by the CNN with concatenate layers in the model as well as post-processing is presented in Table 6.23. The average results are 12.07% for VOE, -1.96% for RVD, 2.25mm for ASD, 2.60mm for RMSD and 43.01mm for MaxASD. All the results except MaxASD decreased significantly which proves that the framework using concatenate layers improves the results. This CNN with concatenate layers within the network achieved a DSC of 93.57%.

The examples of the final results of each CNN model are shown in Fingure

Table 6.22: Final Results of CNN without Concatenate Layers

Image	VOE[%]	RVD[%]	ASD[mm]	RMSD[mm]	MaxASD[mm]
1	14.86	1.56	2.28	2.30	31.97
2	12.52	0.20	2.13	2.29	16.61
3	20.81	-16.76	5.79	4.06	51.24
4	10.86	-2.35	3.32	3.06	47.47
5	10.40	-0.55	1.42	1.76	16.67
6	17.57	-7.56	2.64	3.11	38.85
7	10.69	-5.50	1.52	2.09	27.22
8	35.91	-34.81	9.54	4.70	152.12
9	13.64	-6.78	2.07	2.00	29.85
10	15.98	-11.99	1.93	2.01	24.19
Average	16.33	-8.45	3.26	2.74	43.59

Table 6.23: Final Results of CNN with Concatenate Layers

Image	VOE[%]	RVD[%]	ASD[mm]	RMSD[mm]	MaxASD[mm]
1	14.27	6.57	2.56	2.79	23.19
2	10.33	3.37	1.77	2.41	32.34
3	12.83	-7.07	3.10	3.19	49.30
4	8.67	2.24	2.50	2.48	39.22
5	8.41	1.86	1.19	1.83	18.25
6	14.90	-3.11	2.33	3.20	40.61
7	9.26	-3.24	1.31	1.94	23.26
8	18.58	-14.39	4.29	3.65	145.88
9	11.54	-2.51	1.82	2.05	24.29
10	11.69	-3.55	1.59	2.41	33.76
Average	12.07	-1.96	2.25	2.60	43.01

6.10 where the segmentation outlined in green represent the results obtained by the CNN without concatenate layers, blue represents the results obtained by the CNN with concatenate layers and red represents the ground truth segmentation.

Table 6.24: Comparison of Results of Liver Segmentation Frameworks

Related Work	VOE[%]	Dice[%]
Peng et al [81]	4.53	97.70
Wu et al [106]	7.87	-
Goryawala et al [41]	72.38	92.00
Stawiaski et al [94]	29.49	-
Furukawa et al [37]	12.80	-
Hu et al [52]	5.36	-
Lu et al [70]	5.90	-
Vivanti et al [103]	16.75	90.47
Dou et al [28]	5.42	-
Massoptier and Casciaro [73]	-	95.00
Massoptier and Casciaro [74]	-	94.20
Luo et al [?]	-	97.30
Li et al [65]	-	99.86
Huang et al [53]	-	99.00
Goryawala et al [40]	-	98.27
Singh et al [93]	-	95.00
Christ et al [21]	-	89.50
Christ et al [21]	-	87.00
Ben-Cohen et al [14]	-	89.00
Ben-Cohen et al [14]	-	87.00
Ben-Cohen et al [14]	-	88.00
Proposed	**12.07**	**93.57**

The final results of the framework using the CNN with concatenate layers in the network model are compared to those of related work. The results are presented in Table 6.24. The results in terms of volume overlap error and Dice similarity, were better than those results of Goryawala et al [41], Stawiaski et al [94], Furukawa et al [37], Vivanti et al [103] as well as the works of Christ et al [21] and Ben-Cohen et al [14].
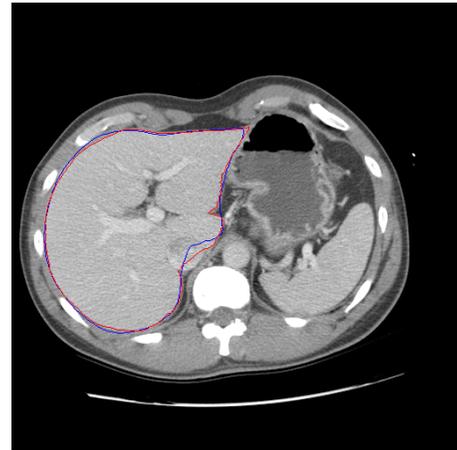
6.8 Conclusion

In this chapter, the dataset, the liver segmentation framework and the results achieved from the framework were discussed. Due to the dataset containing images of variance, the nuances associated with the dataset were discussed. In order to understand how the processes combined to reach the final results, the liver segmentation framework was represented in its entirety. The results of the region of interest detection models were presented, discussed and compared and the results of the liver location and segmentation models were presented, discussed and compared. Thereafter, the results of the combined processing of the region of interest

detection, liver location and segmentation and the segmentation refinement were presented and discussed. Finally, these results were compared to those in related literature.

