UNIVERSITY OF TWENTE.

ANTONI VAN LEEUWENHOEK
NEDERLANDS KANKER INSTITUUT

UNIVERSITY OF TWENTE

DEPARTMENT OF TECHNICAL MEDICINE

Automated vascular region segmentation in ultrasound to utilize surgical navigation in liver surgery

Chairman & clinical supervisor:
prof. dr. T.J.M. RUERS

Daily supervisor:
dr. J. NIJKAMP

Author:
Bart R. Thomson

Technical supervisor:
dr. ir. F. VAN DER HEIJDEN

Mentor:
drs. A. LOVINK

External member:
dr. I.E. ALLIJN

A thesis submitted for the degree of

Master of Science in Technical Medicine
from the Faculty of Science and Technology (TNW)

August 21, 2019

Abstract

The liver is a common location for primary cancer and metastatic disease, often originating from colorectal, lung, breast and pancreatic tumors. Nowadays, surgical resections, when compared to other treatment plans, provide the best patient outcome for various types of liver malignancies. Due to high complexity and inter-patient variability of underlying hepatic vascular anatomy, planning and execution of safe resection is challenging in surgery. Currently, US is the only imaging modality that is widely accepted and integrated into a surgical workflow, making it the most suitable imaging modality for intraoperative visualization of hepatic vasculature. Despite many advantages of intraoperative ultrasound, it is still a primary 2D imaging modality, which complicates precise localization of each 2D image in 3D for a surgeon. Automatic registration with preoperative imaging would provide great value in determining a resection plan. In this thesis, the goal was to realize automatic registration between pre- and intraoperative imaging.

For that purpose, a 3D U-Net is trained to automatically segment intraoperative vasculature. Training on a combined dataset of stacked 2D and 3D imaging gives the most promising results, with a Dice of 0.773 ($\pm$ 0.10) and a Jaccard index (JI) of 0.640 ($\pm$ 0.12), comparing to an inter-observer variability of respectively 0.879 ($\pm$ 0.02) and 0.785 ($\pm$ 0.02). The centerline of this intraoperative segmentation is then registered with a preoperative, semi-automatically segmented, vasculature model. An initial registration is performed based on the US probe orientation and one point translation to crop a similar point cloud from the preoperative model, as is segmented intraoperatively. With visually successful registrations we acquire an automatic target registration error (TRE) of 12.29 ($\pm$ 4.93), however, 55 % of the registrations fail expectantly due to a relatively big cropping volume with respect to the US information that is acquired. Manually adjusting the cropping volumes reduces the TREs over all volumes from 47.32 ($\pm$25.71) to 25.66 ($\pm$10.48).

In conclusion, we demonstrate a fast (69.74 $\pm$ 14.6 seconds) deep learning based hepatic vasculature registration pipeline. Given that the US acquisitions do not contain the vena cava or gallbladder, and span a large part of the hepatic vasculature, our approach looks promising. Further optimization of automatically acquiring similar point clouds is expected to stimulate the adaptation of surgical navigation on a regular basis.

Acknowledgements

The research line of which this thesis is part, in the clinical setting of the NKI-AvL at Amsterdam was set out by prof. dr. Theo Ruers, to whom I express my gratitude for the opportunities received during my graduation research. He kept the clinical goal in mind, was approachable as a supervisor, and kept a critical mindset on where we would like to go intraoperatively.

The overall technical perspective was retained by dr. Jasper Nijkamp from the NKI-AvL, who guided my thinking process in the right direction and steered where needed. The weekly (Skype) meetings were great as a fixed moment of support.

Additional technical support was given by dr. Matteo Fusaglia, who enforced my critical thinking and was of great support during the development and writing process.

Hands-on clinical expertise was brought to my fingertips by Jasper Smit, MSc. Working together, he showed me the ropes of hands-on surgical navigation and all difficulties that are brought with it. Along my technical work, he was of great value educating my clinical self.

Technical guidance in terms of what is relevant and demarcating what is required to achieve certain goals was retained by my supervisor from the University of Twente, dr. ir. Ferdi van der Heijden. My thanks for his guidance in keeping perspective on the best way to achieve my research goals.

Besides the technical and clinical supervision, Annelies Lovink, MSc, mentored me excellently in developing professional soft skills, as well as growing on a personal level.

Furthermore, my colleagues, of both the radiotherapy and surgery department, at the NKI-AvL made it an enjoyable and instructive year filled with good coffee (tea) breaks, educational sessions and fun friday afternoons.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Clinical background

In the Netherlands, approximately 830 patients are diagnosed with primary liver cancer each year [1]. Liver diseases, such as the final stage of liver fibrosis, liver cirrhosis, radically increase the risk for hepatic cancer. Even though a malignant liver mass most likely represents a metastatic hazard instead of a primary hepatic malignancy [2], hepatocellular carcinoma (HCC) is the most common primary cancer of the liver, arising mainly in patients with chronic liver disease [3]. Furthermore, HCC is the third-most common cause of cancer-related deaths and the sixth-most common cancer worldwide [4]. A chronically damaged liver commonly gives rise to the molecularly and genetically highly heterogeneous group of cancers comprised by HCC [5].

Moreover, the biggest group of liver lesions are metastasis originating from colorectal cancer (CRC). In 2017 about 13.800 patients were diagnosed with CRC in the Netherlands [1], of which half face liver metastases [6]. In 30-70% of the cases liver metastases will develop in patients with advanced CRC of which 25% have metastases at presentation [7], causing two thirds of the deaths in CRC patients [6]. Thus, primary tumor staging routinely analyzes the liver and its lesions. After lymph nodes, the liver is the organ most likely to be invaded by colorectal liver metastases (CRLM), therefore regular imaging is necessary [2]. Other primary hazards for liver metastases include lung, breast, stomach and pancreatic cancer [5, 8].

1.1.1 Diagnosis

In hepatic cancer, imaging is divided in surveillance and diagnostic imaging of a previously discovered hazard. Both in metastatic deposits and in primary tumors, accurate detection of malignant hazards is crucial in patient management [9]. B-mode ultrasound (US) is used as the first diagnostic modality for patients with an elevated risk, i.e. chronic hepatitis B or cirrhosis. Imaging strategies should include lesion characterization since benign lesions are very common [10, 11]. The fast, non-invasive and cost-effective properties of US make it the primary screening test to examine the liver parenchyma, which can be done as often as needed [12]. A possible lesion can be localized in the liver, vascularity within and around the lesion can be monitored with color Doppler and abnormalities can be characterized as cystic or solid [13, 14]. Possible thrombosis or vascular infiltration can also be determined, without anesthesia and no downsides for frequent follow up.

Conventional grey scale US has a relatively poor sensitivity to depict a hepatic metastasis (53-77 %) [15, 16] compared to MR imaging and contrast-enhanced CT (80-95 %) [17]. The relatively small difference in backscatter between the hepatic parenchyma and the lesion can result in challenging contrast differentiation in US [18]. US contrast agents such as microbubbles of air or low solubility gasses stabilized by a lipid, increase echogenicity of the liver as they accumulate within the normal parenchyma, thus increasing the visibility of critical structures and hepatic metastases [19, 20]. However, US-based screening is sub-optimal when cirrhosis is present in terms of sensitivity and specificity [21]. Moreover, US is highly operator-dependent and sensitivity can be as low as 20% in sub centimeter lesions [16, 22, 23]. Also, MRI or CT is preferred for precise relation with surrounding critical anatomy, further characterization and in case of malignant neoplasms, detection of associated metastatic disease. Thus, patients with an abnormal liver on US often

undergo contrast-enhanced CT or MR examination when they are diagnosed with cirrhosis [21].

CT offers the ability to study the entire liver and its surroundings whilst offering the best spatial resolution in a single-breath hold [22]. Iodine contrast is routinely used in liver imaging, improving the contrast-to-noise ratio between normal liver tissue and focal liver lesions, thus aiding detection. Based on enhancement patterns during various phases of contrast circulation, contrast media help to characterize liver lesions [24]. At the same time it provides useful information about vascular anatomy, quality of the liver parenchyma, partial and total liver volumes and many other clinical parameters [25]. However, it also exposes the patient to ionizing radiation. Sahani et al. [26] found that MRI offers greater specificity and sensitivity than CT, especially in <1cm lesion detection. Also, diffusion-weighted MRI has shown to improve diagnostic accuracy because of the proton diffusion differences between malignant and benign tissue. The capacity of MRI for detecting and characterizing small lesions has further improved due to recent introduction of new liver-specific MRI contrast agents [26]. In the NKI-AvL, multi-phase MR sequences with liver specific gadolinium-based contrast agent gadoxetic acid (Gd-EOB-DTPA, Primovist) are used in the diagnostic MR protocol. Although CT and MR provide superior imaging compared to US, they are rarely used intraoperatively.

1.1.2 Anatomy and pathology

Nowadays, even tumors less than 1 cm can be characterized on CT and MRI techniques [27, 28], easing the removal of smaller lesions. Guidelines state that CT is not sufficient enough for lesions smaller than 1 cm, therefore MRI is recommended for those lesions [29]. Knowledge of segmental anatomy as described by Couinaud [30] is essential when inspecting the relation of the tumors to the liver vasculature and is shown in Figure 1.1. This classification divides the liver into eight functionally independent segments, where each segment has its own biliary drainage and vascular in and outflow. The center of each segment is marked by a branch of the bile duct, portal vein and hepatic artery. Outflow of each segment happens through the hepatic veins in the periphery. Whilst surgery's primary goal is radical resection, the segmental branches have to be identified in order to preserve essential liver tissue, of which the borders are difficult to determine intraoperatively [31]. The regenerative ability of the liver has shown an operative mortality of less than 5%, for resections up to 80% of the liver [32–36].

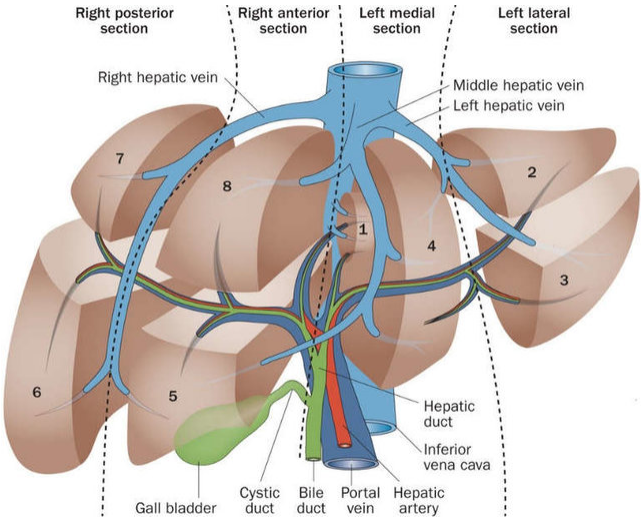


Figure 1.1: Classification as defined by the Couinaud [30] model (adapted from [37]).

1.1.3 Treatment

In patients with CRLM, there are three main options for treatment: radiotherapy, surgical resection or systemic therapy. To prolong survival, liver resection is the treatment of choice as it currently provides the best prognosis [6, 38]. However, tumor location, major vascular contact, insufficient liver remnant, bilaterality or patient co-morbidity frequently impede with the resection feasibility. The majority (70-80%) of patients with liver lesions were considered unsuitable for

resection in recent years at diagnosis [39]. Nowadays, due to the significant improvements in surgical techniques, anesthesia, chemotherapy, imaging modalities and the expansion of resectability criteria among surgeons, a greater number of patients undergo surgery [39]. Due to these factors, the vast majority of patients will undergo liver resection after downstaging of the lesions with alternative treatment. While the most optimal treatment option is surgical resection, not all lesions can be removed surgically. Local treatment options such as microwave ablation (MWA) and radio-frequency ablation (RFA) are becoming more common and can be performed percutaneously by intervention radiologists [6, 40, 41]. A combination of surgical intervention and the aforementioned techniques, performed on a daily basis in the NKI-AvL, can also be performed when at least one lesion appears unresectable during surgery. Intraoperative ablation techniques heavily rely on optimal localization and visualization of the target lesion, a satisfactory resection margin can solely be achieved by ensuring accurate needle placement in the center of the lesion.

External radiation therapy exceeds the tolerance of non-tumorous liver and therefore has had limited success in the past [42–44]. In the last two decades, image-guidance has improved due to technological developments, leading to increased accuracy of dose delivery, allowing for more effective focused high-dose liver radiotherapy [45, 46]. However, the best survival rate of patients is achieved with surgery, which is prone to various criteria and rules for partial resection, limiting the operable patients to 50% [47]. Complex liver surgeries can be aided by detailed knowledge of the patient-specific vasculature and biliary structures, simultaneously contributing to successful surgical resection and higher preservation of functional liver tissue [48–50]. In the NKI-AvL, the surgeon is often provided with a preoperative 3D model, visualizing patient-specific anatomy based on a preoperative contrast-enhanced MR scan. Information based on a preoperative model is of added benefit when a patient underwent chemotherapy with a good response and it is therefore difficult to localize lesions. Moreover, it is beneficial when patients present with centrally located lesions (Figure 1.1 Couinaud segments 4, 5 or 8) or unusual arterial or biliary tract anatomy. It however remains difficult to optimally use the model during live surgery, due to the high natural flexibility and mobility of the liver [51, 52].

1.2 Technical background

Another means of using imaging as support during an intervention is seen in the percutaneous approach to liver lesions. Several clinical applications, such as ablations and biopsies, utilize US-guided navigation [53]. The user is provided with a more detailed view of the anatomical structures surrounding the lesion, by means of tracking the US transducer and biopsy or ablation tools. Although intraoperative application for liver surgery has not made an introduction into regular practice, this principle is used on a regular basis in interventional radiology to facilitate ablation and biopsy guidance. In the NKI-AvL the EPIQ7 US platform with PercuNav software (Philips, The Netherlands) lets clinicians register different diagnostic scans with live percutaneous US imaging. Currently, US is the most widely used method of guidance for percutaneous ablations of carcinomata in the liver [54, 55]. However, as elaborated on in section 1.1.1, cross-sectional modalities (e.g., MRI and CT) are less limiting. In the liver specifically, US can lead to insufficient distinction from surrounding tissue due to isoechogenicity. On pretreatment US, Kim et al. [56] reported that 25.3% of target tumors were undetectable, with distance between the diaphragm and tumor, tumor size and liver cirrhosis as significant factors. Contrast enhanced US has been reported to enhance lesion conspicuity and findability compared to US [57, 58]. However, it is still reported as a major cause of mistargeting [59]. Accordingly, it is of interest to combine advantages of different imaging modalities, which can be achieved by image-guided surgery.

1.2.1 Image-guided surgery

Any surgery using tracked surgical instruments combined with advanced imaging to monitor, localize, control and target procedures is spanned by the concept of image-guided surgery (IGS). Imaging complements direct visualization and procedures to allow for better targeting and improved outcome. Prior to a procedure, routine diagnostic imaging is performed on the patient. The acquired imaging is converted into 3D images and processed into a 3D model representing the patient's anatomy. This 3D information can then be used for preoperative planning, and after registration of the 3D model of the preoperative imaging to the intraoperative organ position, it enables and guides intraoperative surgical decision making. Therewith, the tracking of surgical

instruments, during the surgical procedure, aims to minimize complications and allow for accurately navigating towards targeted tissue or lesions. The orientation and position of the tracked surgical instruments are mapped to an artificial space or 3D scene, where their motion is precisely visualized with respect to the patient's anatomy. Navigating in 3D helps the surgeon's visualization inside the body in relation to the actual surgical instrument's position. Image-guided navigation helps surgeons to perform surgery accurately and minimizes guesswork that is often involved in complicated procedures.

In order to establish the spatial relationship between the artificial and surgical field, the images have to be registered. Usually, specific points in the imaging dataset are matched with the corresponding point in the surgical field. To achieve registration, a minimum of three points should be matched or registered [60]. Section 3.4 further elaborates on this.

Tracking methods

In IGS, tracking is the process of making localization possible in the patient's coordinate system, where optical and electromagnetic (EM) tracking are the two main methods used. Active optical trackers use several video cameras to triangulate the 3D position of flashing LEDs, which can be mounted on any surgical instrument. Passive optical tracking uses infrared light reflectance to calculate the precise location of the instrument. These systems are wireless but require a direct line of sight between the camera and sensors. EM tracking circumvents this limitation by placing small electromagnetic sensors on instruments in a pulsed magnetic field of known geometry, allowing detection of position and orientation in 3D space, whilst being virtually transparent to the surgeon, when compared to optical trackers. EM tracking systems serve the purpose to provide a 3D Cartesian coordinate system of markers attached to instruments and patient anatomy. A drawback of EM tracking is that large ferromagnetic objects can distort the EM generated field and diminish accuracy. EM tracking is preferred in the NKI-AvL as this system does not require a clear line of sight of optical imaging systems. Feasibility of EM tracking during surgery has been shown by Nijkamp et al. [61] in the NKI-AvL. However, the use of skin-bound EM-sensors was concluded as the major error source for inaccuracies. An overview of the components that are used in the current study, to realize EM tracking, is given in section 2.3.

