

DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

**EXTRACTION OF VASCULAR TREES FOR
LIVING DONATED LIVER TRANSPLANTATION**

by
Parvin HUSEYNOVA

July, 2019

İZMİR

EXTRACTION OF VASCULAR TREES FOR LIVING DONATED LIVER TRANSPLANTATION

**A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Master of Science
of Electrical and Electronics Engineering**

**by
Parvin HUSEYNOVA**

July, 2019

İZMİR

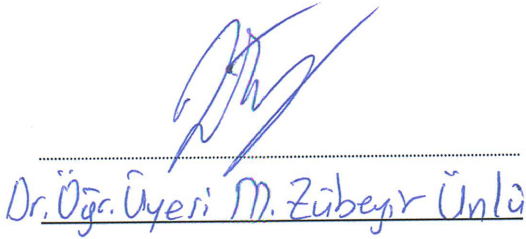
M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled “**EXTRACTION OF VASCULAR TREES FOR LIVING DONATED LIVER TRANSPLANTATION**” completed by **PARVIN HUSEYNOVA** under supervision of **ASSOC. PROF. DR. ALPER SELVER** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

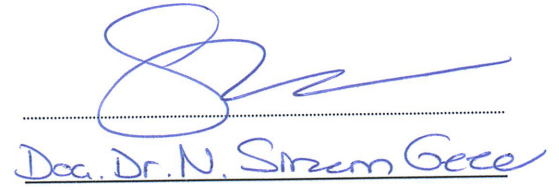


Assoc. Prof. Dr. Alper SELVER

Supervisor



(Jury Member)



(Jury Member)



Prof. Dr. Kadriye ERTEKİN

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my advisor, Assoc. Prof. Dr. Alper Selver for his constant support, valuable suggestions and guidance throughout this work. In addition, I would like to thank Assoc. Prof. Dr. Naciye Sinem Gezer for her shared expertise on medical subjects, her valuable time and patience.

This thesis was supported in part by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under Grant 116E133.

Last but not least, I would like to thank my dear husband Özgür Bulucu, for his encouragement, patience and trust.

Parvin HUSEYNOVA

EXTRACTION OF VASCULAR TREES FOR LIVING DONATED LIVER TRANSPLANTATION

ABSTRACT

Pre-surgical evaluations of living donated liver transplantation require accurate segmentation of liver vasculature. Expert radiologists carry out this procedure either manually or using semi-automatic software tools. Manual delineation is very time consuming and tedious work and segmentation accuracy is largely dependent on the expert's abilities and very susceptible to human error. Semi-automatic methods are faster, however require advanced interaction mechanisms and iterative optimization. Thus, there is a need for automated methods.

Unfortunately, the vascular tree of the liver is very complex and show high variability. Moreover, the contrast-enhanced images may contain significant amount of artifacts and task associated difficulties. Therefore, the development of a fully automatic method becomes a challenging task. Currently, there is no well-established datasets for comparative analysis of existing methods. This makes it hard to propose improvements due to the lack of qualitative analysis of different techniques on a benchmark dataset.

In this thesis, first, a database, which consists of 35 abdominal computed tomography angiography datasets, is collected and hepatic and portal veins are annotated manually. The process is carefully supervised by an experienced radiologist. Some of the well-known vessel segmentation methods were tested and their performances were analyzed. Finally, deep learning based methods were applied to reflect the performance of emerging deep models. After extensive experimentation, DeepMedic architecture is shown to achieve the best performance. An automatic system, which employs combinations of multi planar reconstructions, is developed. The obtained results are shown to outperform both the existing methods and the individual utilization of deep models.