The next chapter will draw more indepth conclusions as to what the results mean. It will also look at what future work can be done to improve the presented liver segmentation framework described in this dissertation.
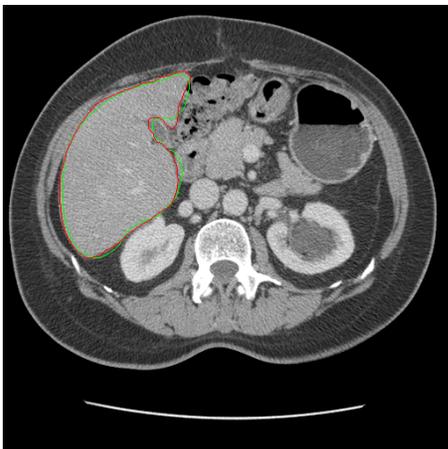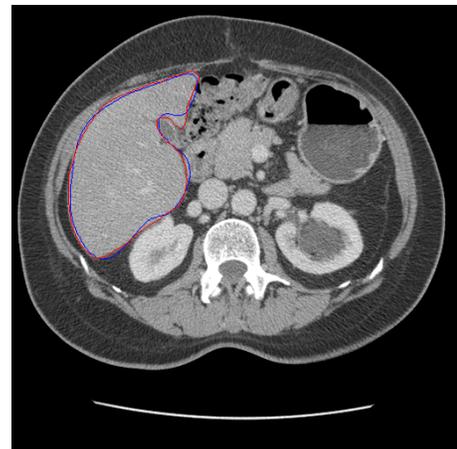
Figure 6.10: Images showing comparison between results of both CNN segmentation models and ground truth segmentation

Chapter 7

Conclusion

7.1 Summary

Segmentation of organs are of great importance for the success on CAD systems in medicine as accurate segmentation of an organ allows for better analysis and a more accurate diagnosis. The method currently used for segmentation is manual segmentation from medical imaging that is conducted by a radiologist because research has not been able to provide an accuracy rate of 100%. However, research has enabled the use of computers to assist in the diagnostic process as a second-opinion while an expert physician makes the final decision.

The use of deep-learning has achieved state-of-the-art performance in medical image analysis. Deep learning techniques have been applied to medical image analysis with the intention of letting computers learn the features that optimally represent the problem data at hand. CNNs, one of the most successful techniques in deep learning, have been used in classification tasks, object or organ detection and segmentation. Although, deep learning techniques have opened up a new avenue for the success of medical image analysis, there is still some room for improvement to increase the accuracy of results.

A method for liver segmentation using 3D CT scans is proposed in this research work. CNNS are implemented to extract the liver from its background in the CT image. Firstly, a CNN is used to classify each slice of a 3D scan in order to remove slices that do not belong to the abdomen. This is done to obtain a region of interest that contains the liver which is processed further during liver segmentation. The abdominal slices from the region of interest are processed by a CNN to locate and segment the liver. The resulting segmented liver slices are then reassembled into a volume for post-processing. The proposed method does not require user-interaction and performs automatic segmentation.

7.1.1   Contributions

- In this research work, a method to align serially acquired 2D slices is presented [48]. The method consists of four stages where the first stage is pre-processing where the entire 3D CT volume undergoes smoothing using a median filter. The second stage applies a Laplacian weighted histogram to determine the optimal liver threshold which is used in the next stage for skeletonization of the entire image volume. The last stage includes aligning the slices of the skeletonized volume with respect to neighbouring slices. The main contribution of this approach is to combine the use of skeletonization and the randomized alignment algorithm for the alignment of 2D slices. The images are aligned with their neighbouring slices in a random fashion rather than a sequential formulation which prevents the propagation through the volume of errors. This approach performed well as the MSD improved by an average of 32.85% in experiment 3 is which is significantly higher than the improvement in the other two experiments. This method also ensures that no particular direction is privileged in the method avoiding any global offsets, biases in the estimation and error propagation.