1.2.2 Clinical application of surgical navigation

As aforementioned, intraoperative US is routinely used in liver surgery for tumor localization to assist resection [62]. A discrepancy between the preoperative imaging modality and the US imaging, possibly hindered by isoechoic or vanishing lesions, might lead to inconclusive observations [63]. US-based navigation can improve adequate lesion localization in surgery. In current practice, all preoperative imaging information is processed by the surgeon, by mentally reconstructing the preoperative information guides the surgery. The preoperative information could gain importance when it is available during surgery and is directly related to the surgical tool positions, in real-time. Surgical tools can be registered to an image, and then be used to display orthogonal views of the patient's preoperative image. Additionally, damage to vital structures can be prevented. Structures which are poorly visible on US can be seen on preoperative imaging [63]. Navigation performed in a three dimensional, preoperative imaging based environment facilitates better assessment of ablation zones and resection planes during open surgery. Traditionally, this is done by cone beam computed tomography (CBCT), allowing for 3D visualization of both the organs at risk and the target volume, but also introduces a non-negligable additional dose to the patient [64]. Therefore, US imaging appears to be an interesting alternative since it is non-irradiating and non-invasive. Thus, it does not imply any additional risk for the patient [65].

Current systems

In recent years, several groups have developed US-based navigation systems in the field of liver surgery. Van Belle et al. [66] developed and evaluated a system with optical sensors for navigated liver segment resections using intraoperatively acquired 3D US data. However, they did not register to preoperative imaging. Fusaglia et al. [67] use a 3D volume, reconstructed from 2D laparoscopic US images and aligned it with a CT volume by means of a stochastic optimizer, where they performed accuracy assessment on a phantom. Penney et al. [68] introduced manual annotation-based US-MR registration, where they acquired a TRE of 9.95 ± 3.83 mm using ICP.

The average time taken for a registration was 300 seconds. Another approach is presented by Haque et al. [69], which achieves high accuracy, but requires breath hold. Weon et al. [70] try to mitigate breath holding by presenting a real-time registration method. However, they have a setup that is significantly different from the common clinical workflow. Wei et al. [71] show that, by means of automatically segmenting liver vasculature and parenchyma, they are able to achieve a TRE of 1.97 ± 1.07 mm. They report that limited vascular US information has a significant effect on their accuracy. All of the aforementioned studies base their findings on imaging performed percutaneous or on a phantom.

Commercially, multiple systems are presented in literature which realize registration of pre- and intraoperative imaging. Banz et al. [72] initially developed a system as part of a research project, but subsequently evolved it to a commercially available platform (CAS-One, CAScination AG, Bern, Switzerland). This guidance system utilizes optical tracking and explored three different registration approaches which developed over time: liver surface landmark based, surface combined with parenchyma landmarks and an US-based volume registration. They show feasibility of the system in more than 65 patients with an ultimate accuracy of 4.5 ± 3.6 mm in 22 patients with a combination of landmark and US-based volume registration. Furthermore, another liver navigation system; Explorer (Analogic, Inc., Boston, MA), has shown an error of 2-6mm in target surface areas [51]. More recent studies report usage of this system in open procedures and ablation guidance, but do not report on accuracy [73, 74]. An overview of the aforementioned literature is presented in Table 1.1.

The commercial solutions rely on landmark selection for registration purposes. However, even for an expert it can be challenging to select the exact corresponding points on both modalities. Furthermore, the aforementioned software is based on optical tracking systems, requiring a direct line of sight, which might not always be possible in an operating theatre setting. Moreover, it is not well suited for laparoscopy as the tip of many instruments can flexibly move in relation to the markers on the device. Due to the extensive size of an optical tracker it is very challenging to properly secure an optical tracker to the liver.

CustusX (SINTEF, Trondheim, Norway) [75] is an image-guided therapy research platform allowing the user to implement custom functionality to circumvent restrictions imposed by the aforementioned commercially available software. In addition to optical tracking, it supports EM tracking, allows for a personalized graphical user interface and has been used in navigation for liver phantom use and research during liver and abdominal surgery [76, 77]. The user is able to extract an acquired US volume which can then be processed as the user wishes. The vasculature that is present in the US volume can be used to align the preoperative coordinate system with the US coordinate system, by means of registration. A necessary step in this process is segmentation of vasculature in both modalities, which is elaborated on in the following section.

Table 1.1: Overview of related work.

Authors	Accuracy (mm)	Time (s)	Type	Applied registration	Success rate
Fusaglia et al. [67]	8.2 ± 1.63	720	3D-3D	Phantom	
Penney et al. [68]	9.95 ± 3.83	300	3D-3D	Percutaneous	
Haque et al. [69]	3.88 ± 1.38	40	3D-3D	Percutaneous	73%
Weon et al. [70]	2.80 ± 1.44	0.06	3D-3D	Percutaneous	56%
Wei et al. [71]	1.97 ± 1.07	0.5	2D-3D	Percutaneous	80%
Banz et al. [72]	4.5 ± 3.6		3D-3D	Intraoperative	
Cash et al. [51]	2 - 6		3D-3D	Intraoperative	

1.2.3 Medical image segmentation

To detect vessels or edges in medical imaging, intensity and gradient features have traditionally been used. Despite encouraging results, these techniques are directly influenced by the image quality of US imaging [78, 79]. Well known 3D techniques such as region growing and deformable models are often used in the segmentation of vessel trees in other modalities than US [80, 81]. However, typical shadowing, speckle of US images and missing boundaries due to image orientation make it difficult to perform accurate segmentation [82]. Manual intervention is usually required at some stage in general purpose segmentation methods, which impedes with an

intraoperative workflow. These traditional segmentation methods often lead to lower accuracy and are thus considered to be limited. Features are included, but the model is not able to influence feature definition. In recent years they have become less popular [83], due to the manually defined features when compared to machine learning models. Linde [84] argues that artificial intelligence (AI) systems need to acquire their own knowledge, by extracting patterns from raw data. Self acquisition of behaviour imitating human knowledge, by a computer, is referred to as machine learning. This involves real world knowledge in problem solving leading to computers making decisions that appear subjective [85]. Machine learning can be used to discover the mapping from representation to output. Deep learning simplifies this even more by learning highly abstract image features [86].

In the past, AI projects were sought in logistic regression, handcrafted features and logical knowledge about the world. Deep learning can circumvent the cumbersome and time-consuming task of manual intervention when segmenting live data. Also, it has demonstrated to increase performance in image segmentation and classification tasks, providing aid in many computer vision tasks [87].

1.2.4 Convolutional neural networks

The growth in popularity of deep learning models mostly is due to hardware improvements, allowing for faster training of convolutional neural networks (CNNs). But also due to data availability and accessibility. Employing CNNs in the processing of volumetric data has taken a lot of effort, 2D CNNs have been used to aggregate 3D features in adjacent slices [88], multi-view planes [89] or orthogonal planes [90]. 2D networks benefit from lower computational costs and thus faster processing, whereas 3D networks benefit from information of an added dimension, potentially increasing accuracy.

LeCun et al. [91] introduced the concept of CNNs, as we know it today, based on a self-organizing artificial network [92] that was able to recognize patterns regardless of spatial shift. Despite successes in industrial technology, CNNs were largely forsaken in automatic image recognition tasks until the ImageNet competition in 2012, where Krizhevsky et al. [93] halved the error rate of traditional approaches. They showed that a large, deep CNN is capable to achieve record breaking results using purely supervised learning on a highly challenging dataset. Three main types of layers can be identified when building a CNN architecture: convolutional layer, max pooling layer and the fully connected layer. The convolutional and max pooling layer are discussed in the following sections.

Convolutional layer

The core building block with the most computational effort is the convolutional layer. Its parameters consist of a collection of learn-able filters, which are spatially small (i.e. 3x3 pixels) but generate output whilst sweeping over the whole input image. During the forward pass each filter slides across the input volume and computes dot products between the input and the entries of the filter at any position (Figure 1.2). The composition of the filters is initially determined random, but is optimized over the course of training a network. The necessary amount of filters often depends on the complexity of the task, whereas all filters produce output and thus rapidly increase the number of weights and biases. These filters are then stacked along the depth dimension before they are passed into the next layer. Two other important parameters in the design of a CNN are the stride and zero-padding. First, the stride determines the movement with which the filter slides over the image, a stride of 1 corresponds with a movement of 1 pixel. Bigger strides would lead to a smaller output volume. Second, it might be convenient to pad the input volume with zeros around the border, allowing for control over the spatial size of the output volumes. It is most common to apply a zero-padding of 1 with a stride of 1, so that the output volume is the same size as the input volume (with a filter of 3x3). When processed by a convolutional layer, the feature map is passed through an activation function, which nowadays most commonly is the Rectified Linear Unit (ReLU), presented in Figure 1.3. When processed by a ReLU activation function, only positive values are propagated further through the network, as described by the formula $ReLU(z) = max(0, z)$. This function is so commonly used because it has proved faster to train, whilst often also improving discriminative performance [94].

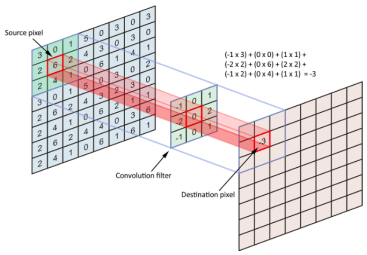


Figure 1.2: Convolutional layer (adapted from [95]).

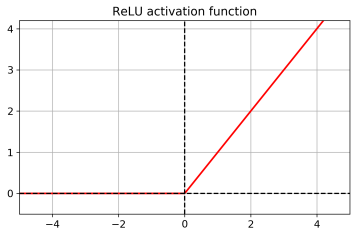


Figure 1.3: Activation function (adapted from [96]).

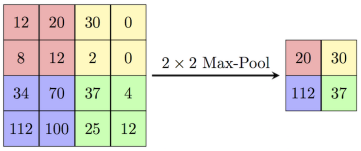


Figure 1.4: Max pooling layer (adapted from [95]).

Max pooling layer

Convolutional layers with many filters lead to a rapid increase in trainable parameters and thus computational effort and time needed to train the network, whilst also reducing over-fitting. Max pooling layers are introduced in-between successive convolutional layers to reduce the load caused by the increasing amount of trainable parameters. A max pooling layer is independently applied to every feature map, where it spatially resizes the feature map using a max operation within the field of view. Only the highest value from within the field of view is passed on to the succeeding layer. Max pooling layers with filters of size 2x2 and stride 2 are most commonly used, reducing the number of activations by 75% (Figure 1.4).

Network training

Feeding data through the aforementioned layers comprises training of a neural network. The purpose of training a neural network is to minimize the difference between the networks' prediction and the ground truth, especially when predicting on new data. This can be achieved by using the optimal set of values for the network weights and biases, which are computed by means of back-propagation. Back-propagation is an algorithm which computes the partial derivative of the cost function that is chosen to minimize the difference between the networks' prediction and the ground truth. During the forward pass of a single iteration all weights (values in filters in case of a CNN) are applied (first iteration is randomly initialized). During the backward pass weights that have contributed the most to the overall error will undergo the largest change, they are identifiable by larger derivation values.

The three most important steps in back-propagation, the forward pass, the backward pass and the weight updates respectively are expressed in equations 1.2 - 1.4 [97]. In these equations, the state of layer k for pattern p is denoted by $X_p(k)$, with the global state of the network for pattern p denoted by X_p. The non-linear transformation associated with layer k is denoted with F_k, which is typically an activation function. The vector of total input to units in layer k (weighted sums) is denoted by $A_p(k)$, of which the value is given by equation 1.1. Here, layer $k-1$ is connected to layer k through a connection matrix $W(k)$.

$$A_p(k) = W(k)X_p(k-1) \quad \forall p \in [1, P] \tag{1.1}$$

$$X_p(k) = F_k(A_p(k)) \quad \forall k, p \in [1, N][1, P] \tag{1.2}$$

Equation 1.3 gives the usual method for computing the gradient variables Y by backward propagation. Whereas equation 1.4 presents the weight updates, where we are looking for a minimum in the output cost function with respect to W. According to literature [97], the method of steepest descent is the most common and easiest method to do so. Hence, it is presented, where λ is the step size. In a CNN, the individual filter values are the weights, which are updated based on the gradient determined with the derivative, leading to an improved prediction with the next iteration.

$$Y_p(k) = \bigtriangledown F_k(A_p(k))W^T(k+1)Y_p(k+1) \quad \forall k, p \in [0, n-1][1, P] \tag{1.3}$$

$$W(k) \leftarrow W(k) + \lambda \sum_{p=1}^{P} Y_p(k)X_p^T(k-1) \quad \forall k, p \in [1, N][1, P] \tag{1.4}$$

1.2.5 3D modeling

Given what is presented in the foregoing sections, an essential development in most surgical fields has been brought by the advancement of radiological imaging and segmentation techniques. CT and MR imaging make it possible to visualize the size and location of lesions in organs, subsequently allowing for planning of surgery. Whereas deep learning allows for automatic delineation of surgically relevant structures. Without knowledge of important structures related to the lesion or major vessels, surgery cannot be performed safely and curatively at the same time [31]. Significant improvements have been made in the surgical equipment to resect liver tissue, as well as in post and intraoperative treatment [31]. However, the largest contributor has been the improvement of imaging techniques, with intraoperative US further improving the localization of lesions [98]. Open surgery is performed in 3D, whereas resection line planning may cause difficulties when based on 2D imaging, even though 3D information is present. Creating a 3D model bridges the gap between a 2D mental representation of the surgeon and a 3D visualization. Data segmentation is a prerequisite to construct a visualizable 3D model. It should, however, be as automatic as possible without compromising the accuracy in an intraoperative setting. In preoperative segmentation, speed is less important and manual corrections can still be applied. An example of a preoperative model as it is used during surgery is presented in Figure 1.5.

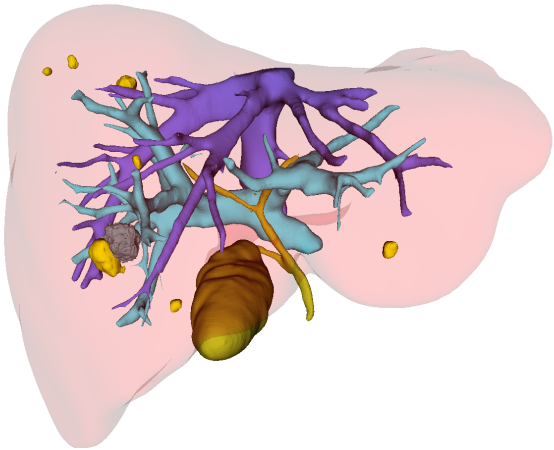


Figure 1.5: Example 3D model, hepatic vein depicted in purple, portal vein in light blue, the gallbladder in yellow brown, and the lesions in yellow.