Keywords: Liver vascular tree, segmentation, computed tomography

CANLI VERİCİLİ KARACİĞER NAKLİ İÇİN KARACİĞER DAMAR AĞACI BÖLÜTLEME

ÖZ

Canlı vericili karaciğer nakli öncesinde karaciğer damar ağacının doğru bir şekilde bölütlenmesi gerekir. Uzman radyologlar bu prosedürü manuel olarak veya yarı otomatik yazılım araçlarını kullanarak gerçekleştirmektedir. Elcil etiketleme çok zaman alıcı ve sıkıcı bir işlemdir ve yöntemin doğruluğu, büyük ölçüde uzmanın yeteneklerine bağlıdır ve insan hatasına oldukça duyarlıdır. Yarı otomatik yöntemler daha hızlıdır, ancak gelişmiş etkileşim mekanizmaları ve yinelemeli optimizasyon gerektirir. Bu nedenle, otomatik yöntemlere ihtiyaç vardır.

Ne yazık ki, karaciğerin damar ağaçları oldukça karmaşık ve hastalar arası değişkenlik gösteren yapılardır. Üstelik, kontrastı artırılmış görüntüler önemli miktarda gürültü içerebilmektedirler. Bu nedenlerden dolayı, tam otomatik bir yöntemin geliştirilmesi zorlu bir görev haline gelir. Halihazırda, mevcut yöntemlerin karşılaştırmalı analizi için sağlam bir veri seti bulunmamaktadır. Bu durum, bir değerlendirme veri kümesi üzerinde farklı tekniklerin nicel analizlerinin yapılmamış olmasından dolayı iyileştirme önermeyi zorlaştırmaktadır.

Bu tez çalışmasında ilk olarak 35 abdominal bilgisayarlı tomografi anjiyografi veri kümesinden oluşan bir veri tabanı toplanmış, hepatik ve portal damarlar elcil olarak etiketlenmiştir. İşlem deneyimli bir radyolog tarafından dikkatle denetlenmiştir. Bilinen damar bölütleme yöntemlerinden bazıları test edilmiş ve performansları analiz edilmiştir. Son olarak, geliştirilen derin modellerin performansını yansıtmak için derin öğrenme temelli yöntemler uygulanmıştır. Kapsamlı deneylerden sonra, DeepMedic mimarisinin en iyi performansı sağladığı görülmüştür. Çok düzlemsel rekonstrüksiyonların kombinasyonlarını kullanan bir otomatik sistem geliştirilmiştir. Elde edilen sonuçların hem mevcut yöntemlerden hem de derin modellerin bireysel kullanımından daha iyi performans gösterdiği gözlemlenmiştir.

Anahtar kelimeler: Karaciğer damar ağacı, bölütleme, bilgisayarlı tomografi

CONTENTS

	Page
M. Sc THESIS EXAMINATION RESULT FORM.....	ii
ACKNOWLEDGEMENTS	iii
ABSTRACT.....	iv
ÖZ	v
LIST OF FIGURES	viii
LIST OF TABLES	x
CHAPTER ONE – INTRODUCTION	1
CHAPTER TWO – LITERATURE ANALYSIS.....	6
CHAPTER THREE – EVALUATION METRICS	10
CHAPTER FOUR - VESSEL EXTRACTION AND EXTRICATION FOR LIVER ANALYSIS DATASET - VEELA.....	13
3.1 VEELA.....	13
3.2 Extraction of Liver Vascular Trees	13
3.3 Challenges Encountered During Labeling	18
CHAPTER FIVE – APPLICATION OF COMMON VESSEL SEGMENTATION METHODS AND RESULTS	19
5.1 Used Methods.....	19

5.1.1 Hessian Eigenvalues – Frangi Method	19
5.1.2 Hessian Eigenvalues – Jerman Method	20
5.1.3 Multilevel Thresholding – Otsu Method	21
5.1.4 K-Means Clustering.....	22
5.2 Results	22
 CHAPTER SIX – TRAINING DEEPMEDIC	25
6.1 Training DeepMedic.....	25
6.2 Combining DeepMedic Results.....	31
 CHAPTER SEVEN – CONCLUSION	34
 REFERENCES	35