- A method to obtain the region of interest for liver segmentation is proposed [49]. Since the liver occurs in the abdomen, the 2D slices that contain the pelvis and the chest are detected using two CNNs. These slices are then removed from the further processing, which localizes further processing to the abdomen where the liver is located. This also increases the chance of a higher accuracy rate as well as reduces the computation time of segmentation. Two different CNNs are implemented and trained to perform the chest slice detection and pelvis slice detection. One CNN is a small network with a few layers and the other is a deeper network with many layers where a filter size is repeated for more than one layer. These networks are trained and tested on the same dataset in order to determine which network performs better and can be used prior to the liver segmentation step. Experiments were carried out and the deeper networks produced better results. The deeper networks achieved an accuracy rate of 97% for chest slice detection and 99% for pelvis slice detection while the small networks achieved an accuracy rate of 95% for chest slice detection and 98% for pelvis slice detection. This proves that deeper networks produce more accurate results for medical image detection.

- A CNN is implemented to locate and segment the liver from each 2D slice [50]. This is due to the ability of CNNs to learn where the liver is located and segment liver tissue. The use of concatenate layers in the network and its effect on liver segmentation is investigated. Experiments are conducted on two CNNs to investigate this. One CNN does not contain any concatenate layers while the other contains two concatenate layers among the convolution layers in order to test its effectiveness on the segmentation of the liver to determine whether this addition improves the results produced by the convolutional neural network. The use of concatenate layers improved the VOE from 18.94%

to 16.39%, the RVD from 2.61% to -8.77%, the ASD from 6.38mm to 3.32mm, the RMSD from 5.10mm to 2.98mm and the MaxASD from 144.64mm to 120.97mm. It is evident from these results that the addition of concatenate layers in a CNN improves the accuracy of the segmentation task.

- Finally, experiments were carried out to test the effect of the region of interest detection method combined with the liver location and segmentation with concatenate layers. The accuracy improved the VOE from 16.39% to 12.07%, the RVD from -8.77% to -2.33%, the ASD from 3.32mm to 2.26mm and the RMSD from 2.98mm to 2.86mm. From these results, it is evident that the use of region of interest detection that implements deep learning techniques improves the accuracy of the entire liver segmentation method. Compared to the results obtained in related literature, the results achieved in this work performed better than some. It can be concluded that the proposed liver segmentation framework performed well.

7.2   Limitations and Future Work

The proposed liver segmentation framework uses deep learning in the form of CNNs to learn features in the images in order to locate and segment the liver. The methods presented are successful in their tasks however; there are limitations which can be improved upon. Therefore, future work will involve finding ways to overcome these limitations.

Firstly, the segmentation method is strongly tied to the training of the convolutional network models during which important features are learned from scratch which enables the network model to successfully analyse test images and perform its task. The model needs enough data to be able to generalize well as well as an adequate amount of data to validate against. Since the dataset used was relatively small, future work will improve the model's ability to generalize and produce more accurate results by increasing the size of the dataset significantly.

Secondly, during the investigation of the region of interest detection methods it was concluded that deeper networks produce better results. In future studies, the use of deeper networks with more convolution layers could be explored as well as comparisons to the performance of other neural networks that were trained on medical images.

List of References

[1] Abdominal ct. `https://www.radiologyinfo.org/en/info.cfm?pg=abdominct`. Accessed 2018-08-20.

[2] Body ct. `https://www.radiologyinfo.org/en/info.cfm?pg=bodyct`. Accessed: 2018-08-18.

[3] Body ct. `https://www.radiologyinfo.org/en/info.cfm?pg=bodyct`. Accessed 2018-08-18.

[4] Ct scan. `https://www.emedicinehealth.com/ct_scan/article_em.htm#what_is_a_ct_scan`. Accessed: 2018-08-18.

[5] The difference between real (optical) images and synthetic (computed) images. `https://blog.fluidimaging.com/blog/real-optical-versus-synthetic-computed-images/`. Accessed 2019-03-02.

[6] Fully convolutional networks (fcn) for 2d segmentation. `http://deeplearning.net/tutorial/fcn_2D_segm.html`. Accessed 2018-08-20.

[7] How do ct scans work. `https://www.floridamedicalclinic.com/blog/how-do-ct-scans-work/`. Accessed 2018-08-20.