1.2.6 Registration

Based on the (automatically) extracted models, one can combine the information from both modalities. Here, it is important to realize that the preoperative images and the intraoperatively acquired US data have their own coordinate systems. Registration is the process resulting in a geometrical mapping between data represented in different modalities (coordinate systems). Registering preoperative imaging to intraoperative imaging gives more insight to the anatomical structures and enables the possibility of surgical navigation. The warping of a source volume (preoperative imaging) to align with a fixed volume (intraoperative US) is the typical formulation of volume registration. In the preoperative imaging the xy-plane is defined as the transversal plane, with the z-axis oriented from the cranial to caudal direction. Both the preoperative and intraoperative coordinate system are Cartesian. However, there is a difference in orientation of the z-axis in the EM-tracked field, it is orthogonal to the opening of the generator and thus depending on its orientation relative to the patient. Figure 1.6 presents the steps that are necessary to perform a correct registration. A reference sensor can be fixed in place near the liver in order to have a point of reference as close to the region of interest, eliminating possible EM inaccuracies. Another sensor is clipped and calibrated on the US probe. The calibration transformation between the 2D US plane and its location and orientation in space is defined as T_{cal}. The transformation between the tracked probe and the reference sensor is established by continuously updating T_{track}, resulting

in transformation matrix T_{tot} (Formula 1.5). Registration is finalized by applying transformation matrix T_{reg}, in order to express the US coordinates in MR coordinates. When applied to a point US_p, position MR_p in the preoperative model can be expressed as presented in Formula 1.6.

$$T_{tot} = T_{reg}T_{track}T_{cal} \tag{1.5}$$

$$MR_p = T_{tot}US_p \tag{1.6}$$

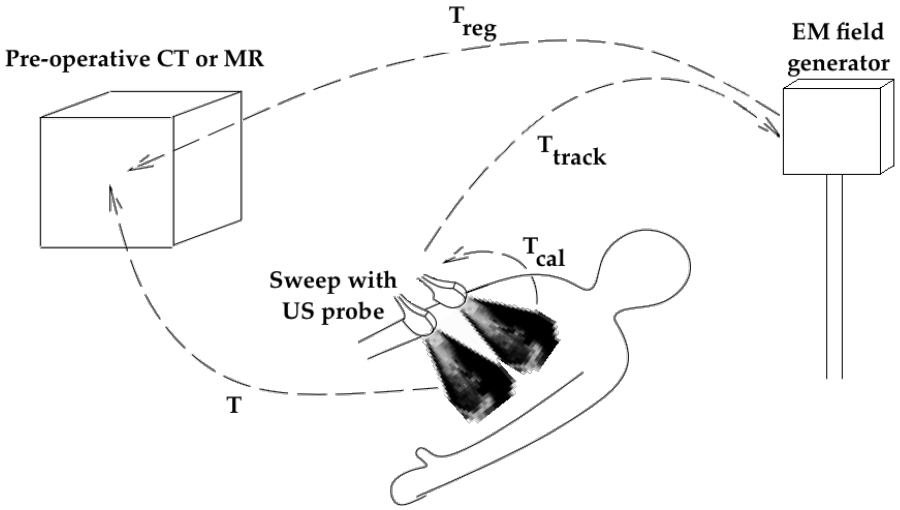


Figure 1.6: Overview of coordinate systems and transformations (modified from [99]).

In many computer vision tasks registration of point sets is a key component, where the goal is to recover the transformation that maps one point set to the other. Often the points in a point set are features extracted from an image, such as boundary points, locations of corners or salient regions. Practically, three main desirable properties can be identified with registration algorithms:

- Robustness to outliers, noise and missing points resulting from sub-optimal feature extraction and image acquisition
- Accurate modeling of the transformation necessary for aligning of the point sets using manageable computational complexity.
- The ability to handle dimensionality of point sets (2D/3D)

The variable in our overview (Figure 1.6) representing the registration from one coordinate system to another is T_{reg}. Usually T_{reg} is performed in a rigid or non-rigid manner, where a rigid transformation solely allows for rotation and translation. One of the most commonly used registration algorithms is the iterative closest point algorithm [100] (ICP), it iteratively determines the sum of distances between all corresponding points in a point cloud and can either be rigid or affine. The transformation matrix is iteratively adjusted and ultimately used when a global minimum is found. Another registration algorithm, which has proven to outperform ICP [101] in a registration task between vascular centerlines, is coherent point drift [102] (CPD). It is used in this study and further elaborated on in section 2.7.1. An affine transformation is the most simple non-rigid transformation, which also allows for anisotropic skews and scaling whilst preserving parallel lines. In order to improve diagnostics and monitoring there is a demand for finding better ways to fuse and compare corresponding images in US technology.

Deformable image registration can be applied in addition to rigid image registration techniques [103–105] in order to find a more accurate registration. Different approaches have been established within the last few years for deformable image registration. [103] present an overview of recent methods for US registration, including several deformable approaches. In general it is assumed that muscular activity, external forces or weight displacements cause elastic movements of tissue. Therefore, all models have to preserve tissue topology and represent a physiologically plausible situation when they are applied. The three main approaches include: knowledge

based transformations, models derived from interpolation and geometric models based on physical models [106]. Deformable registration strategies generally comprise a slow deformable transformation with many degrees of freedom preceded by affine transformation for global alignment. It is, however, difficult due to spatial and temporal variability between both modalities to provide a model that is sufficiently robust for clinical use [107]. Moreover, it typically requires significant processing time as well as the use of computationally intensive approaches. Multiple machine learning approaches [108–111] that try to solve this challenge, based on labeled data, argue that faster models can be developed while maintaining clinical robustness. Hu et al. [112] propose a weakly supervised network with only sparse annotations in registering preoperative MR images to transrectal US. Other deep learning based approaches are illustrated by [109, 113, 114].

Multidimensional point sets in real world problems are common and most registration algorithms are well suited for 2D and 3D cases. However, outliers, noise and missing points complicate the registration task. Given that the point clouds presented in this thesis are acquired by centerline extraction (section 2.7), from automatic segmentation, missegmentations are possible. Over-segmentation results in outliers and noise, whereas under-segmentation results in missing points. In order to realize an accurate registration, the point set registration method should be robust to these degradations.

1.3 Problem definition

As elaborated on in section 1.1.3, surgical resections, when compared to other treatment plans, provide the best patient outcome for various types of liver malignancies [38]. Due to high complexity and inter-patient variability of underlying hepatic vascular anatomy, planning and execution of safe resection is challenging in surgery. Therefore, repetitive intraoperative imaging is required to monitor surgery progress and assess the tumor-vessel relationship in 3D. Currently, US is the only imaging modality that is widely accepted and integrated into a surgical workflow, because it is an easy to use, real-time, non-ionizing and relatively cheap modality, compared to e.g. CBCT or MRI. Additionally, even though intraoperative ultrasonography is sensitive to image artifacts, i.e., reverberation, ghosting, signal loss due to the presence of the air, it results in high soft tissue contrast and spatial resolution. Therefore, ultrasonography is the most suitable imaging modality for intraoperative visualization of hepatic anatomy.

Despite many advantages of intraoperative US, it is still a primary 2D imaging modality, which complicates precise localization of each 2D image in 3D for a surgeon. Even when 3D reconstructions of 2D ultrasound images are performed, assessment of these 3D volume requires scrolling in three anatomical slice orientations, which is cumbersome in a surgical environment. An interactive visualization of automatically segmented vasculature in 3D would be of great value, yet challenging due to the complexity of US segmentation [82, 115]. CT or MR images ease identification of basic hepatic anatomy, containing all required information in tumors, major vessels and biliary tracts. Given this information, surgeons can find it difficult to estimate relations during surgical planning. Easing interpretation of conventional images seems fundamental to improve surgical outcome.

However, the deformability of the liver impedes with proper correlation to a 3D model, resulting in extended surgery time as the lesions have to be discovered manually. Improper localization potentially results in incorrect ablations or insufficient resection margins. Improvement of surgical planning has largely been a consequence of modern image processing and computer-based operation planning systems [116]. A patient specific visual illustration of an organ allows for operation planning by the surgeon. The spatial relation between the liver surface and lesions, the vasculature and other relevant structures can be shown by different visualization methods [31, 117]. Intraoperative US with registered image quality of CT and MR, can provide superior information, allowing the surgeon to spare large vessels based on real time feedback of a 3D model of the vascular topography. Additionally, a negative resection margin can be achieved with greater accuracy, thus increasing patient benefit. Surgical navigation is already implemented in several applications, such as facial surgery, neuro-surgery and orthopedics [118–120]. In literature, surgical navigation is mostly realized by preoperative CT or MR imaging combined with intraoperative tracked US or CBCT [61, 62, 73, 121]. Image guided surgery which is established this way provides real-time information of the surgical site. CBCT however increases the radiation load, is bulky and expensive, whereas US can be used freely at any point in time. In this work, an

attempt to alleviate these challenges using 3D ultrasound imaging, in conjunction with automatic vasculature segmentation is proposed.

1.3.1 Rationale of this thesis

These potential benefits of US-based registration lead to our research question: is it possible to automatically register preoperative imaging with intraoperative US imaging based on vascular centerline registration in patients with hepatic lesions? In order to answer this question, several sub goals have been defined where automatic segmentation of the ultrasound vasculature comprises the first step, followed by post processing to make the segmentations more similar in both modalities, expectantly improving registration. The centerlines of the automatic segmentations are used to register the US imaging to the preoperative imaging, providing optimal information to the surgeon during surgery. The subgoals are indicated as follows, with a schematic overview presented in Figure 2.1:

1. Development of a neural network to acquire a model that is capable of automatic vascular segmentation in 3D US images of the liver. Here we will also investigate the effects of training on datasets from different sources as well as a combination of sources.
2. Quantification of the performance of these models, and ultimately using the best model for construction of a patient-specific 3D vasculature model.
3. Automatic registration of preoperative volume to US volume based on automatically extracted centerlines of both modalities.
4. Evaluation of registration accuracy in a (post)clinical setting.

1.3.2 Thesis outline

This thesis presents a segmentation and a registration challenge. The outcome of the segmentation performance serves the registration performance, hence both challenges are presented together. Chapter 2 presents the materials and methods that are used in this thesis. Chapter 3 presents the results, followed by a discussion and general conclusion in chapter 4. The thesis is concluded with recommendations in chapter 5.

Chapter 2

Materials and methods

This study proposes to automatically register intraoperative US imaging with preoperative (MR or CT) imaging, thus providing additional information for localizing lesions and their location with respect to the major hepatic vasculature. In previous studies at the NKI-AvL, 35 3D US volumes have been acquired. For a study running in parallel we expanded that dataset by 34 additional stacked 2D US scans that are acquired in 7 patients during hepatic surgery. Registration is based on the vasculature that is present in both imaging modalities. The vasculature of the preoperative imaging is segmented semi-automatically and adjusted manually, whereas the US vasculature is segmented automatically by using a CNN. After an initial registration by recording the orientation and location of the ultrasound probe, the centerlines of both segmentations are used in the fine registration process. Figure 2.1 provides a visual overview of the developed registration framework. The Euclidean distance between the lesion in the US model and the registered preoperative model is used as a measure of accuracy.

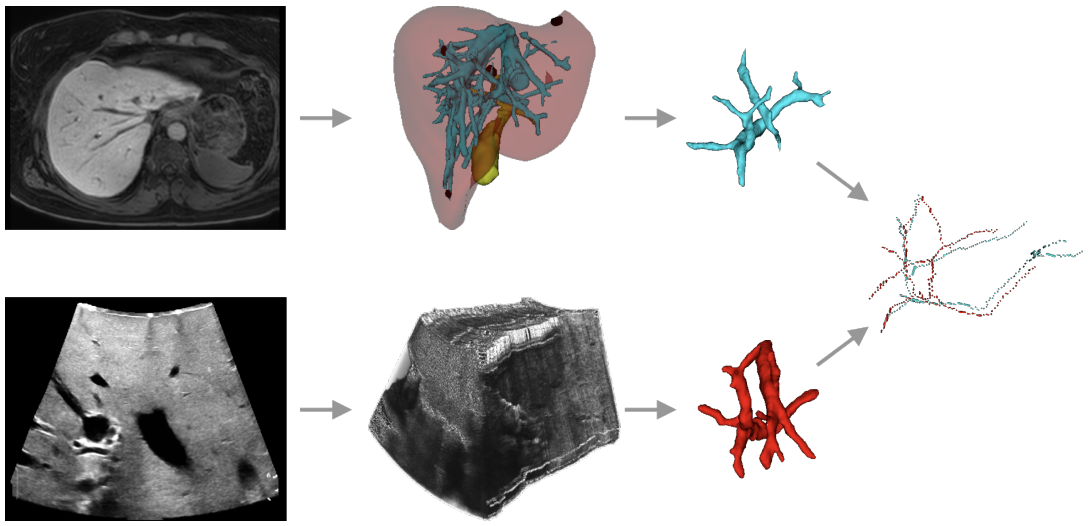


Figure 2.1: Vasculature is extracted from the preoperative scan prior to surgery. During surgery vasculature is extracted from a reconstructed US volume. Centerlines from both modalities are used for registration.

2.1 Patients

Inclusion to the additional patient group of whom stacked 2D US scans are acquired were bound by certain criteria. Patients scheduled for open surgery, for primary or secondary liver lesions from any origin, for whom MWA/RFA is required during surgery are included in the study population when the in- and exclusion criteria are satisfied.

2.1.1 Inclusion criteria

In order to be eligible for the study, a patient had to meet the following criteria:

- Age >= 18 years
- Patient provides written informed consent
- Lesion located within 5cm of the liver surface
- Lesion diameter under 8cm
- Presence of at least one centrally located liver lesion
- Patient is scheduled for ablation, open liver resection or both

2.1.2 Exclusion criteria

Patients who met the following criteria were not included in the population:

- Pregnancy
- Pacemaker
- Presence of large cysts near the target lesion
- Preoperative scan older than 2 months at the time of surgery
- Lesions with a complete radiological response or isoechoic liver lesions
- Metal implants in the thoracic or abdominal area, or other influences, that could interfere with the EM tracking

Table 2.1: Patient characteristics.

Characteristic	3D	Stacked 2D
Sex – no. (%)		
Male	8 (50)	10 (59)
Female	8 (50)	7 (41)
Age – yr		
Median (interquartile range)	61 (57-68)	66 (53-71)
Range	43-80	45-82
Number of lesions per patient	3.5	3.3
Usable US acquisitions per patient	2.2	3.3

2.2 Data

In total 115 US scans were collected in the study running parallel, of which 50 volumes were considered of sufficient quality. The main reason of exclusion was incorrect recombination of the stacked 2D volume due to either an EM-field error or due to too fast movement of the US-probe during acquisition. From these 50 volumes, 34 have been delineated. CustusX [75] was used for acquisition of the stacked 2D volumes, where the US operator was instructed to acquire a volume as large as possible, in one straight path. These instructions were sometimes misconceived as acquiring imaging for a clinical purpose could follow a different trajectory. The stacked 2D volumes were constructed, based on the 2D US slices, using the pixel nearest neighbor algorithm.

The readily available 3D dataset contained 35 scans and was expanded with 34 stacked 2D volumes (stacked based on EM tracking, Aurora Northern Digital — Ontario, Canada). Data distribution is presented in Table 2.2 and visualizations of the differences in volumes are presented in Figure 2.5. The test set of the stacked 2D volumes is twice as big with regard to the 3D dataset because these are used in the clinical setting, hence the segmentation performance on these scans

is of higher importance. Original 3D volume sizes were $512 \times 400 \times 256$ pixels and stacked 2D volumes ranged from $293 \times 396 \times 526$ to $404 \times 572 \times 678$ pixels, depending on the zoom of the 2D slices, but were downsampled to 40% prior to training. US acquisitions were performed by five different operators at the NKI-AvL.

Each acquisition was delineated in 3D Slicer [122] by one out of four annotators and has been validated with an expert radiologist. The hepatic and portal veins were segmented separately, as well as the liver parenchyma, however, in this study we combined the hepatic and portal vein labels adn did not utilize the liver parenchyma. To estimate the performance of automatic segmentation, a benchmark was established. This was done by computing the inter-observer variability of four scans, delineated by two users. For registration accuracy assessment additional scans have been used from the parallel study, of which the majority lack a ground truth delineation and thus are not reported for segmentation accuracy.

Table 2.2: Overview of number of US volumes per modality.

Dataset	Training	Validation	Test	Total
3D	26	6	3	35
Stacked 2D	22	6	6	34
Combined	48	12	9	69
Inter-observer			4	4

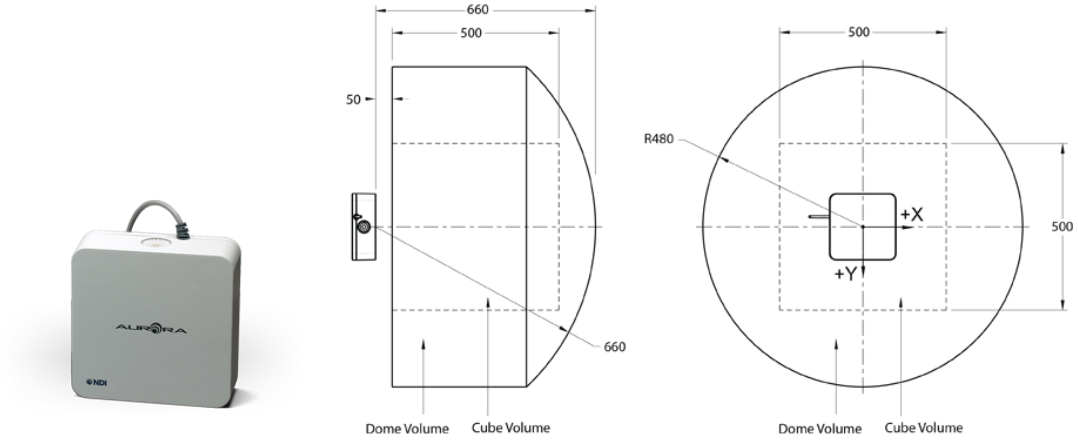
2.3 Components

This Section describes the components that are used in realization of tracking of the US probe in order to create stacked 2D volumes. The Aurora V2 from Northern Digital Inc. (Waterloo, Canada) is one of the popular commercially available tracking systems today and is also used in this study. Two types of field generators are used during data acquisition. The planar field generator in Figure 2.2a is mounted on a positioning arm, offering flexible setup options around the patient. The tabletop field generator in Figure 2.3a is positioned in a Plexiglas casing underneath the mattress of the operating bed. The generated EM field has a predetermined field of view where the sensors positioned on the tools can be tracked (Figures 2.2b and 2.3b). The field generator generates a well-defined EM field, the coils in the sensors, when placed within the field, deform it. Specific deformations are then related to a specific position and angle of the coil. In-house sensor holders have been developed, for stacking 2D US slices, and allow unique positioning onto the US probe (BK FlexFocus 5000, T- I145T US, Figure 2.4b), which have been calibrated using the methodology described in [123]. The electric signal coming from the sensors is amplified and digitized by a Sensor Interface Unit (SIU), increasing the distance that can be spanned by the signal and minimizing noise. The amplified signal is collected by the System Control Unit (SCU), which then calculates the position and orientation of each sensor and connects with the host computer. The 35 3D US volumes that were available in the NKI-AvL have previously been acquired with the Philips EPIQ7, X6-1 probe.