LIST OF FIGURES

	Page
Figure 1.1 Examples of quality losses and artifacts. (a) Quality loss due to miscalculation of timing, (b) artefact causing top part of the parenchyma to appear brighter, (c) high inter slice thickness, (d) transient hepatic attenuation, (e) non uniform attenuation, (f) beam hardening	5
Figure 4.1 An example labeling process in 3D slicer	14
Figure 4.2 Labeled example slices.....	15
Figure 4.3 Example 3D models for separate vein systems for (a) a hepatic vein system, (b) a portal vein system, (c) close up of a hepatic vein, (d) close up of a portal vein.....	16
Figure 4.4 Example 3D models	17
Figure 5.1 Results of the segmentation algorithms. (a) Ground truth, (b) Frangi, (c) Jerman, (d) multilevel Otsu, (e) K-means clustering	23
Figure 6.1 Model of DeepMedic The kernel sizes of both pathways are given as 53 for the purpose of reducing figure's size and actual kernels are of 33 size (Kamnitsas, et al., 2017).....	25
Figure 6.2 Training process for datasets in three orthogonal planes: (a) Mean accuracy, (b) sensitivity, (c) precision and (d) mean dice values over 35 epochs. Smaller figures only show mean results for each epochs and ignores individual subepoch results.	26
Figure 6.3 Final DSC values for training sets, test sets and all sets in (a) axial, (b) coronal and (c) sagittal planes	27
Figure 6.4 Difference between ground truth and results. Thin lines in images indicate disagreement in vessel borders between algorithm results and ground truth labels.....	27
Figure 6.5 Comparison between DSC value and Area values with $\alpha = 1$ and $\alpha = 2$. (a) Results on axial plane, (b) coronal plane and (c) sagittal plane	28
Figure 6.6 Example results of training process for axial, coronal and sagittal planes, and volumes that were obtained from the intersection of these three results	29

Figure 6.7 DSC results of all training that were performed in axial, coronal and sagittal planes, and DSC results of the volumes that were obtained with intersections and unions of these results	30
Figure 6.8 DSC results of axial, coronal, sagittal planes, intersection volumes and union volumes where FPs are removed.....	30
Figure 6.9 Change in parameters in terms of (a) Dice coefficient, (b) true positive rate, (c) false positive rate and (d) false negative rate.....	32
Figure 6.10 Change in parameters in terms of (a) Dice coefficient, (b) true positive rate, (c) false positive rate and (d) false negative rate for all datasets	33



LIST OF TABLES

	Page
Table 5.1 Mean DSC results and standard deviations for vessel segmentation methods	24



CHAPTER ONE

INTRODUCTION

Accurate analysis of liver vasculature in three dimensions (3D) is essential for variety of medical procedures – e.g. preoperative planning, treatment of hepatic diseases, computer-aided diagnosis. Moreover, it is also used for medicine education, which significantly benefits from advanced models and realistic simulations. The details of some particular applications can be given as follows:

Living donated liver transplantation is a surgical operation where part of the donor's liver is transplanted to another patient. Pre-surgical planning for transplantation requires precise knowledge of liver vascular morphology (Selver et al., 2008). Furthermore, the donor's suitability for the operation can be determined by carrying out a volumetric approximation of the liver with vasculature analysis, thus providing an understanding of postoperative liver function (Selle, Preim, Schenk, & Peitgen, 2002).

Localization of the liver lesions is based on the lesion's position relative to the surrounding hepatic vessels (Fasel et al., 1988). Hepatocellular carcinoma is the most common liver cancer and is the primary source of cancer-related deaths worldwide (Balogh et al., 2016). The most common therapy for hepatocellular carcinoma is radiofrequency treatment. Blood vessels around the tumor can disturb the thermal process by acting as coolers. Therefore, the knowledge of the vessels neighboring the tumor is crucial in order to determine the most suitable path for the insertion of the interstitial applicator (Esneault, Lafon, & Dillenseger, 2010).