[8] Liver- anatomy and function of the human liver. `http://www.innerbody.com/image_digeov/card10-new2.html`. Accessed 2018-08-24.

[9] Liver anatomy and functions. `http://www.hopkinsmedicine.org/healthlibrary/conditions/liver_biliary_and_pancreatic_disorders/liver_anatomy_and_functions_85,P00676`. Accessed: 2017-03-01.

[10] To what extent is cadcam technology influencing the medical industry. `https://www.indiacadworks.com/blog/to-what-extent-is-cadcam-technology-influencing-the-medical-industry/`. Accessed: 2018-08-18.

[11] Writing a systematic literature review. `https://blog.efpsa.org/2018/01/03/writing-a-systematic-literature-review/`. Accessed 2019-03-03.

[12] Marwan Alshipli and Norlaili A. Kabir. Effect of slice thickness on image noise and diagnostic content of single-source-dual energy computed tomography. *Journal of Physics*, 2017.

[13] Reinhard Beichel, Christian Bauer, Alexander Bornik, Erich Sorantin, and Horst Bischof. Liver segmentation in ct data: A segmentation refinement approach. 9 2017.

[14] Avi Ben-Cohen, Idit Diamant, Eyal Klang, Michal Amitai, and Hayit Greenspan. Fully convolutional network for liver segmentation and lesions detection. pages 77–85, 2016.

[15] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14, 2 1992.

[16] D. Brown and K. Bradshaw. Improved fingercode alignment for accurate and compact fingerprint recognition. *2016 IEEE Symposium on Technologies for Homeland Security (HST)*, pages 1–6, 2016.

[17] Jinzheng Cai, Le Lu, Zizhao Zhang, Fuyong Xing, Lin Yang, and Qian Yin. Pancreas segmentation in mri using graph-based decision fusion on convolutional neural networks. *MICCAI 2016, Part 2, LNCS 9901*, pages 442–450, 2016.

[18] Paola Camadelli, Elena Casiraghi, and Andrea Esposito. Liver segmentation from computed tomography scans: A survey and a new algorithm. *Artificial Intelligence in Medicine*, 45:185–196, 7 2009.

[19] Chung-Ming Chen, Yi-Hong Chou, Norio Tagawa, and Younghae Do. Computer-aided detection and diagnosis in medical imaging. *Computational and Mathematical Methods in Medicine*, 7 2013.

[20] Francois Chollet. Xception: deep learning with depthwise separable convolutions. 12 2015.

[21] Patrick Ferdinand Christ, Florian Ettlinger, Felix Grun, Mohamed Ezzeldin A. Elshaer, Jana Lipkova, Sebastian Schlecht, Freba Ahmaddy, Sunil Tatavarty, Marc Bickel, Patrick Bilic, Markus Rempfler, Felix Hofmann, Melvin D. Anastatsi, Seyed-Ahmad Ahmadi, Georgios Kaissis, Julian Holch, Weiland Sommer, Rickmer Baren, Volker Heinemann, and Bjoern Menze. Automatic liver and tumor segmentation of ct and mri volumes using cascaded fully convolutional neural networks. *Medical Image Analysis*, 2 2017.

[22] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Seutens, and G. Marchal. Automated multi-modality in image registration based on information theory. *Information processing in medical imaging*, 3(6):263–274, 1995.

[23] D. L. Collins, G. L. Goualher, and A. C. Evans. Non-linear cerebral registration with sulcal constraints. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 974–984, 1998.

[24] The International Liver Congress. Background media information: Fast facts about liver disease. 4 2016.

[25] Laura Fernandez de Manuel, Jose L. Rubio, Maria J. Ledesma-Carbayo, Javier Pacau, Jose M. Tellado, Enrique Ramon, Manuel Desco, and Andres Santos. 3d liver segmentation in preoperative ct images using a level-sets active surface method. *31st Annual Internation Conference of the IEEE EMBS*, 8 2009.

[26] Bob D. de Vos, Jelmer M. Wolterink, Pim A. de Jong, Tim Leiner, Max A. Viergever, and Ivana Isgum. Convnet-based localization of anatomical structures in 3d medical images. 4 2017.

[27] Kunio Doi. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Comput Med Imaging Graph*, 31:198–211, 6 2007.