The aforementioned components are combined using dedicated software on a navigation trolley. Open-source software CustusX [75] is used for the navigation and visualization and is dedicated to ultrasound imaging and intraoperative navigation in a phantom or research setting [124]. This software is able to reconstruct a 3D volume based on the 2D US acquisition combined with the spatial information acquired with the EM system.

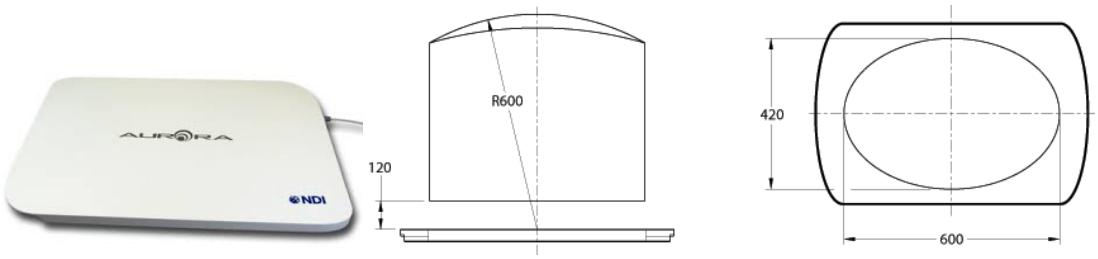
2.4 Initial registration

Prior to acquiring US volumes, the orientation of the tracked US probe is used for setting an approximate patient orientation, one landmark is then used for the translation part of the registration. This initial registration allow for partial overlapping of the images, thus preventing possible local minima during the fine registration.



(a) (b)

Figure 2.2: Aurora NDI planar system and generated EM field, dimensions are in mm.

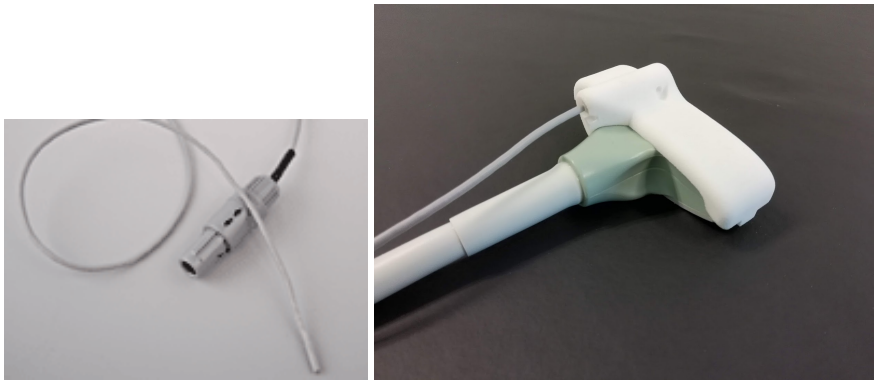


(a) (b)

Figure 2.3: Aurora NDI tabletop system and generated EM field, dimensions are in mm.

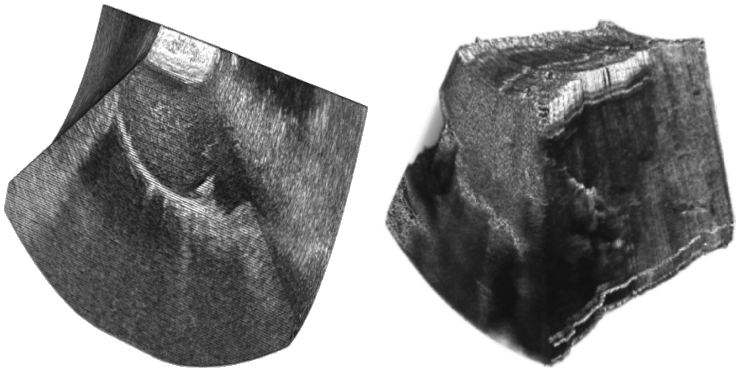
2.5 Pre-processing

As US imaging inherently contains speckle noise, a 3x3 median filter is applied prior to training of the network. Median filtering is commonly used in segmentation tasks in order to reduce noise and improve segmentation performance [125]. Moreover, since US imaging that is acquired for training of the network, originates from different sources, the pixel spacing of both modalities (true 3D and stacked 2D) is normalized to 1. In order to allow for a bigger batch size, all US images were down-sampled to 40 % of their original size prior to training, reducing burden on the GPU memory.



(a) (b)

Figure 2.4: Aurora 6DOF sensor (a) and calibrated US probe grip with 6DOF sensor (b).



(a) 3D volume. (b) Stacked 2D volume.

Figure 2.5: Two types of US volumes acquired.

2.6 Segmentation

Once the data has been pre-processed, it can be used to train a CNN. In this study, a reduced filter 3D U-Net, chosen due to its popularity in medical image segmentation [126], is proposed to achieve accurate vessel segmentation in both true 3D (Figure 2.5a) and stacked 2D (Figure 2.5b) US volumes. Segmentation performance is reported based on training and testing of the 3D U-Net on solely 3D US volumes, solely stacked 2D US volumes and a combination of both datasets (3D + stacked 2D). Significance regarding performance measures is tested with relation to the combined dataset based on a t-test. In the ultimate clinical workflow, stacked 2D volumes will be used for registration. Based on the segmentation performance, the best performing model will be used for segmentation of the vasculature that is used for registration.

2.6.1 3D U-Net

Figure 2.6 illustrates the 3D U-Net [127] architecture used in this study, like the standard U-Net [128], it has an analysis and synthesis path each with four resolution steps. Each layer in the analysis path contains two $3 \times 3 \times 3$ convolutions, which are followed by a rectified linear unit (ReLU) activation function and a $2 \times 2 \times 2$ max pooling with strides of two in each dimension. Each layer in the synthesis path consists of an upconvolution of $2 \times 2 \times 2$ with strides of 2 in each dimension, followed by two $3 \times 3 \times 3$ convolutions which are followed by a ReLU activation function. Skip connections from layers of equal resolution, transfer the essential high-resolution features to the synthesis path. The final layer reduces the number of output channels to 2 by means of a $1 \times 1 \times 1$ convolution. Moreover, batch normalization is performed before each ReLU activation function. The 3D U-Net architecture that is used in this study is a NiftyNet [129] Tensorflow implementation similar to Cicek et al. [127], however with an eighth of the amount of filters in every layer compared to the original implementation (Figure 2.6), to avoid memory related bottlenecks. Training using the Dice loss was performed on four NVIDIA (Nvidia cooperation, Santa Clara, California) 1080 GTX GPUs. Every epoch the network parameters are stored in a checkpoint, the 5 checkpoints with the lowest loss on the validation set are used to asses performance on the test set, of which the best performing checkpoint is used to report the ultimate results. Performance measures used for determining segmentation accuracy are elaborated on in section 2.6.3.

2.6.2 Hyper-parameter optimization

Depending on the model and on how many hyper-parameters the experimenter chooses to optimize, neural networks have from ten to fifty hyper-parameters [130]. A combination of grid search and manual search is the most widely used strategy [131] in optimizing hyper-parameters. Grid search requires choosing a set of values for each variable (e.g., learning rate, type of optimizer, amount of patches taken from volume), resulting in an exponentially growing number of combinations with the number of hyper-parameters. Manual search identifies regions that are promising while at the same time developing the intuition that is necessary for further optimization. Where manual search suffers from difficulty in reproducibility, grid search suffers

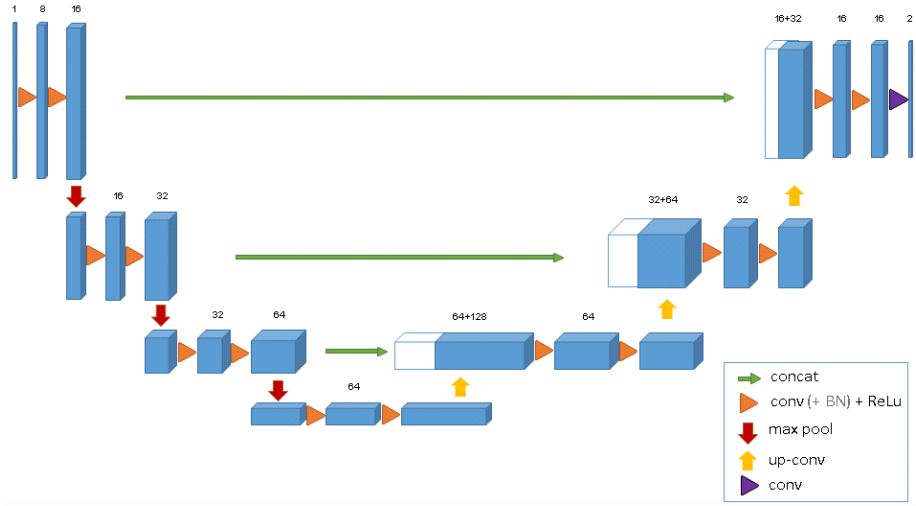


Figure 2.6: 3D U-Net architecture [127] used in segmentation of liver vasculature.

from mostly ineffective computing time, whilst performing poorly. Random search is proposed in literature because of its practicality and robustness [131]. Despite decades of research into hyper-parameter optimization algorithms [132, 133], manual search has no technical overhead and gives a degree of insight into the model's behavior, whereas grid search is easy to implement and reliable [131]. Hence, in this study a combination between grid and random search is used for tuning of the hyper-parameters. Hyper-parameters that are optimized are the amount of patches per volume and their size, learning rate, regularization type, batch size, padding, type of optimizer and the amount of feature maps. The hyper-parameters are optimized based on their Dice loss performance ($1-$ mean Dice score over fore- and background classes), ultimate performance is reported in Dice and Jaccard index.

2.6.3 Performance measures

After hyper-parameter optimization, ultimate performance is reported in metrics similar to those used in literature [134], based on wide use in the evaluation of segmentation algorithms. Namely:

- the Dice similarity coefficient (DSC, Equation 2.1) [135, 136]
- Jaccard index (JI, Intersection over Union, Equation 2.2)

Volume metrics are specifically chosen, to give an overview of segmentation accuracy [134]. Boundary metrics are highly sensitive to outliers [137] and it is expected that parts of the smaller vasculature are more challenging to segment due to the downsampling of the data. DSC and JI are volume based metrics that are optimal when equal to 1, indicating full overlap of the volumes [138]. Although the two measures appear similar, in JI poor classifications are weighted more strongly and in literature both metrics are used separately.
DSC is defined as:

$$DSC = \frac{2(A \cap B)}{|A| + |B|} \tag{2.1}$$

where A is the number of segmented voxels in the ground truth and B is the number of voxels in the segmentation result.

JI gives the similarity between the ground-truth and predicted region and is defined as the size of the overlap divided by the union of the two regions:

$$JI = \frac{TP}{FP + TP + FN} \tag{2.2}$$

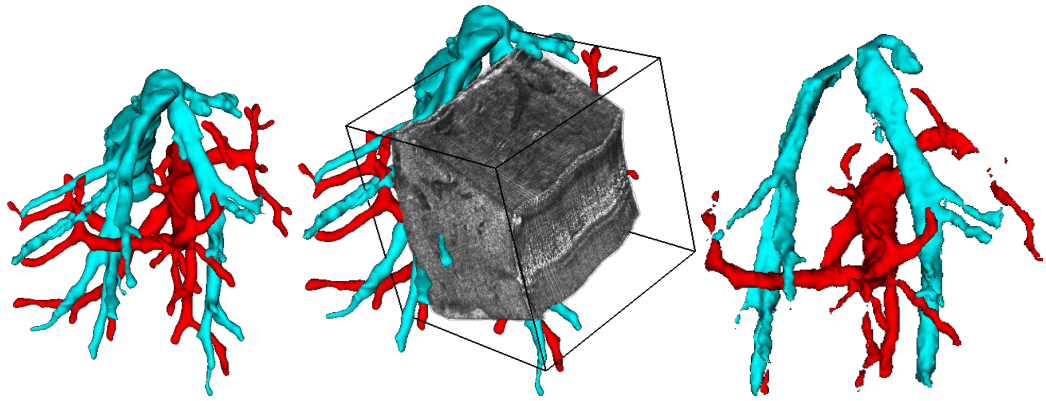
Where TP, FP and FN respectively signify the true positive, false positive and false negative.

2.6.4 Post-processing

Prior to surgery, the vasculature is segmented from preoperative imaging by using a vesselness filter and manual post processing. This step is performed so that the surgeon is only exposed to the relevant vasculature, hence smaller branches do not occlude the overview. This step is not time-critical and can be iterated over multiple times. Predicted (US) segmentations often contain small islands and other noise that do not belong to a vascular model and can be discarded by post-processing the output of the network. A median filter with kernel size 3 is used in both modalities to remove excess noise and make the point clouds that represent vasculature more similar with respect to the vessel sizes. From the median filtered output the two largest structures are identified, followed by extraction of their centerlines. An implementation of [139] is used to automatically extract the centerlines from both the preoperative and the intraoperative segmentation, similar to that of [140]. The centerlines are represented as a set of discrete points. Centerlines are used because they represent a smooth path of the vessels, whilst at the same time reducing the size of the point cloud, hence reducing computational load and improving speed of the registration process.

2.7 Fine registration

After the centerlines have been extracted, successful navigation depends on accurate registration between the pre- and intraoperative centerlines. Given the initial registration, a region of interest based on the volume spanned by the stacked 2D US reconstruction, can be cropped out of the preoperative point cloud. A visual impression of this is given in Figure 2.7. This step is essential because it results in two similar point clouds that can then be automatically registered by aligning the centerlines. The centerlines of both modalities are registered based on the CPD algorithm. In all registrations, manual corrections are made to shed a light on possible improvement of the registration pipeline. These manual corrections comprise adjusting of the cropped volume from the preoperative model, and removing of over-segmentations in the intraoperative model.



(a) (b) (c)

Figure 2.7: Automatically determined cropbox (black lines) around the US volume, based on US acquisition. (a) Preoperative model with hepatic vein in blue and portal vein in red. (b) US volume overlaid on preoperative model after initial registration, crop volume indicated by black box. (c) Cropped preoperative vasculature used for fine registration based on CPD.

2.7.1 Coherent point drift

Iterative closest point (ICP) algorithm [141] is the most used registration algorithm because of its efficiency and simplicity. However, it can be prone to converge to a local minima, especially if there is a lot of noise [142]. Although post-processing is applied to make both point clouds as similar as possible, it is expected that some differences in the amount of vascular branches can be observed.

Myronenko et al. [102] introduce the CPD algorithm, which is a robust probabilistic multi-dimensional point set registration algorithm suitable for rigid transformations. They consider

the alignment of two point clouds as a probability density estimation problem, where the target point cloud represents data points and the source point cloud is represented by the Gaussian Mixture Model (GMM). The GMM centroids (US vasculature model) are fitted to the target point cloud (preoperative vasculature segmentation) by expectation-maximization (E-M) optimization. The E-step comprises finding out from which Gaussian the observed point cloud was sampled from, providing a correspondence probability. In the M-step, maximization of the negative log-likelihood that the observed points were sampled from the GMM is performed. This provides transformation parameters as soon as the correspondence probabilities are known, resulting in a rotation matrix and a translation vector, which can be combined into a transformation matrix.

The point sets are aligned at the optimum, where the correspondence is the maximum of the GMM posterior probability for any given data point. The core of this method is that the GMM centroids move coherently as a group, preserving the topological structure of the point sets. The coherence constraint is imposed by explicit reparameterization of the GMM centroid locations. Two distinct advantages are provided by this approach. First, it allows to assign points to each other, so that the registration can be solved as a Procrustes problem. Moreover, depending on the distance to the GMM centroid, proximity weighting can be applied to the loss.