In the late stages of the liver cancer, the only cure is the liver resection, which is a surgical operation that requires the doctors to remove the diseased tissue. Such an operation necessitates the patients to meet a list of preconditions. Patients' suitability for the procedure depends on the position and the size of the tumor, as well as the postoperative liver function, which is another critical point to consider (Reitinger, Bornik, Beichel, & Schmalstieg, 2006). In this regard, detailed analysis of the liver vasculature should consist of all three vessel sub-systems (Abdel-Misih, & Bloomston, 2010):

1. Hepatic vessels that transport de-oxygenated blood from the liver to the heart,
2. Portal vessels that transport nutrient-rich blood from intestines to the liver,
3. The hepatic artery carries oxygenated blood from the heart to the liver.

Occlusion of one of these vessels can obstruct the drainage or supply for some liver parts (Lehmann et al., 2008).

Computer tomography angiography (CTA) is the most common radiographic imaging technique to acquire the liver vascular tree. An intra-vascular contrast agent is injected to the patient before the scanning, which is more visible on x-ray images and causes vessels to appear brighter compared to the liver parenchyma. From image processing point of view, this process enables enhanced intensity levels for blood filled regions and makes segmentation easier for the main branches.

During routine clinical workflow, vessel segmentation is mostly carried out by radiologists, who manually annotate all vessels on abdominal multi-slice CTA slices. This process is a tedious, time-consuming work that is susceptible to human error. Therefore, automatic liver-vessel segmentation is an area of interest for a lot of researchers. Vessel structure of the liver is very complex and varies a lot for each individual, to the point where it differs from commonly accepted structures. Additionally, the existence of tumors and previous operations on the liver can alter the usual vessel structure (Conversano et al., 2011).

Semi-automatic methods like region growing, fast marching or centerline tracking ease and speed up the task, reduce the manual interaction and lower the inter-expert variability (Lesage, Angelini, Bloch, & Funka-Lea, 2009). Region growing is method with lower computational complexity and relatively higher speed, whose main principle is to collect voxels that have similar intensity levels together. Jiang et al. (2013), proposed a region growing vessel segmentation method based on the Fourier transform of the vascular region. However, these methods require advanced interaction mechanisms with at least one expert clinician and iterative optimization of the segmentation results to reach the desired level of accuracy.

Unfortunately, angiographic acquisitions show high variations in image quality and, contrast level of vessels due to the changes in imaging time and conditions. We classify the reasons of these variations in image quality as follows:

1. The miscalculation of timing: The time period between the injection of the contrast agent and the beginning of the scanning is miscalculated. Therefore, contrast agent is not in the desired location of the vessel during the scanning and enhancement of the vessels are not achieved. As a result, obtained CTA images show poor quality and vessel borders are not clearly identifiable. Figure 1.1 (a) demonstrates an example CTA slice with miscalculated timing, where vessels are almost invisible.
2. Artefacts: Artefacts and noise might further drop the quality of rendering. Hepatic and portal vessels may look connected, some parts of parenchyma may appear brighter and look like vessels, noise may obscure the appearance of vessels. As an example, Figure 1.1 (b) presents a CTA slice, where top part of the liver appears even brighter than the actual vessels.
3. Interslice distance: High slice thickness alters the adequate appearance of vessels in 3D. Figure 1.1 (c) is the coronal view of a liver that was reconstructed from axial plane images. High interslice distance causes low quality vessel appearance, which is more visible on the left side of the inferior vena cava.
4. Transient hepatic attenuation: The liver is the only organ with dual blood supply. Approximately 70% of its perfusion comes from the portal vein, and 30% comes from the hepatic artery (Wong, Desser, & Jeffrey, 2008). Transient hepatic attenuations are areas of enhancement on CTA that occur as results of localized variations in the proportion of hepatic arterial and portal venous blood supply. Figure 1.1 (d) demonstrates an example, where transient hepatic attenuation is marked with red arrows. Enhanced area is similar looking to the surrounding vessel tissue and may cause false vessel segmentation results.
5. Non-uniform attenuation: Non-uniform blood velocity in vessel complicates the vascular imaging (Murphy, Aghayev, & Steigner, 2018). A short acquisition time is preferable in order to obtain uniform opacification on vessels. Investigation the example given in Figure 1.1 (e), it can be seen that vessels on the right hand side of the liver have lower intensity values. This phenomenon

may misguide the methods whose operation principle are based on vessels' intensities.