[28] Qi Dou, Hao Chen, Yueming Jin, Lequan Yu, Jing Qin, and Pheng-Ann Heng. 3d deeply supervised network for automatic liver segmentation from ct volumes. 7 2016.

[29] Marius Erdt, Sebastian Steger, Matthias Kirschner, and Stefan Wesarg. Fast automatic liver segmentation combining learned shape priors with observed shape deviation. *2010 IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS)*, pages 249–254, 2010.

[30] G. D. Evangelidis and E. Z. Psarakis. Parametric image alignment using enhanced correlation coefficient maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30, 10 2008.

[31] G. D. Evangelidis and E. Z. Psarakis. Projective image alignment using enhanced correlation coefficient maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10):1858–1865, 2008.

[32] Alun Evans, Tryphon Lambrou, Alf Linney, and Andrew Todd-Pokropek. Automatic segmentation of liver from computerised tomography(ct) images. *Proceeding of Medical Image Understanding and Analysis*, pages 57–60, 2004.

[33] B. Fei, J. L. Duerk, D. T. Boll, J. S. Lewin, and D. L. Wilson. Slice-to-volume registration and its potential application to interventional mri-guided radio-frequency thermal ablation of prostate cancer. *IEEE Transactions on Medical Imaging*, 22, 4 2003.

[34] E. Ferrante and N. Paragios. Slice-to-volume medical image registartion: a survey. *Medical image analysis*, 39:101–123, 2017.

[35] B. Fischer and J. Modersitzki. A unified apporach to fast image registration and a new curvature based registration technique. *Science Direct, Linear Algebra and its application*, 10 2003.

[36] Charles Florin, Nikos Paragois, Gareth Funka-Lea, and James Williams. Liver segmentation using sparse 3d prior models with optimal data support. *Biennial International Conference on Information Processing in Medical Imaging*, pages 38–49, 2007.

[37] Daisuke Furukawa, Akinobu Shimizu, and Hidefumi Kobatake. Automatic liver segmentation method based on maximum a posterior probability estimation and level set method. *3D Segmentation in The Clinic: A Grand Challenge*, 117, 2007.

[38] M. Gasser. Transfer functions for volume rendering applications and implementation results with vtk. *Institue of computer graphics and algorithms*.

[39] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks.

[40] Mohammed Goryawala, Magno R. Guillen, Mercedes Cabrerizo, Armando Barreto, Seza Gulec, Tushar C. Barot, Rekha R. Suthar, Ruchir N. Bhatt, Anthony McGoron, and Malek Adjouadi. A 3d liver segmentation method with parallel computing for selective internal radiation therapy. *IEEE Transactions on Information Technology in Biomedicine*, 16, 1 2012.

[41] Mohammed Goryawala, Seza Gulec, Ruchir Bhatt, Anthony J. McGoron, and Malek Adjouadi. A low-interaction automatic 3d liver segmentation method using computed tomography for selective internal radiation therapy. *BioMed Research International*, 7 2014.

[42] Payel Gosh and Melanie Mitchell. Segmentation of medical images using a genetic algorithm. pages 1171–1178, 2006.

[43] Akshat Gotra, Lojan Sivakumaran, Gabriel Chartrand, Kim-Nhien Vu, Franck Vandenbroucke-Menu, Claude Kauffman, Samuel Kadoury, Benoit Gallix, Jacques A. de Guise, and An Tang. Liver segmentation: indications, techniques and future directions. *Insights Imaging*, 8:377–392, 2017.

[44] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. Mask r-cnn. *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.

[45] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. 12 2015.

[46] Thomas Heeneman. Lung nodule detection by using deep learning. 1 2018.

106

[47] A. Hill, A. Thornham, and C. J. Taylor. Model-based interpretation of 3d medical images. *4th British Machine Vision Vonference*, 9 1993.

[48] A. Hiraman, S. Viriri, and M. Gwetu. Efficient 3d liver segmentation using enhanced image alignment. *IEEE International Conference on Information Communication Technology and Society*, pages 1–8, 3 2018.

[49] A. Hiraman, S. Viriri, and M. Gwetu. Efficient region of interest detection for liver segmentation using 3d ct scan. *IEEE International Conference on Information Communication Technology and Society*, 11 2018.

[50] A. Hiraman, S. Viriri, and M. Gwetu. Liver segmentation using 3d ct scans. *Image Analysis and Stereology Journal*, 12 2018.