In the original paper the authors have also shown how the computational complexity of the method can be reduced to linear, making it applicable for large data sets. In this study, rigid CPD registration is performed, the algorithm pseudo code is presented in Figure 2.8, where the following notations are used:

- N, M - number of points in the points sets,
- D - dimension of the point sets,
- $\mathbf{X}_{NxD} = (\mathbf{x}_1, ..., \mathbf{x}_N)^T$ - the preoperative segmentation (the data points),
- $\mathbf{Y}_{NxD} = (\mathbf{y}_1, ..., \mathbf{y}_N)^T$ - the automatic US segmentation (the GMM centroids),
- $\mathcal{T}(\mathbf{Y}, \theta)$ - Transformation $\mathcal{T}$ applied to $\mathbf{Y}$, where θ is a set of the transformation parameters,
- $\mathbf{1}$ - column vector of ones,
- $\mathbf{I}$ - identity matrix,
- $d(a)$ - diagonal matrix formed from the vector $\mathbf{a}$

2.7.2 Performance measures

Performance of the fine registration using the CPD algorithm is assessed postoperatively. A differentiation is made between successful and unsuccessful registrations based on visual inspection of the overlapping centerlines after the fine registration. In both successful and unsuccessful registrations, the US lesion is delineated manually and is then used for determining the registration accuracy. The registration accuracy is defined as the target registration error (TRE) between the centroid of the US lesion and the centroid of the registered preoperative lesion, expressed in Euclidean distance (mm, Figure 2.9).

Rigid point set registration algorithm:

- Initialization: $\mathbf{R} = \mathbf{I}, \mathbf{t} = 0, \mathbf{s} = 1 \leq w \leq 1$
  $\sigma^2 = \frac{1}{DNM} \sum_{n=1}^{N} \sum_{m=1}^{M} ||\mathbf{x}_n - \mathbf{y}_m||^2$
- EM optimization, repeat until convergence:
  - E-step: Compute $\mathbf{P}$,
    $p_{mn} = \frac{exp^{-\frac{1}{2\sigma^2}||\mathbf{x}_n - (\mathbf{R}\mathbf{y}_m + \mathbf{t})||^2}}{\sum_{k=1}^{M} exp^{-\frac{1}{2\sigma^2}||\mathbf{x}_n - (s\mathbf{R}\mathbf{y}_k + \mathbf{t})||^2} + (2\pi\sigma^2)^{D/2} \frac{w}{1-w} \frac{M}{N}}$
  - M-Step: Solve for $\mathbf{R}$, s, $\mathbf{t}$, σ^2:
    - $N_{\mathbf{P}} = \mathbf{1}^T\mathbf{P}\mathbf{1}, \mu_x = \frac{1}{N_p}\mathbf{X}^T\mathbf{P}^T\mathbf{1}, \mu_{\mathbf{y}} = \frac{1}{N_p}\mathbf{Y}^T\mathbf{P}\mathbf{1},$
    - $\hat{\mathbf{X}} = \mathbf{X} - \mathbf{1}\mu_{\mathbf{x}}^T, \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{1}\mu_{\mathbf{y}}^T,$
    - $\mathbf{A} = \hat{\mathbf{X}}^T\mathbf{P}^T\hat{\mathbf{Y}}$, compute SVD of $\mathbf{A} = \mathbf{U}SS\mathbf{V}^T,$
    - $\mathbf{R} = \mathbf{U}\mathbf{C}\mathbf{V}^T$, where $\mathbf{C} = d(1, .., 1, det(\mathbf{U}\mathbf{V}^T)),$
    - $s = \frac{tr(\mathbf{A}^T\mathbf{R})}{tr(^T d(\mathbf{P}\mathbf{1})\hat{Y})},$
    - $\mathbf{t} = \mu_x - s\mathbf{R}\mu_y,$
    - $\sigma^2 = \frac{1}{N_{\mathbf{P}}D}(tr(\hat{\mathbf{X}}^T d(\mathbf{P}^T\mathbf{1})\hat{\mathbf{X}}) - str(\mathbf{A}^T\mathbf{R})).$
- The aligned point set is $\mathcal{T}(\mathbf{Y}) = s\mathbf{Y}\mathbf{R}^T + 1\mathbf{t}^T,$
- The probability of correspondence is given by $\mathbf{P}$.

Figure 2.8: Summary of rigid point set registration algorithm, adapted from [102]

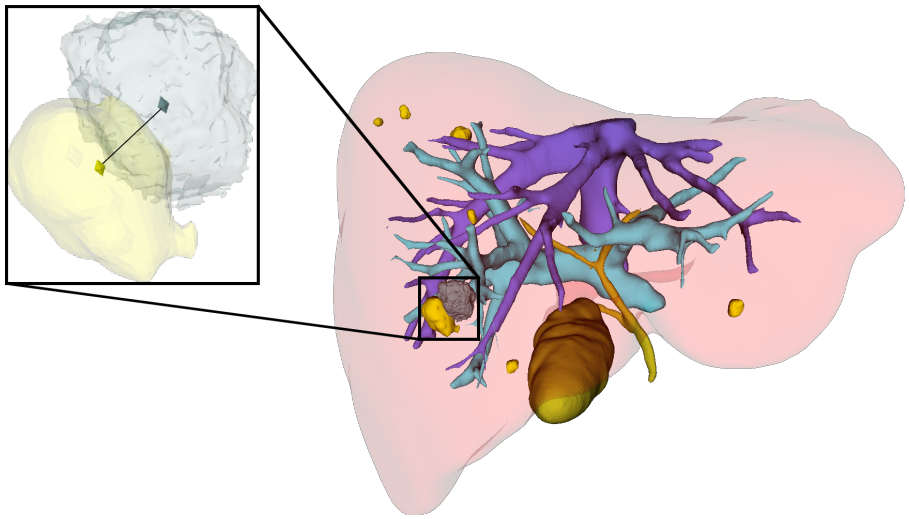


Figure 2.9: Registration evaluation is computed as the Euclidean distance between the registered preoperative lesion (yellow), and the reconstructed US lesion (blue). Distance is expressed in mm, Figure inspired by [67].

Chapter 3

Results

This chapter presents the study results, starting with the optimal hyper-parameter settings for training of the network. Then, the segmentation performance over the different datasets is followed by presentation of the registration results based on the automatic segmentations. The registration results are also compared between the different segmentation models and with inter-observer variability.

3.1 Hyper-parameter optimization

As described in Section 2.6.2 we performed a grid search over the hyper-parameters in order to optimize the training process. The hyper-parameters that were iterated over are presented in Table 3.1. Optimization was performed over the combined dataset, the best performing values were used during training of the network on all datasets. Adam optimizer, with learning rate 5×10^{-3}, L1 regularization with 10^{-5} weight decay, and a batch size of 2 was used. Twenty patches with size $152 \times 152 \times 96$ were used per mean value normalized volume. All volumes were padded with a volume of $32 \times 32 \times 32$ and were augmented by rotating (between -10° and 10°), scaling (between -10% and $+10\%$) and elastically deforming (S.D. 1).

Table 3.1: Table of hyper-parameters which was iterated over.

Hyper-parameter	Range	Best performing value
Samples per volume	[4, 20, 32, 50, 500]	20
Learning rate	[0.005, 0.01, 0.02]	0.005
Patch size	[[96, 96, 96], [104, 104, 96], [152, 152, 152]]	[152, 152, 152]
Regularization	[None, L1]	L1
Feature maps	[[8, 16, 32, 64, 128], [32, 64, 128, 256, 512]]	[8, 16, 32, 64, 128]
Batch size	[1, 2]	2
Zero padding	[None, [32, 32, 32]]	[32, 32, 32]
Optimizer	[None, Adam]	Adam

3.2 Training

The network, using the best performing parameters presented in Table 3.1, was trained on the 3D dataset for 320 epochs ($\pm$ 35 hours), where the training and validation loss gradually respectively converged to 0.044 and 0.126. When trained on the stacked 2D dataset for 223 epochs ($\pm$ 50 hours) the training and validation loss respectively converged to 0.038 and 0.068. A combination of the datasets led to loss values of 0.052 and 0.081 for the training and validation respectively in 231 epochs ($\pm$ 49 hours). Figure 3.1 shows the loss function values from training and validation on the three different datasets. During training on the different datasets a similar trend is seen, where fast converging is observed in both the training and validation loss up to the $\pm$ 40th epoch. In all cases training was ceased $\pm$ 50 epochs after apparent converging of the validation loss stalled to exclude further possible converging of the validation loss. The five checkpoints with the lowest

validation loss were used to evaluate segmentation performance on the test set, all analyses are based on the best performing checkpoint of which the loss is presented in Table 3.2.

Table 3.2: Training and validation loss at most optimal checkpoint when trained on solely stacked 2D, solely 3D or combined dataset.

Dataset	Training loss	Validation loss	Epoch used for segmentation	Time trained (hours)
3D	0.044	0.126	288	31
Stacked 2D	0.038	0.068	222	50
Combined	0.052	0.081	213	45

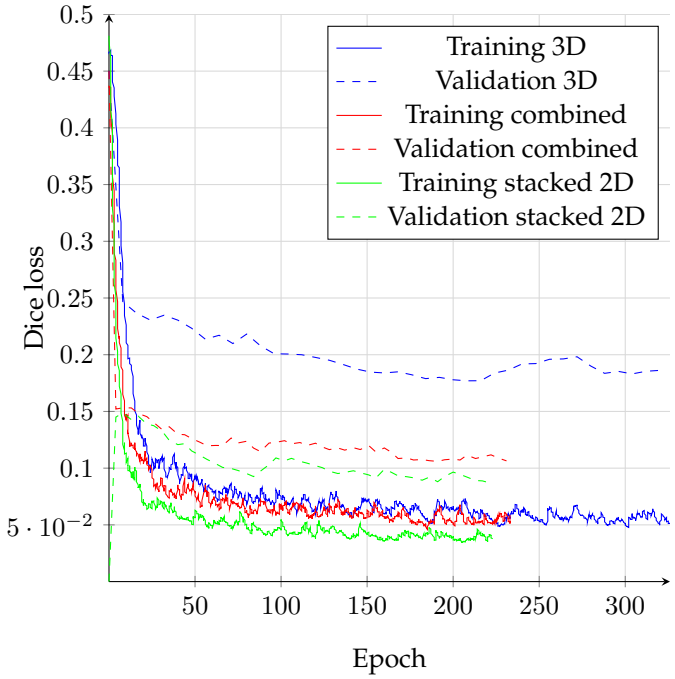


Figure 3.1: Validation losses for different settings, note that these are running averages and therefore do not exactly match the values in Table 3.2.

3.3 Segmentation performance on different datasets

This section presents the performance metrics of the best performing checkpoint on the test set. When trained on the 3D dataset, the reduced filter 3D U-Net obtained average Dice scores of 0.743 ($\pm$0.09), 0.655 ($\pm$0.12) and 0.684 ($\pm$0.12) for the 3D, stacked 2D and combined dataset respectively. Mean JI is reported at 0.623 ($\pm$0.08), 0.435 ($\pm$0.13) and 0.498 ($\pm$0.15) respectively.

When trained on the stacked 2D dataset the mean Dice is reported at 0.593 ($\pm$0.08), 0.747 ($\pm$0.13) and 0.696 ($\pm$0.13) for the 3D, stacked 2D and combined dataset respectively. The mean JI performance is respectively reported at 0.423 ($\pm$0.09), 0.610 ($\pm$0.14) and 0.548 ($\pm$0.15).

In case of training on the combined dataset the mean Dice is reported at 0.753 ($\pm$0.07), 0.783 ($\pm$0.10) and 0.773 ($\pm$0.10) for the 3D, stacked 2D and combined dataset respectively. In this training setup the mean JI is respectively reported at 0.607 ($\pm$0.10), 0.657 ($\pm$0.13) and 0.640 ($\pm$0.12). A visualization of the segmentation performance of training on the combined model is presented in Figure 3.2.

The inter-observer mean Dice is reported at 0.879 ($\pm$0.02) with an JI of 0.785 ($\pm$0.02) based on a comparison between 4 (stacked 2D) US volumes. A complete overview of the aforementioned numbers is presented in Table 3.3.

When comparing the model trained on the combined dataset with the models trained on the separate datasets, the combined model outperforms either of the separate models. This change

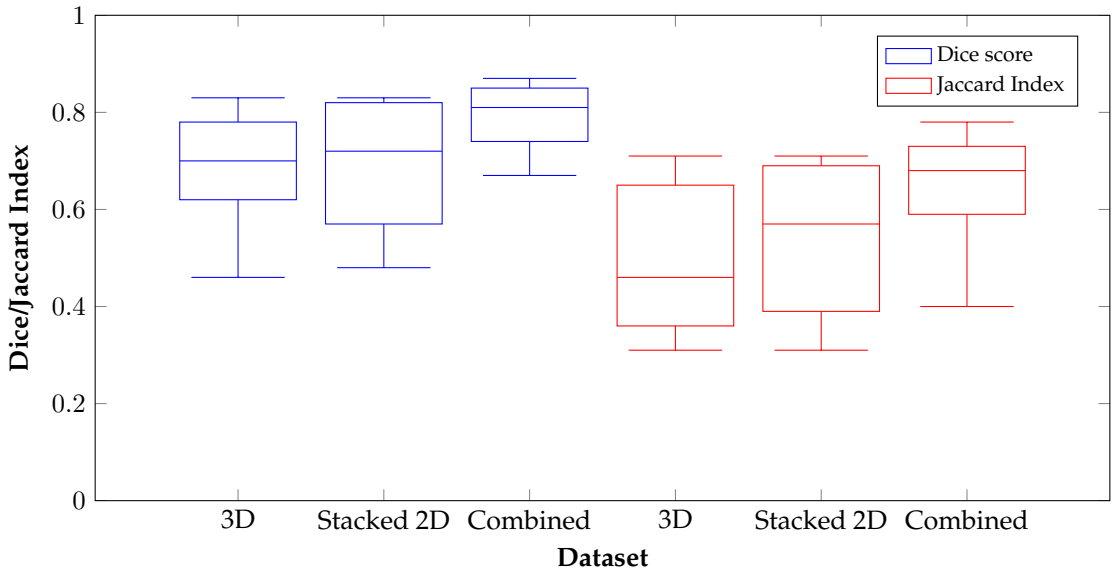


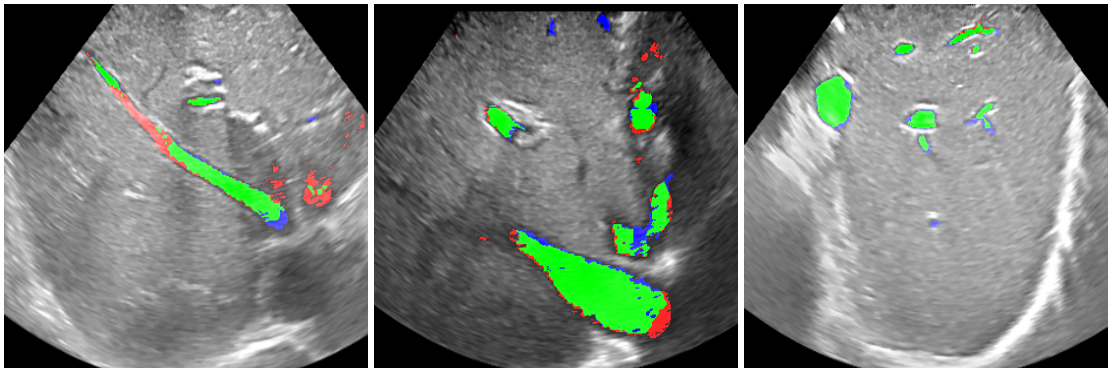
Figure 3.2: Segmentation performance, of the model trained on the combined dataset, on the seperate and combined datasets.

in performance is only significant when comparing the JI between the stacked 2D and combined dataset performance between the models trained on the 3D dataset and the combined dataset. The JI performance is significantly lower on the stacked 2D volumes and combined dataset, when compared to the segmentation performance based on the combined dataset model.

Figure 3.3 (3D) and Figure 3.4 (stacked 2D) present visual elaboration on performance for six subjects based on training on the combined dataset, ranging from poor to excellent segmentation performance according to the performance metrics.