[51] Assaf Hoogi, John W. Lambert, Yefeng Zheng, Dorin Comaniciu, and Daniel L. Rubin. A fully-automated pipeline for detection and liver lesions and pathological lymph nodes. *arXiv preprint arXiv:1703.06418*, 2017.

[52] Peijun Hu, Fa Wu, Jialin Peng, Ping Liang, and Dexing Kong. Automatic 3d liver segmentation based on deep learning and globally optimized surface evolution. *Physics in Medicine and Biology*, 12 2016.

[53] Lianfen Huang, Minghui Weng, Haitao Shuai, Yue Huang, Jianjun Sun, and Fenglian Gao. Automatic liver segmentation from ct images using single-block linear detection. *BioMed Research International*, 7 2016.

[54] Weimin Huang, Ning Li, Ziping Lin, Guang-Bin Huang, Weiwei Zong, Jiayin Zhou, and Yuping Duan. Liver tumour detection and segmentation using kernel-based extreme learning machine. *34th Annual Internation Conference of the IEEE EMBS*, 7 2013.

[55] Alex Pappachen James and Belur V. Dasarathy. Medical image fusion: A survey of the state of the art. *Information Fusion*, 19:4–19, 2 2014.

[56] Jae Hyun Jeon, Jae Young Choi, Sihyoung Lee, and Yong Man Ro. Multiple roi selection based focal liver lesion classification in ultrasound images. *Expert Systems with Applications*, 40:450–457, 2013.

[57] Ola Jigin and Tova Linder. Organ detection and localization in radiological image volumes. 2017.

[58] Dagmar Kainmuller, Thomas Lange, and Hans Lamecker. Shape constrained automatic segmentation of the liver based on a heuristic intensity model. *Proc. MICCAI Workshop 3D Segmentation in the Clinic: A Grand Challenge*, pages 109–116, 2007.

[59] M. S. Khan, V. H. Mankar, G. Prashanthi, and G. Sathya. Skeletonization of 3d images using 2.5d and 3d algorithms. *2015 1st International Conference on Next Generation Computing Technologies(NGCT-2015)*, 2015.

[60] S. Krinidis, C. Nikou, and I. Pitas. 3d volume reconstruction by serially acquired 2d slices using a distance transform-based global cost function. *Hellenic Conference on Artificial Intelligence*, pages 390–400, 2002.

[61] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, pages 1097–1105, 2012.

[62] Paras Lakhani, Daniel L. Gray, Carl R. Pett, Paul Nagy, and George Shih. Hello world deep learning in medical imaging. *Journal of Digital Imaging*, 31:283–289, 2018.

[63] Juelin Leng, Guoliang Xu, and Yongjie Zhang. Medical image interpolation based on multi-resolution registration. *Computers and Mathematics with applications*, 4 2013.

[64] Tianjie Li, Yuanyuan Wang, Cai Chang, and Yongping Zheng. Color-appearance-model based fusion of gray and pseudo-color images for medical applications. *Information Fusion*, 19:103–114, 7 2014.

[65] Xuechen Li, Suhuai Luo, and Jiaming Li. Liver segmentation from ct image using fuzzy clustering and level set. *Journal of Signal and Information Processing*, 4:36–42, 8 2013.

[66] Seong-Jae Lim, Yong-Yeon Jeong, and Yo-Sung Ho. Automatic liver segmentation for volume measurement in ct images. *Journal of Visual Communication and Image Representation*, 17:860–875, 9 2006.

[67] Haibin Ling, S. Kevin Zhou, Yefeng Zheng, Bogdan Georgescu, Michael Suehling, and Dorin Comaniciu. Hierarchical, learning-based automatic liver segmentation. *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[68] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bram van Ginneken, and Clara I. Sanchez. A survey on deep learning in medical image analysis. 2 2017.

[69] Ziwei Liu, Xiaoxiao Li, Ping Luo, Chen Change Loy, and Xiaoou Tang. Semantic image segmentation via deep parsing network. 9 2015.

[70] Fang Lu, Fa Wu, Peijun Hu, Zhiyi Peng, and Dexing Kong. Automatic 3d liver location and segmentation via convolutional neural networks and graph cut. *International journal of computer assisted radiology and surgery*, 12(2):171–182, 2017.

[71] Suhuai Luo, Xuechen Li, and Jiaming Li. Improvement of liver segmentation by combining high order statistical texture features with anatomical structural features. *Engineering*, 5:67–72, 5 2013.