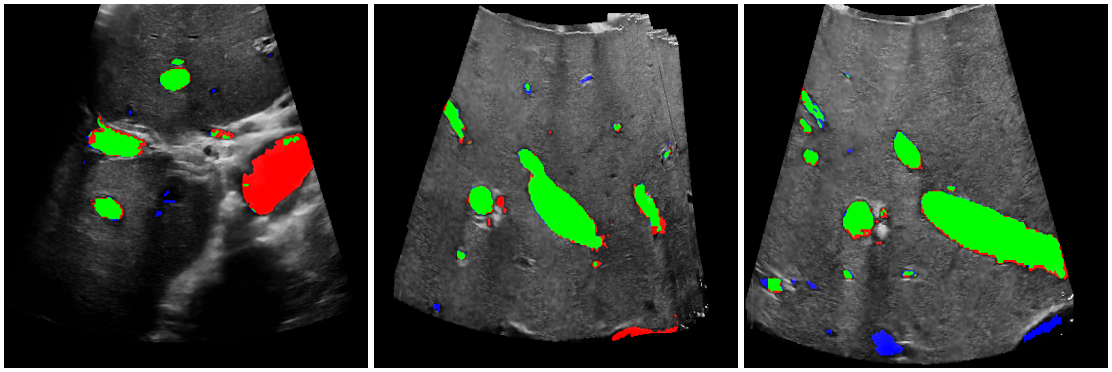
Table 3.3: Performance metrics for vessel segmentation in 3D, stacked 2D, the combined dataset and inter-observer. Note that all UVI US volumes are acquired with the 3D probe and the ULN volumes are acquired with the stacked 2D probe. P-values comparing the 3D and stacked 2D with the combined dataset are reported in parentheses with the mean values. Significance compared to training on the combined dataset is indicated in **bold**.

Trained on dataset:	3D		Stacked 2D		Combined	
Patient	**Dice**	**JI**	**Dice**	**JI**	**Dice**	**JI**
UVI_004	0.83	0.71	0.70	0.54	0.85	0.73
UVI_012	0.78	0.64	0.57	0.39	0.74	0.59
UVI_040	0.62	0.45	0.51	0.34	0.67	0.50
ULN_002003	0.46	0.30	0.48	0.31	0.57	0.40
ULN_003009	0.83	0.71	0.83	0.71	0.87	0.78
ULN_004004	0.72	0.56	0.83	0.71	0.86	0.76
ULN_004005	0.59	0.42	0.82	0.69	0.84	0.72
ULN_005001	0.70	0.54	0.80	0.67	0.81	0.68
ULNt_08005	0.63	0.46	0.72	0.57	0.75	0.60
Mean 3D	0.743 (0.91)	0.623 (0.86)	0.593 (0.10)	0.423 (0.11)	0.753	0.607
Mean stacked 2D	0.655 (0.09)	**0.435** (0.02)	0.747 (0.62)	0.610 (0.60)	0.783	0.657
Mean combined dataset	0.684 (0.11)	**0.498** (0.04)	0.696 (0.20)	0.548 (0.20)	0.773	0.640
Mean inter-observer	0.879	0.785	0.879	0.785	0.879	0.785



(a) Dice = 0.75 (b) Dice = 0.80 (c) Dice = 0.84

Figure 3.3: Examples of 3D test set segmentation results, true positives are colored green, false positives red and false negatives blue, Dice is measured over total volume. The indicated Dice score is reported based on the complete volume.

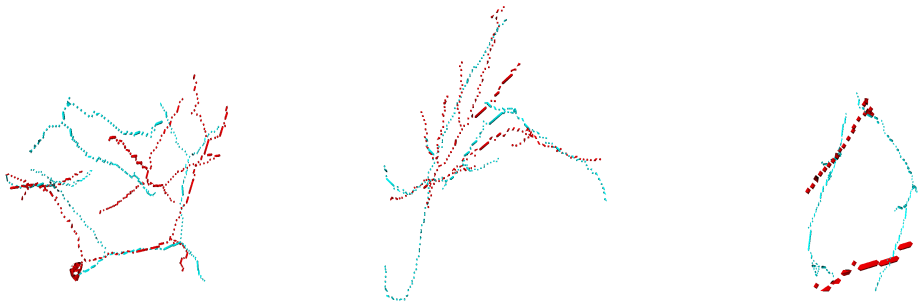


(a) Dice = 0.63 (b) Dice = 0.73 (c) Dice = 0.80

Figure 3.4: Examples of stacked 2D test set segmentation results, true positives are colored green, false positives red and false negatives blue. The indicated Dice score is the score over the complete volume.

3.4 Registration

Due to the higher segmentation performance when trained on the combined dataset, this model is used for segmentation of the vasculature that is used for registration with the preoperative model. The TREs present a spread in accuracy based on whether the registration was successful on visual inspection. The average successful TRE for automatic fine registration was 12.29 $\pm$ 4.93 mm. This is a slight improvement when compared to the initial registration (15.77 $\pm$ 5.92 mm), based on orientation of the probe and a single point translation. Out of 11 target volumes, three showed a TRE < 10 mm, which is considered a safety margin in [143]. Figure 3.5 shows examples of alignment of the centerlines after fine registration. Figure 3.5a and 3.5b present successful registrations. Unsuccessfully automatically registered volumes such as Figure 3.5c, have a mean TRE of 47.32 $\pm$ 25.7 indicating a large spread and misalignment. It is noted that volumes with more US information relative to the size of the volume that is used for cropping, perform better. In this small test set it appears that when there is twice as much US information in the cropping US volume, performance is near the 10 mm threshold. Figure 3.6 visualizes this relation. It is shown that manual adjustments improve the registration accuracy, where the mean of the successful registrations is reported at 13.23 $\pm$ 3.93 compared to initial (22.51 $\pm$ 13.33) and fine (31.40 $\pm$ 25.99) registration. A complete overview of the TRE values acquired per volume are given in Table 3.4. The manually adjusted registrations that fail often present with correct alignment on a single blood vessel (i.e. middle hepatic vein), but are rotated. The manual adjustments were made in such a manner that a the preoperative volume is cropped by a volume more specific to the US segmentation, which improves the accuracy of the registration.



(a) TRE = 6.74 mm (b) TRE = 9.72 mm (c) TRE = 34.98 mm

Figure 3.5: Examples of registered centerlines of stacked 2D US, preoperative centerline is visualized in blue, US is visualized in red

Table 3.4: TRE after coarse and fine registration per patient, it is also reported whether the registration was successful on visual inspection, dimensions are in mm.

Patient	TRE initial (mm)	TRE fine (mm)	Manually adjusted (mm)	US to crop volume ratio	Successful initial/fine (manual)
ULN_007006	13.75	29.02	15.19	1.2	no (no)
ULN_006001	27.17	33.36	25.71	1.2	no (no)
ULN_006002	29.92	34.98	45.48	1.2	no (no)
ULN_002003	9.93	100.71	22.03	1.4	no (no)
ULN_006003	31.52	28.84	19.91	1.4	no (no)
ULN_003009	56.56	57	12.4	1.5	no (yes)
ULN_004001	13.17	19.51	15.30	1.6	yes (yes)
ULN_004004	13.17	9.18	12.18	2	yes (yes)
ULN_005004	11.76	16.76	20.11	2	yes (yes)
ULN_004005	13.17	6.75	12.4	2.2	yes (yes)
ULN_006004	27.57	9.26	7.01	2.2	yes (yes)
Mean unsuccessful	28.14 $\pm$15.07	47.32 $\pm$25.71	25.66 $\pm$ 10.48	1.35 $\pm$0.17	55 %
Mean successful	15.77 $\pm$5.92	12.29 $\pm$4.93	13.23 $\pm$ 3.93	2.13 $\pm$0.1	45 %
Mean	22.51 $\pm$13.33	31.40 $\pm$25.99	18.88 $\pm$9.83	1.69 $\pm$0.41	100 %

3.5 Workflow efficiency

The current registration approach is set up in such a manner that volume reconstruction and segmentation of the US vasculature take around 60 seconds, the extra time that is added by the registration of the centerlines is 10.8 ($\pm$ 1.5) seconds. During the average cumulative time of 69.74 ($\pm$ 14.6) seconds needed to perform the whole registration procedure, the only action required from the surgeon is making an US sweep after laparotomy. An overview of the time needed is given in Table 3.5.

Table 3.5: Overview of average time taken for automatic registration from US sweep to registration, time is indicated in seconds.

Step	Time taken (s)
Mean volume reconstruction	18.2 $\pm$ 8.99
Mean automatic segmentation	40.77 $\pm$ 11.4
Mean automatic registration	10.8 $\pm$ 1.5
Sum	69.74 $\pm$ 14.6

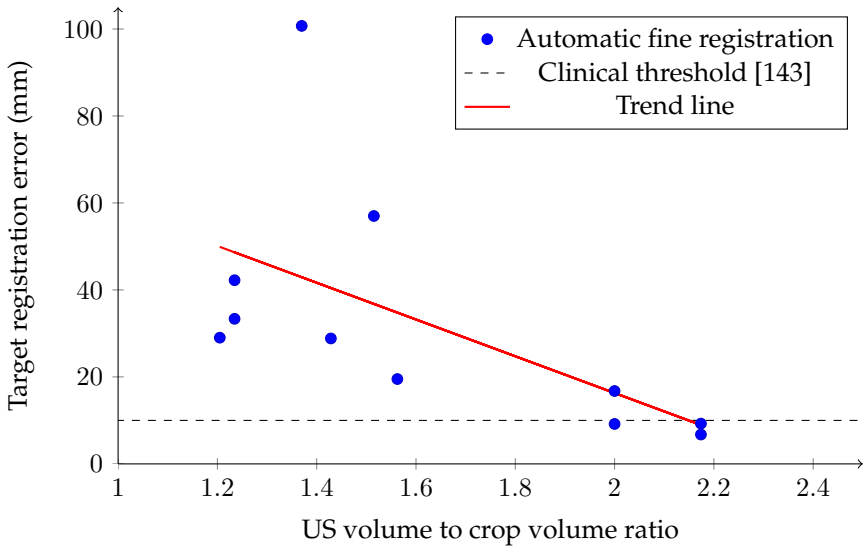


Figure 3.6: Influence of US volume to crop volume ratio on TRE measured in the lesion, after automatic fine registration.

Chapter 4

Discussion and conclusion

4.1 Discussion

This thesis addressed the task to realize automatic vessel-based registration of intraoperative US imaging with preoperatively acquired (CT or MR) imaging during hepatic surgery. For that purpose, a 3D U-Net was trained to automatically segment intraoperative vasculature. Training on a combined dataset of stacked 2D and 3D imaging gave the most promising results, with a Dice of 0.773 ($\pm$0.10) and a JI of 0.640 ($\pm$0.12), comparing to an inter-observer variability of respectively 0.879 ($\pm$0.02) and 0.785 ($\pm$0.02). The intraoperative segmentation was then registered with a preoperative, semi-automatically segmented, vasculature model. This has been achieved by means of extracting the centerlines from the post-processed segmentations and determining the transformation matrix between the two using CPD algorithm. Ultimately, the registration performance was quantified by computing the Euclidean distance between the center of the intraoperative lesion and the center of the lesion in the registered preoperative imaging. The registration achieved decent results when the scanned US volume satisfied several requirements, with a TRE of 12.29 $\pm$ 4.93 in visually successful registrations. However in many cases (55%) the algorithm was prone to failure (overall TRE of 47.32 $\pm$ 25.71 mm), when the preoperative model was adjusted manually, an improvement was seen (overall TRE 25.66 ± 10.48 mm).

4.1.1 Segmentation

The CNN that has been used to acquire automatic segmentations is adapted from literature but has been reduced in the amount of filters to handle larger patches of 3D data. It proves to accurately segment hepatic vessels on US imaging with a relatively small dataset, but deviates from the inter-observer performance. We show that training on the combined dataset improves performance, when compared to training on either separate dataset. However, the only measure that shows significant underperformance is the JI on the stacked 2D and combined dataset, when segmentation is done with the model trained on the 3D dataset. These results indicate that similar features are present in the 3D and stacked 2D volumes, leading to increased performance. However, when using the combined dataset, training is performed on twice the amount of US volumes, expectantly contributing to the performance gain. Furthermore, our results seem favorable [71, 144], but also slightly under perform [145] when compared to 2D segmentation literature.

A limitation of the segmentation framework is the overall under-segmentation the inferior vena cava, especially near the edges of the volume (Figure 3.4a). We suspect that this is caused by incomplete vessel information (i.e. incomplete visibility of vessel cross section), strongly influencing the Dice similarity index due to its large volume. The vena cava is often identified at the edge of an US volume, resulting in fewer paths to propagate its information during the forward pass [146]. Under-segmentation of a major vessel, weighting heavily on the points in the point cloud, is prone to reducing the registration accuracy as the registration is performed between partial point clouds. It would be interesting to investigate whether further optimization of the learning rate hyper-parameter contributes to a gain in segmentation performance.

As stated in Section 2.6.2, the loss function of the network was defined as 1 - mean Dice score (over fore - and background classes). When comparing the validation loss value, trained on the combined dataset (0.081, hence mean Dice 0.838, Table 3.2), with the actual performance on the

test set (Dice 0.773), under-performance is observed. Given that training on the combined dataset achieves higher Dice scores than either of the models trained on their own dataset, we expect that increasing the size of the dataset can improve performance.

Comparing the performance of the proposed CNN to the original 3D U-Net was not possible due to GPU memory limitations. However, we expect that the under-segmentation issues can be mitigated by either excluding the structures from the US acquisition, or by enlarging the receptive field of the network. The vena cava is not specifically necessary for accurate registration due to the distance with respect to the central lesions in the liver, hence excluding it might be beneficial. The receptive field of the network can be expanded by downsampling the US acquisitions even further, but also by implementing more convolutional layers or using a larger stride [146]. Performing more downsampling has an additional benefit of reducing computational load and time needed for segmentation predicting. Moreover, it is expected that smaller vessels become unidentifiable due to interpolation. This expectantly does not influence the registration heavily as the smaller vessels are removed during post-processing in the current setup.

4.1.2 Registration

The mean accuracy of our successful registrations is less favorable when compared to the values presented in literature (Table 1.1). Most of the presented literature assesses their performance based on the fiducial registration error (FRE), which is based on expert placement [69–71] of points on anatomical features. These anatomical landmarks are by definition located closer to the vessels or surface that is used during registration, making a skewed comparison to our TRE. Moreover, most of the better performing solutions are based on percutaneous US imaging. In transplantation literature [147, 148] it is well known that the volume of a liver imaged preoperatively can deviate up to 10 % when compared to an intraoperative measurement. Hence, there might be an influence of the laparotomy on the US vasculature volume and thus the similarity of the two point clouds. With respect to the amount of time it takes to perform a complete registration (69.74 ± 14.6 seconds, Table 3.5) is not as fast as presented by others in literature [69–71]. However, during the course of this study, the surgeons use the time after the initial US acquisition for clinical guidance of their resection approach. Hence, we do not foresee any clinical implementation issues with the amount of time needed for registrations in the presented setup.

One of the major factors influencing the TRE is the amount of similarity between the two point clouds generated from the extracted centerlines. As mentioned before, it is often seen that the preoperative model contains less vascular information than the intraoperative model. It is also seen that in certain cases the gallbladder and the inferior vena cava are recognized as vasculature by the segmentation algorithm. In this work we compensate for that by post-processing as described in section 2.6.4, eliminating the smaller vessels in both modalities and thus providing point clouds with similar information. What is often seen after manually adjusting the preoperative model is a seemingly overlapping registration on a single vessel (i.e. middle hepatic vein). However, it is then the case that there is a complete mismatch with the other vasculature. This supports the claim of further exploring post-processing. Instead of applying post-processing to both modalities, one could also provide a preoperative model of similar detail as the intraoperative model initially. This would lead to creation of two models, as the amount of detail needed to create this model exceeds the detail that is necessary for the surgeon to determine his resection strategy. The additional work involved with the creation of a secondary model can possibly be circumvented by training a neural network to automatically segment the preoperative scans. With some manual post-processing it is expectantly less effort to acquire two models.

However, even after post-processing, the algorithm is still prone to a lot of failed registrations. Only under certain circumstances (US to crop volume ratio larger than 2 and a big US acquisition excluding the vena cava and the gallbladder), satisfying (< 10 mm) registration results are achieved in certain cases, but not on average (mean successful TRE 12.29 ± 4.93). We have shown that manual adjustment of volume that is cropped after the coarse registration has a beneficial effect on the overall registration accuracy. However, improvement does not yet lead to clinically applicable results. A major factor influencing the cropping are the tool orientation and the one point translation that are used for the initial registration (Section 2.4). The initial registration determines which points are used for registration by means of cropping. Therefore, if the wrong points are included in the cropped volume, registration is deemed to fail. Figure 3.6 indicates that the quality of the registration improves when the US volume contains more US information. Sec-

tion 3.4 presents the gain that can be made by manually adjusting the region that is cropped. This indicates that there still is a lot of room for improvement in the amount of successful registrations that were performed in this study.