[72] Suhuai Luo, Xuechen Li, and Jiaming Li. Review on the methods of automatic liver segmentation from abdominal images. *Journal of Computer and Communications*, 2:1–7, 9 2014.

[73] Laurent Massoptier and Sergio Casciaro. Fully automatic liver segmentation through graph-cut technique. *Conference of the IEEE EMBS*, 2007.

[74] Laurent Massoptier and Sergio Casciaro. A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumours from ct scans. *European Society of Radiology*, 18:1658–1665, 3 2008.

[75] Maciej A. Mazurowski, Mateusz Buda, Ashirbani Saha, and Mustafa R. Bashir. Deep learning in radiology: an overview of the concepts and a survey of the state of the art. *arXiv preprint arXiv:1802.08717*, 2018.

[76] Arne Militzer, Tobias Hager, Florian Jager, Christian Tietjen, and Joachim Hornegger. Automatic detection an segmentation of focal liver lesions in contrast enhanced ct images. *2010 International Conference on Pattern Recognition*, 2010.

[77] Jan Hendrik Moltz, Lars Bornemann, Jan-Martin Kuhnigk, Volker Dicken, Elena Peitgen, Stephan Meier, Hendrick Bolte, Michael Fabel, Hans-Christian Bauknecht, Markus Hittinger, Andreas Jiebling, Michael Pusken, and Heinz-Otto Peitgen. Advanced segmentation techniques for lung nodules, liver metastases and enlarged lymph nodes in ct scans. *IEEE Journal of Selected Topics in Signal Processing*, 3, 2 2009.

[78] M. Murgasova, G. L. Estrin, M. Rutherford, D. Rueckert, and J. Hajnal. Distortion correction in fetal epi using non-rigid registration with laplacian constraint. *IEEE transactions on medical imaging*, 37(1):12–19, 2018.

[79] J. Alison Noble and Djamal Boukerroui. Ultrosound image segmentation: A survey. *IEEE Transactions on Medical Imaging*, 25, 8 2006.

[80] S. Osechinskiy and F. Kruggel. Slice-to-volume nonrigid registration of histological sections to mr images of the human brain. *Hindawi Publishing Corporation, Anatomy Research International*, 2011.

[81] Jialin Peng, Peijun Hu, Fang Lu, Zhiyi Peng, Dexing Kong, and Hongbo Zhang. 3d liver segmentation using multiple region appearances and graph-cuts. *The International Journal of Medical Physics Research and Practice*, 42, 11 2015.

[82] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual-information-based registration of medical images: A survey. *IEEE Transactions on Medical Imaging*, 22, 8 2003.

[83] Yingyi Qi, Wei Xiong, Wee Kheng Leow, Qi Tian, Jiayin Zhou, Jiang Liu, Thazin Han, Sudhakar K. Venkatesh, and Shih chang Wang. Semi-automatic segmentation of liver tumors from ct scans using bayesian rule-based 3d region growing. *MICCAI workshop*, 41(43):201, 2008.

[84] Olaf Ronneberger, Phillipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. 5 2015.

[85] Holger R. Roth, Le Lu, Amal Farag, Andrew Sohn, and Ronald M. Summers. Spatial aggregation of holistically-nested networks for automated pancreas segmentation. pages 451–459, 2016.

[86] Laszlo Rusko, Gyorgy Bekes, Gabor Nemeth, and Marta Fidrich. Fully automatic liver segmentation for contrast-enhanced ct images. *3D Segmentation in a Clinic: A grand challenge*, pages 143–150, 2007.

[87] Shraddha Sangewar, Atish A. Peshattiwar, Vilas Alagdeve, and Rupali Balpande. Liver segmentation of ct scan images using k means algorithm. *2013 International Conference on Advanced Electronic Systems*, 2013.

[88] Dieter Seghers, Pieter Slagmolen, Yves Lambelin, Jeroen Hermans, Dirk Loeckx, Frederik Maes, and Paul Suetens. Landmark based liver segmentation using local shape and local intensity models. 3 2017.

[89] M. Alper Selver, Aykut Kocaoglu, Guleser K. Demir, Hatice Dogan, Oguz Dicle, and Cuneyt Guzelis. Patient oriented and robust automatic liver segmentation for pre-evaluation of liver transplantation. *Computers in Biology and Medicine*, 38:765–784, 3 2008.

[90] Azrulhizam Shapi'i, Riza Sulaiman, Mohammed Khatim Hasan, Abd Yazid Mohd Kassim, and Hamsaini Abd Hamid. Applications of computer aided design in medical image technology. *Proceeding of the International Conference on Advanced Science, Engineering and Information Technology 2011*, 1 2011.

[91] Evan Shelhamer and Jonathan Long andTrevor Darrell. Fully convolutional networks for semantic segmentation. 5 2016.

[92] Karen Simonyan and Andrew Zissermani. Very deep convolutional networks for large-scale image recongition. *ICLR*, 4 2015.

[93] Mandeep Singh, Sukhwinder Singh, and Savita Gupta. An information fusion based method for liver classification using texture analysis of ultrasound images. *Information Fusion*, 19:91–96, 5 2013.

[94] Jean Stawiaski, Etienne Decenciere, and Francois Bidault. Interactive liver tumor segmentation using graph-cuts and watershed. 1 2008.

[95] Hiroyuki Sugimori. Classification of computed tomography images in different slice positions using deep learning. *Journal of Healthcare Engineering*, 2018.

[96] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. 9 2014.

[97] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. 12 2015.

[98] William Thong, Samuel Kadoury, Nicolas Piche, and Christopher J. Pal. Convolutional networks for kidney segmentation in contrast-enhanced ct scans. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, 6 2016.

[99] J.K. Udupa, V. R. Leblanc, Y. Zhuge, C. Imielinska, H. Schmidt, L. M. Currie, B. E. Hirsch, and J. Woodburn. A framework for evaluating image segmentation algorithms. *Comput Med Imaging Graph*, pages 75–87, 3 2006.

[100] Bram van Ginneken, Tobias Heimann, and Martin Styner. 3d segmentation in the clinic: A grand challenge. 2007.

[101] Eva van Rikxoort, Yulia Arzhaeva, and Bram van Ginneken. Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching. *Proceedings of the MICCAI Workshop*, 3:101–108, 2007.

[102] Akshay P. Vartak and Dr. Vijay Mankar. Morphological image segmentation analysis. *International Journal of Computer Science and Applications*, 6, 4 2013.

[103] R. Vivanti, A. Ephrat, L. Joskowitcz, O. A. Karaaslan, N. Lev-Cohain, and J. Sosna. Automatic liver tumor segmentation in follow-up ct studies using convolutional neural networks. *Proc. Patch-Based Methods in Medical Image Processing Workshop*, 2, 2015.

[104] Jason Wang and Luis Perez. The effectiveness of data augmentation in image classification using deep learning.

[105] Damon Wong, Jiang Liu, Yin Fengshou, Qi Tian, Wei Xiong, Jiayin Zhou, Yingyi Qi, Thazin Han, Sudhakar K. Venkatesh, and Shih-Chang Wang. A semi-automated method for liver tumour segmentation based on 2d region growing with knowledge-based constraints. *MICCAI workshop*, 41(43):159, 2008.

[106] Weiwei Wu, Zhuhuang Zhou, Shuicai Wu, and Yanhua Zhang. Automatic liver segmentation on volumetric ct images using supervoxel-based graph cuts. *Computational and Mathematical Methods in Medicine*, 3 2016.

[107] Varduhi Yeghiazaryan and Irina Voiculescu. An overview of current evaluation methods used in medical image segmentation. 11 2015.

[108] Oussema Zayane, Besma Jouini, and Mohamed Ali Mahjoub. Automatic liver segmentation method in ct images. *Canadian Journal on Image Processing and Computer Vision*, 2, 12 2011.

[109] Xing Zhang, Kexin Deng, Yongfang Wu, and Xiuli Li. Automatic liver segmentation using a statistical shape model with optimal surface detection. *IEEE Transactions on Biomedical Engineering*, 57, 10 2010.

[110] Jiayin Zhou, Wei Xiong, Qi Tian, Yingyi Qi, Jiang Liu, Wee Keng Leow, Thazin Han, Sudhakar K. Venkatesh, and Shih chang Wang. Semi-automatic segmentation of 3d liver tumours from ct scans using voxel classification and propagational learning. *MICCAI workshop*, 41:43, 2008.

[111] Qikui Zhu, Bo Du, Baris Turkbey, Peter L. Choyke, and Pingkun Yan. Deeply-supervised cnn for prostate segmentation. pages 178–184, 2017.