A shortcoming of this thesis, making it more challenging to compare with certain literature, is the sole use of TRE as a measure of registration accuracy. The main motivation for this study is to perform more accurate surgery, hence TRE was chosen because it is the clinical parameter which is most relevant in case of surgical navigation. It reports on how well the lesions overlap in the pre- and intraoperative modality. This study could be expanded with other performance measure to give more insight in the accuracy of the registration. As aforementioned, in literature [69–71] the FRE is often used when reporting on registration accuracy and would therefore be a viable option. Although, when used intraoperatively, which is the ultimate goal of this thesis, both the TRE and FRE require manual intervention to be calculated.

Many of the articles described in section 1.2.2 utilize breath holding to mitigate deformation in the vasculature segmentation of US. In our approach, we did not account for potential deformation due to respiratory motion. It would be interesting to compare the amount of liver motion before and after laparotomy, to give an insight into the necessity of motion compensation during open surgery. Given that the results of this study compare less favorable to literature [67, 69–72], it might be of interest to perform acquisition of the US data during surgery with the 3D probe, instead of making a 2D sweep. The 3D probe acquires a volume within 1.5 seconds, when properly timed within one out-breath of the patient. Our model shows similar performance on the 3D dataset as it does on the stacked 2D dataset, when trained on the combined dataset. Hence, we do not expect difficulties with the automatic segmentation of vasculature on 3D imaging. Another strategy would be attaching the reference sensor to the liver surface, in this manner movement of the liver can be compensated for. Although, this does not account for heterogeneous compression of the liver parenchyma.

4.2 Conclusion

This thesis presented a framework that can register hepatic vessels from preoperative imaging to intraoperative US imaging using a fast rigid transformation, utilizing deep learning based segmentation. Contrary to other registration techniques this foremost relies on the US sweep in the vicinity of the target lesion. The hepatic vasculature that is segmented automatically using a 3D U-Net acquires a Dice score of 0.773, when trained on a dataset containing stacked 2D and 3D volumes. Although lower than inter-observer variability, it is suited for the registration task using the CPD algorithm. The TREs of visually successful registrations compare less favorable to literature, but are near the clinical threshold of 10 mm. A challenge lies in registering all volumes successfully, as in this study 55 % of the registrations failed. It is often observed that with smaller US acquisitions the registration task becomes more challenging. Hence, several challenges remain in finding the optimal manner of cropping the right volume from the preoperative model, in order to maximize the similarity of both point clouds. Moreover, the shape of the acquired US volume has a large influence on the quality of the registration and it is therefore necessary to further explore successful characteristics of an US acquisition.

Chapter 5

Recommendations

The ultimate goal of the registration task posed in this thesis is to provide additional value to a surgeon during hepatic surgery, cutting down on surgery time and improving patient benefit. Based on the results presented in this thesis, several technical challenges are posed that expectantly improve accuracy of the automatic registration.

First, the quality of the segmentation, determining the centerlines in the intraoperative imaging, has a big impact on the accuracy of the registration. It is therefore necessary to exclude certain anatomy from the region that is acquired with the US scan, the gallbladder and the vena cava are respectively over- and undersegmented. This respectively leads to noise and missing information, directly influencing the registration. Therefore it would be of interest to standardize the manner in which the US volumes are acquired with respect to scan depth and anatomy that should be contained in the US acquisition. Moreover, due to the improved performance on the combined dataset and the gap between the (Dice) loss performance between the validation set and the Dice on the test, we expect that including more US scans will provide an easy performance gain.

Moreover, prior to training of the 3D U-Net the data is downsampled by means of interpolation, which is a time consuming process, in order to reduce the GPU memory burden. This process also has to take place when using the algorithm during live surgery. A potential way around this step would be increasing the stride that is used during training of the network. This makes that the input patches are downsampled as they are propagated though the network, allowing for input of the original volume during inference. However, it has to be investigated whether the segmentation performance remains similar and what the effect on the segmentation prediction speed of the network is.

Furthermore, the segmentation approach we present is mainly focused on general vascular segmentation. It might be of clinical interest to be able to differentiate between the hepatic and portal veins that are present in the liver. This provides the option of generating a 3D model solely based on US segmentation, which can then be used for guidance during surgery. One can think of this in the same manner as the preoperative model is used now during surgery. However, the full benefit of a 3D model lies in the possibility to estimate the tumor relation to the vasculature. To achieve this it would be optimal if tumor segmentation on US can be realized.

Second, a box spanning the total volume of the US acquisition is used to extract the region of interest from the preoperative model after initial registration. When a US volume is not acquired in a straight path, the volume of the box increases exponentially relative to the extra US data that is acquired (US to crop volume ratio). This results in a cropped preoperative point cloud that is very dissimilar to the intraoperative point cloud, making it more challenging to perform a successful registration. We expect that this can be mitigated by solely cropping from the part where there is US information available and discarding what is left. Here it will be interesting to solely crop the volume that is spanned by the US data, but also look into expanding the region of interest. The overlap of the initial registration might not be optimal and exclude necessary information, requiring a larger region of interest. Determining a FRE by manually selecting points in both modalities, will allow for quantification of a succesful registration, instead of solely visual inspection.

Another approach to acquiring similar point clouds can be pruning of the point clouds based on their graph representations. Chapman et al. [149] present an approach where they crop false branches from pulmonary anatomy, based on directed graphs. In our challenge it would be inter-

esting to extract the graphs from both the pre- and intraoperative imaging, followed by pruning to a similar graph, which is then used for registration.

Moreover, the registration approach in this thesis is based on rigid registration, which is a simplified representation of reality as the liver is prone to deformations during surgery. The elasticity of the liver is very challenging to predict, and one does not want to make false assumptions during surgery. Given that there is coherence in the deformations, it would be interesting to further explore these techniques. Approaches can either be traditional or also deep learning based [112].

Bibliography

[1] Nederlandse Kankerregistratie. *Cijfers over kanker*. 2018. URL: https://www.cijfersoverkanker.nl/.

[2] Khan AN, Macdonald S, and Amin Z. *Liver Metastases Imaging*. 2015. URL: https://emedicine.medscape.com/article/369936-overview.

[3] HB El-Serag and AC Mason. "Rising incidence of hepatocellular carcinoma in the United States". In: *New England Journal of Medicine* 340.10 (1999), pp. 745–750.

[4] J Ferlay et al. "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008". In: *International journal of cancer* 127.12 (2010), pp. 2893–2917.

[5] PF Christ et al. "Automatic liver and tumor segmentation of ct and mri volumes using cascaded fully convolutional neural networks". In: *arXiv preprint arXiv:1702.05970* (2017).

[6] EK Abdalla et al. "Improving resectability of hepatic colorectal metastases: expert consensus statement". In: *Annals of surgical oncology* 13.10 (2006), pp. 1271–1280.

[7] U Stein and PM Schlag. "Clinical, biological, and molecular aspects of metastasis in colorectal cancer". In: *Targeted Therapies in Cancer*. Springer, 2007, pp. 61–80.

[8] NB Ackerman. "The blood supply of experimental liver metastases: Distribution of hepatic artery and portal vein blood to" small" and" large" tumors". In: *Surgery* 66 (1969), pp. 1067–1072.

[9] RC Semelka et al. "Liver lesion detection, characterization, and effect on patient management: comparison of single-phase spiral CT and current MR techniques". In: *Journal of Magnetic Resonance Imaging* 7.6 (1997), pp. 1040–1047.

[10] TC Noone et al. "Common occurrence of benign liver lesions in patients with newly diagnosed breast cancer investigated by MRI for suspected liver metastases". In: *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 10.2 (1999), pp. 165–169.

[11] PJ Karhunen. "Benign hepatic tumours and tumour like conditions in men." In: *Journal of clinical pathology* 39.2 (1986), pp. 183–188.

[12] C Nicolau, L Bianchi, and R Vilana. "Gray-scale ultrasound in hepatic cirrhosis and chronic hepatitis: diagnosis, screening, and intervention". In: *Seminars in Ultrasound, CT and MRI*. Vol. 23. 1. Elsevier. 2002, pp. 3–18.

[13] IN Kochin et al. "Benign liver masses and lesions in children: 53 cases over 12 years". In: *IMAJ-Israel Medical Association Journal* 13.9 (2011), p. 542.

[14] P Jha et al. "Pediatric liver tumors–a pictorial review". In: *European radiology* 19.1 (2009), pp. 209–219.

[15] B Ohlsson et al. "Detection of hepatic metastases in colorectal cancer: a prospective study of laboratory and imaging methods." In: *The European journal of surgery= Acta chirurgica* 159.5 (1993), pp. 275–281.

[16] K Wernecke et al. "Detection of hepatic masses in patients with carcinoma: comparative sensitivities of sonography, CT, and MR imaging." In: *AJR. American journal of roentgenology* 157.4 (1991), pp. 731–739.

[17] T Helmberger et al. "Diagnosis and staging of liver metastases with imaging methods". In: *Der Chirurg; Zeitschrift fur alle Gebiete der operativen Medizen* 70.2 (1999), pp. 114–122.

[18] V Cantisani et al. “Detection of hepatic metastases from colorectal cancer: prospective evaluation of gray scale US versus SonoVue® low mechanical index real time-enhanced US as compared with multidetector-CT or Gd-BOPTA-MRI”. In: *Ultraschall in der Medizin-European Journal of Ultrasound* 31.05 (2010), pp. 500–505.

[19] T Albrecht et al. “Phase-inversion sonography during the liver-specific late phase of contrast enhancement: improved detection of liver metastases”. In: *American journal of roentgenology* 176.5 (2001), pp. 1191–1198.

[20] E Quaia et al. “Comparison of contrast-enhanced ultrasonography versus baseline ultrasound and contrast-enhanced computed tomography in metastatic disease of the liver: diagnostic performance and confidence”. In: *European radiology* 16.7 (2006), pp. 1599–1609.

[21] TJ Vogl et al. “Hepatocellular carcinoma: role of imaging diagnostics in detection, intervention and follow-up”. In: *RoFo: Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin* 174.11 (2002), pp. 1358–1368.

[22] DV Sahani and Sanjeeva P Kalva. “Imaging the liver”. In: *The oncologist* 9.4 (2004), pp. 385–397.

[23] K Kinkel et al. “Detection of hepatic metastases from cancers of the gastrointestinal tract by using noninvasive imaging methods (US, CT, MR imaging, PET): a meta-analysis”. In: *Radiology* 224.3 (2002), pp. 748–756.

[24] S Saini. “Imaging of the hepatobiliary tract”. In: *New England journal of medicine* 336.26 (1997), pp. 1889–1894.

[25] AL Simpson et al. “Texture analysis of preoperative CT images for prediction of postoperative hepatic insufficiency: a preliminary study”. In: *Journal of the American College of Surgeons* 220.3 (2015), pp. 339–346.

[26] DV Sahani et al. “Current status of imaging and emerging techniques to evaluate liver metastases from colorectal carcinoma”. In: *Annals of surgery* 259.5 (2014), pp. 861–872.

[27] KW Kim et al. “Small ($\leq$ 2 cm) hepatic lesions in colorectal cancer patients: detection and characterization on mangafodipir trisodium–enhanced MRI”. In: *American Journal of Roentgenology* 182.5 (2004), pp. 1233–1240.

[28] R Hammerstingl et al. “Diagnostic efficacy of gadoxetic acid (Primovist)-enhanced MRI and spiral CT for a therapeutic strategy: comparison with intraoperative and histopathologic findings in focal liver lesions”. In: *European radiology* 18.3 (2008), p. 457.

[29] E Van Cutsem et al. “Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up”. In: *Annals of oncology* 25.suppl_3 (2014), pp. iii1–iii9.

[30] C Couinaud. “Le foie”. In: *Etudes anatomiques et chirurgicales* (1957).

[31] K Numminen, Outi Sipilä, and Heikki Mäkisalo. “Preoperative hepatic 3D models: virtual liver resection using three-dimensional imaging technique”. In: *European journal of radiology* 56.2 (2005), pp. 179–184.

[32] B Nordlinger et al. “Hepatic resection for colorectal liver metastases. Influence on survival of preoperative factors and surgery for recurrences in 80 patients.” In: *Annals of surgery* 205.3 (1987), p. 256.

[33] P Schlag, P Hohenberger, and CH Herfarth. “Resection of liver metastases in colorectal cancer–competitive analysis of treatment results in synchronous versus metachronous metastases.” In: *European journal of surgical oncology: the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 16.4 (1990), pp. 360–365.

[34] J Scheele et al. “Indicators of prognosis after hepatic resection for colorectal secondaries.” In: *Surgery* 110.1 (1991), pp. 13–29.

[35] S Iwatsuki, B W Shaw Jr, and T E Starzl. “Experience with 150 liver resections.” In: *Annals of Surgery* 197.3 (1983), p. 247.

[36] J-N Vauthey et al. “Is extended hepatectomy for hepatobiliary malignancy justified?” In: *Annals of surgery* 239.5 (2004), p. 722.

[37] JW Rohen and Chihiro Yokochi. *Anatomie des Menschen: Photographischer Atlas der systematischen u. topographischen Anatomie*. Schattauer., 1993.

[38] GP Kanas et al. “Survival after liver resection in metastatic colorectal cancer: review and meta-analysis of prognostic factors”. In: *Clinical epidemiology* 4 (2012), p. 283.

[39] PC Simmonds et al. “Surgical resection of hepatic metastases from colorectal cancer: a systematic review of published studies”. In: *British journal of cancer* 94.7 (2006), p. 982.

[40] SL Wong et al. “American Society of Clinical Oncology 2009 clinical evidence review on radiofrequency ablation of hepatic metastases from colorectal cancer”. In: *Journal of Clinical Oncology* 28.3 (2009), pp. 493–508.

[41] SL Ong et al. “Efficacy and safety of microwave ablation for primary and secondary liver malignancies: a systematic review”. In: *European journal of gastroenterology & hepatology* 21.6 (2009), pp. 599–605.

[42] AEM van der Pool et al. “Stereotactic body radiation therapy for colorectal liver metastases”. In: *British Journal of Surgery: Incorporating European Journal of Surgery and Swiss Surgery* 97.3 (2010), pp. 377–382.

[43] GB Stillwagon et al. “194 hepatocellular cancers treated by radiation and chemotherapy combinations: toxicity and response: a Radiation Therapy Oncology Group Study”. In: *International Journal of Radiation Oncology* Biology* Physics* 17.6 (1989), pp. 1223–1229.

[44] SA Leibel et al. “A comparison of misonidazole sensitized radiation therapy to radiation therapy alone for the palliation of hepatic metastases: results of a Radiation Therapy Oncology Group randomized prospective trial”. In: *International Journal of Radiation Oncology* Biology* Physics* 13.7 (1987), pp. 1057–1064.

[45] AH Russell et al. “Accelerated hyperfractionated hepatic irradiation in the management of patients with liver metastases: results of the RTOG dose escalating protocol”. In: *International Journal of Radiation Oncology* Biology* Physics* 27.1 (1993), pp. 117–123.

[46] W Park et al. “Local radiotherapy for patients with unresectable hepatocellular carcinoma”. In: *International Journal of Radiation Oncology* Biology* Physics* 61.4 (2005), pp. 1143–1150.

[47] L Soler et al. “Real-time 3D image reconstruction guidance in liver resection surgery”. In: *Hepatobiliary surgery and nutrition* 3.2 (2014), p. 73.

[48] S Beller et al. “Image-guided surgery of liver metastases by three-dimensional ultrasound-based optoelectronic navigation”. In: *British Journal of Surgery: Incorporating European Journal of Surgery and Swiss Surgery* 94.7 (2007), pp. 866–875.

[49] O Heizmann et al. “Assessment of intraoperative liver deformation during hepatic resection: prospective clinical study”. In: *World journal of surgery* 34.8 (2010), pp. 1887–1893.

[50] T Lange et al. “3D ultrasound-CT registration of the liver using combined landmark-intensity information”. In: *International journal of computer assisted radiology and surgery* 4.1 (2009), pp. 79–88.

[51] DM Cash et al. “Concepts and preliminary data toward the realization of image-guided liver surgery”. In: *Journal of Gastrointestinal Surgery* 11.7 (2007), pp. 844–859.

[52] J D Westwood et al. “Development of a 3D visualization system for surgical field deformation with geometric pattern projection”. In: *Medicine Meets Virtual Reality 13: The Magical Next Becomes the Medical Now* 111 (2005), p. 172.

[53] J McGrath, David N Siegel, and David L Waldman. “Ultrasonography and GPS technology”. In: *Ultrasound Clinics* 8.2 (2013), pp. 201–212.

[54] H–H Hung et al. “Survival rates are comparable after radiofrequency ablation or surgery in patients with small hepatocellular carcinomas”. In: *Clinical Gastroenterology and Hepatology* 9.1 (2011), pp. 79–86.

[55] T Livraghi, H Mäkisalo, and P-D Line. “Treatment options in hepatocellular carcinoma today”. In: *Scandinavian Journal of Surgery* 100.1 (2011), pp. 22–29.

[56] PN Kim et al. “Planning ultrasound for percutaneous radiofrequency ablation to treat small ($\leq$ 3 cm) hepatocellular carcinomas detected on computed tomography or magnetic resonance imaging: a multicenter prospective study to assess factors affecting ultrasound visibility”. In: *Journal of Vascular and Interventional Radiology* 23.5 (2012), pp. 627–634.

[57] M Claudon et al. “Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (CEUS) in the liver–update 2012”. In: *Ultraschall in der Medizin-European Journal of Ultrasound* 34.01 (2013), pp. 11–29.

[58] L Solbiati et al. “Guidance and monitoring of radiofrequency liver tumor ablation with contrast-enhanced ultrasound”. In: *European journal of radiology* 51 (2004), S19–S23.

[59] MW Lee et al. “Percutaneous sonographically guided radio frequency ablation of hepatocellular carcinoma: causes of mistargeting and factors affecting the feasibility of a second ablation session”. In: *Journal of Ultrasound in Medicine* 30.5 (2011), pp. 607–615.

[60] FA Jolesz. *Intraoperative imaging and image-guided therapy*. Springer Science & Business Media, 2014. Chap. Introduction.

[61] J Nijkamp et al. “Image-guided navigation surgery for pelvic malignancies using electromagnetic tracking”. In: *Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling*. Vol. 9786. International Society for Optics and Photonics. 2016, p. 97862L.

[62] TM Peters. “Image-guidance for surgical procedures”. In: *Physics in Medicine & Biology* 51.14 (2006), R505.

[63] CJ Harvey and T Albrecht. “Ultrasound of focal liver lesions”. In: *European radiology* 11.9 (2001), pp. 1578–1593.

[64] JR Perks et al. “Comparison of peripheral dose from image-guided radiation therapy (IGRT) using kV cone beam CT to intensity-modulated radiation therapy (IMRT)”. In: *Radiotherapy and oncology* 89.3 (2008), pp. 304–310.

[65] M Fargier-Voiron et al. “Ultrasound versus Cone-beam CT image-guided radiotherapy for prostate and post-prostatectomy pretreatment localization”. In: *Physica Medica* 31.8 (2015), pp. 997–1004.

[66] S Beller et al. “Feasibility of navigated resection of liver tumors using multiplanar visualization of intraoperative 3-dimensional ultrasound data”. In: *Annals of surgery* 246.2 (2007), p. 288.

[67] M Fusaglia et al. “A novel ultrasound-based registration for image-guided laparoscopic liver ablation”. In: *Surgical innovation* 23.4 (2016), pp. 397–406.

[68] G P Penney et al. “Overview of an ultrasound to CT or MR registration system for use in thermal ablation of liver metastases”. In: *Proc. Medical Image Understanding and Analysis*. Vol. 1. 2001, p. 6568.

[69] H Haque et al. “Automated registration of 3D ultrasound and CT/MR images for liver”. In: *2016 IEEE International Ultrasonics Symposium (IUS)*. IEEE. 2016, pp. 1–4.

[70] C Weon et al. “Position tracking of moving liver lesion based on real-time registration between 2D ultrasound and 3D preoperative images”. In: *Medical physics* 42.1 (2015), pp. 335–347.

[71] W Wei et al. *Fast Registration for Liver Motion Compensation in Ultrasound-guided Navigation*. Jan. 2019.

[72] VM Banz et al. “Intraoperative image-guided navigation system: development and applicability in 65 patients undergoing liver surgery”. In: *Langenbeck’s archives of surgery* 401.4 (2016), pp. 495–502.

[73] TP Kingham et al. “Image-guided liver surgery: intraoperative projection of computed tomography images utilizing tracked ultrasound”. In: *HPB* 14.9 (2012), pp. 594–603.

[74] T Peter Kingham et al. “3D image guidance assisted identification of colorectal cancer liver metastases not seen on intraoperative ultrasound: results from a prospective trial”. In: *HPB* 20.3 (2018), pp. 260–267.

[75] C Askeland et al. “CustusX: an open-source research platform for image-guided therapy”. In: *International journal of computer assisted radiology and surgery* 11.4 (2016), pp. 505–519.

[76] A Rethy et al. “Anthropomorphic liver phantom with flow for multimodal image-guided liver therapy research and training”. In: *Int J Comput Assist Radiol Surg* 13.1 (Jan. 2018), pp. 61–72.

[77] R Brekken et al. “Registration of Real-Time 3-D Ultrasound to Tomographic Images of the Abdominal Aorta”. In: *Ultrasound Med Biol* 42.8 (Aug. 2016), pp. 2026–2032.

[78] R Manniesing et al. “Level set based cerebral vasculature segmentation and diameter quantification in CT angiography”. In: *Medical image analysis* 10.2 (2006), pp. 200–214.

[79] AAA Youssif, Atef Z Ghalwash, Amr S Ghoneim, et al. “Comparative study of contrast enhancement and illumination equalization methods for retinal vasculature segmentation”. In: *Cairo International Biomedical Engineering Conference*. 2006, pp. 1–5.

[80] J Mille and L D Cohen. “Deformable tree models for 2D and 3D branching structures extraction”. In: *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*. IEEE. 2009, pp. 149–156.

[81] J Yi and J B Ra. “A locally adaptive region growing algorithm for vascular segmentation”. In: *International Journal of Imaging Systems and Technology* 13.4 (2003), pp. 208–214.

[82] JA Noble and D Boukerroui. “Ultrasound image segmentation: a survey”. In: *IEEE Transactions on medical imaging* 25.8 (2006), pp. 987–1010.

[83] M Haft-Javaherian et al. “Deep convolutional neural networks for segmenting 3D in vivo multiphoton images of vasculature in Alzheimer disease mouse models”. In: *arXiv preprint arXiv:1801.00880* (2018).

[84] N Linde. “The machine that changed the world, episode 3”. In: *Documentary miniseries* 2 (1992).

[85] I Goodfellow, Y Bengio, and A Courville. *Deep learning*. MIT press, 2016.

[86] B Van Ginneken et al. “Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans”. In: *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*. IEEE. 2015, pp. 286–289.

[87] A Garcia-Garcia et al. “A review on deep learning techniques applied to semantic segmentation”. In: *arXiv preprint arXiv:1704.06857* (2017).

[88] H Chen et al. “Automatic detection of cerebral microbleeds via deep learning based 3d feature representation”. In: *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*. IEEE. 2015, pp. 764–767.

[89] AAA Setio et al. “Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks”. In: *IEEE transactions on medical imaging* 35.5 (2016), pp. 1160–1169.

[90] A Prasoon et al. “Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network”. In: *International conference on medical image computing and computer-assisted intervention*. Springer. 2013, pp. 246–253.

[91] Y LeCun et al. “Backpropagation applied to handwritten zip code recognition”. In: *Neural computation* 1.4 (1989), pp. 541–551.

[92] K Fukushima. “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position”. In: *Biological cybernetics* 36.4 (1980), pp. 193–202.

[93] A Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. “Imagenet classification with deep convolutional neural networks”. In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.

[94] GE Dahl, Tara N Sainath, and Geoffrey E Hinton. “Improving deep neural networks for LVCSR using rectified linear units and dropout”. In: *2013 IEEE international conference on acoustics, speech and signal processing*. IEEE. 2013, pp. 8609–8613.

[95] P Protopapas and KA Rader. *CS109A: Introduction to Data Science*. 2018. URL: `https://harvard-iacs.github.io/2018-CS109A/`.

[96] SR Rath. *Activation Functions in Neural Networks*. 2019. URL: `https://debuggercafe.com/activation-functions-in-neural-networks/`.

[97] Y LeCun et al. “A theoretical framework for back-propagation”. In: *Proceedings of the 1988 connectionist models summer school*. Vol. 1. CMU, Pittsburgh, Pa: Morgan Kaufmann. 1988, pp. 21–28.

[98] J Zacherl et al. “Current value of intraoperative sonography during surgery for hepatic neoplasms”. In: *World journal of surgery* 26.5 (2002), pp. 550–554.

[99] GP Penney et al. “Registration of freehand 3D ultrasound and magnetic resonance liver images”. In: *Medical image analysis* 8.1 (2004), pp. 81–91.

[100] Y Chen and G Medioni. “Object modelling by registration of multiple range images”. In: *Image and vision computing* 10.3 (1992), pp. 145–155.

[101] G Kaila et al. “Fusion of CTA and XA data using 3D centerline registration for plaque visualization during coronary intervention”. In: *Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling*. Vol. 9786. International Society for Optics and Photonics. 2016, p. 978606.

[102] A Myronenko and X Song. “Point set registration: Coherent point drift”. In: *IEEE transactions on pattern analysis and machine intelligence* 32.12 (2010), pp. 2262–2275.

[103] C Che, TS Mathai, and J Galeotti. “Ultrasound registration: a review”. In: *Methods* 115 (2017), pp. 128–143.

[104] WM Wells III et al. “Multi-modal volume registration by maximization of mutual information”. In: *Medical image analysis* 1.1 (1996), pp. 35–51.

[105] M Kaar et al. “Automatic patient alignment system using 3D ultrasound”. In: *Medical physics* 40.4 (2013), p. 041714.

[106] M Figl et al. “Deformable registration of 3D ultrasound volumes using automatic landmark generation”. In: *PloS one* 14.3 (2019), e0213004.

[107] JA Noble. *Reflections on ultrasound image analysis*. 2016.

[108] G Wu et al. “Unsupervised deep feature learning for deformable registration of MR brain images”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2013, pp. 649–656.

[109] BD de Vos et al. “End-to-end unsupervised deformable image registration with a convolutional neural network”. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2017, pp. 204–212.

[110] X Cao et al. “Deformable image registration based on similarity-steered CNN regression”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2017, pp. 300–308.

[111] X Yang et al. “Quicksilver: Fast predictive image registration–a deep learning approach”. In: *NeuroImage* 158 (2017), pp. 378–396.

[112] Y Hu et al. “Weakly-supervised convolutional neural networks for multimodal image registration”. In: *Medical image analysis* 49 (2018), pp. 1–13.

[113] S Shan et al. “Unsupervised end-to-end learning for deformable medical image registration”. In: *arXiv preprint arXiv:1711.08608* (2017).

[114] P Fischer et al. “Flownet: Learning optical flow with convolutional networks”. In: *arXiv preprint arXiv:1504.06852* (2015).

[115] X Zhu et al. “A snake-based method for segmentation of intravascular ultrasound images and its in vivo validation”. In: *Ultrasonics* 51.2 (2011), pp. 181–189.

[116] L Fischer et al. “The impact of virtual operation planning on liver surgery”. In: *Imaging Decisions MRI* 11.1 (2007), pp. 39–44.

[117] G Glombitza et al. “Virtual planning of liver resections: image processing, visualization and volumetric evaluation”. In: *International Journal of Medical Informatics* 53.2-3 (1999), pp. 225–237.

[118] MJ Zinser et al. “A paradigm shift in orthognathic surgery? A comparison of navigation, computer-aided designed/computer-aided manufactured splints, and “classic” intermaxillary splints to surgical transfer of virtual orthognathic planning”. In: *Journal of oral and maxillofacial surgery* 71.12 (2013), 2151–e1.

[119] P Grunert et al. “Computer-aided navigation in neurosurgery”. In: *Neurosurgical review* 26.2 (2003), pp. 73–99.

[120] DK Bae and S J Song. "Computer assisted navigation in knee arthroplasty". In: *Clinics in orthopedic surgery* 3.4 (2011), pp. 259–267.

[121] AL Simpson and TP Kingham. "Current evidence in image-guided liver surgery". In: *Journal of Gastrointestinal Surgery* 20.6 (2016), pp. 1265–1269.

[122] R Kikinis, S D Pieper, and K G Vosburgh. "3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support". In: *Intraoperative imaging and image-guided therapy*. Springer, 2014, pp. 277–289.

[123] LE Bø et al. "Versatile robotic probe calibration for position tracking in ultrasound imaging". In: *Physics in Medicine & Biology* 60.9 (2015), p. 3499.

[124] A Rethy et al. "Anthropomorphic liver phantom with flow for multimodal image-guided liver therapy research and training". In: *International journal of computer assisted radiology and surgery* 13.1 (2018), pp. 61–72.

[125] J Premaladha and KS Ravichandran. "Novel approaches for diagnosing melanoma skin lesions through supervised and deep learning algorithms". In: *Journal of medical systems* 40.4 (2016), p. 96.

[126] G Litjens et al. "A survey on deep learning in medical image analysis". In: *arXiv preprint arXiv:1702.05747* (2017).

[127] O Çiçek et al. "3D U-Net: learning dense volumetric segmentation from sparse annotation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2016, pp. 424–432.

[128] O Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2015, pp. 234–241.

[129] E Gibson et al. "NiftyNet: a deep-learning platform for medical imaging". In: *Computer methods and programs in biomedicine* 158 (2018), pp. 113–122.

[130] JS Bergstra et al. "Algorithms for hyper-parameter optimization". In: *Advances in neural information processing systems*. 2011, pp. 2546–2554.

[131] J Bergstra and Y Bengio. "Random search for hyper-parameter optimization". In: *Journal of Machine Learning Research* 13.Feb (2012), pp. 281–305.

[132] I Czogiel, Karsten Luebke, and Claus Weihs. *Response surface methodology for optimizing hyper parameters*. Universitätsbibliothek Dortmund, 2006.

[133] F Hutter. "Automated configuration of algorithms for solving hard computational problems". PhD thesis. University of British Columbia, 2009.

[134] G Litjens et al. "Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge". In: *Medical image analysis* 18.2 (2014), pp. 359–373.

[135] T Heimann et al. "Comparison and evaluation of methods for liver segmentation from CT datasets". In: *IEEE transactions on medical imaging* 28.8 (2009), pp. 1251–1265.

[136] S Klein et al. "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information". In: *Medical physics* 35.4 (2008), pp. 1407–1417.

[137] AA Taha and A Hanbury. "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool". In: *BMC medical imaging* 15.1 (2015), p. 29.

[138] LR Dice. "Measures of the amount of ecologic association between species". In: *Ecology* 26.3 (1945), pp. 297–302.

[139] T–C Lee, Rangasami L Kashyap, and Chong-Nam Chu. "Building skeleton models via 3-D medial surface axis thinning algorithms". In: *CVGIP: Graphical Models and Image Processing* 56.6 (1994), pp. 462–478.

[140] T Lange et al. "Vessel-based non-rigid registration of MR/CT and 3D ultrasound for navigation in liver surgery". In: *Computer Aided Surgery* 8.5 (2003), pp. 228–240.

[141] PJ Besl and N D McKay. "Method for registration of 3-D shapes". In: *Sensor fusion IV: control paradigms and data structures*. Vol. 1611. International Society for Optics and Photonics. 1992, pp. 586–606.

[142] GP Penney et al. "A stochastic iterative closest point algorithm (stochastICP)". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2001, pp. 762–769.

[143] AH Mahnken, Kai E Wilhelm, Jens Ricke, et al. *CT-and MR-guided Interventions in Radiology*. Vol. 22. Springer, 2009.

[144] S Milko, Eigil Samset, and Timor Kadir. "Segmentation of the liver in ultrasound: a dynamic texture approach". In: *International Journal of Computer Assisted Radiology and Surgery* 3.1-2 (2008), p. 143.

[145] D Mishra et al. "Segmentation of vascular regions in ultrasound images: A deep learning approach". In: *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE. 2018, pp. 1–5.

[146] W Luo et al. "Understanding the effective receptive field in deep convolutional neural networks". In: *Advances in neural information processing systems*. 2016, pp. 4898–4906.

[147] PRO Salvalaggio et al. "Liver graft volume estimation in 100 living donors: measure twice, cut once". In: *Transplantation* 80.9 (2005), pp. 1181–1185.

[148] T Schroeder et al. ""All-in-one" imaging protocols for the evaluation of potential living liver donors: comparison of magnetic resonance imaging and multidetector computed tomography". In: *Liver transplantation* 11.7 (2005), pp. 776–787.

[149] BE Chapman, Holly P Berty, and Stuart L Schulthies. "Automated generation of directed graphs from vascular segmentations". In: *Journal of biomedical informatics* 56 (2015), pp. 395–405.