Besides the angiographic acquisition, computed tomography itself has general drawbacks and limitations. Artefacts may seriously degrade the imaging quality. Physics-based artefacts are caused due to the physical process of CT acquisition. Patient-based artefacts result from such reasons as patient movements or metal objects that are present in or on the patient. Scanned-based artefacts are caused due to the imperfections in scanner functions (Barrett & Keat, 2004). Beam hardening, partial volume and undersampling are physics-based artefacts. Figure 1.1 (f) presents a slice with beam hardening, an artefact that occurs when the beam passing through the tissues with different density causes dark density lines (Chen & Chen, 1999).

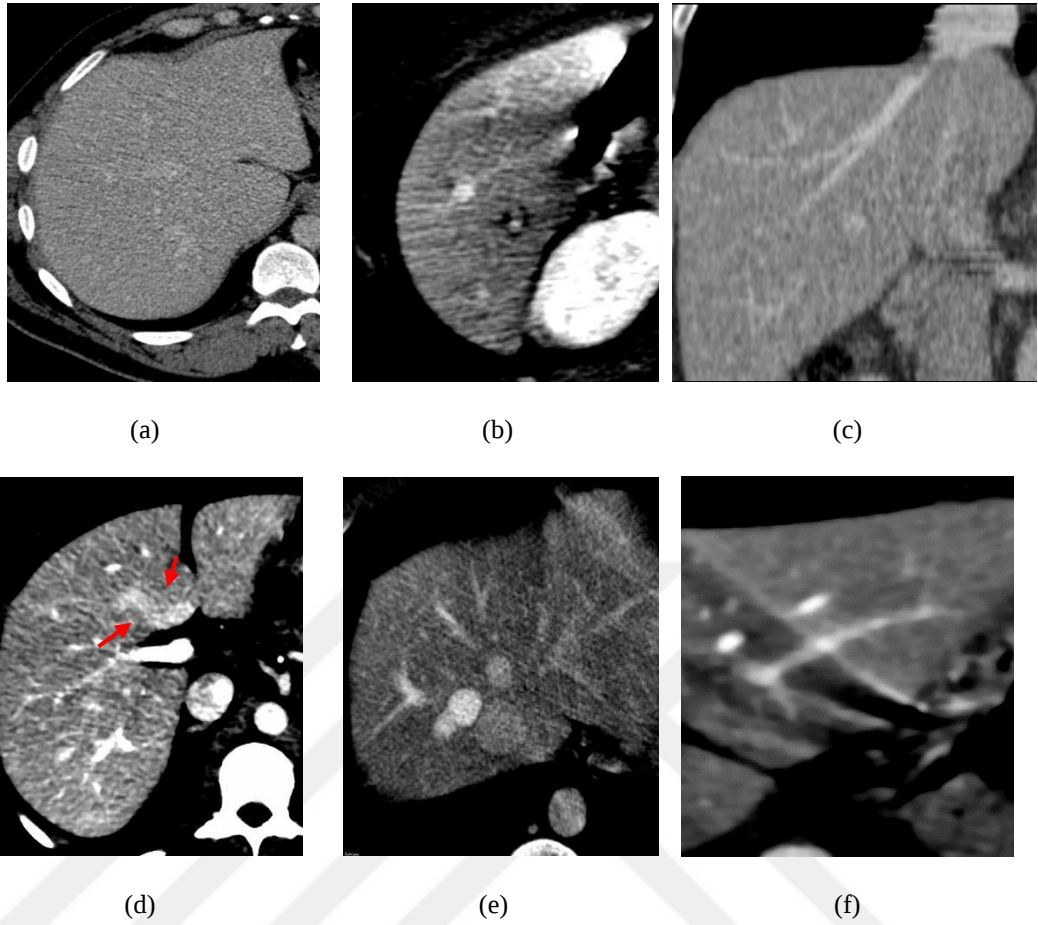


Figure 1.1 Examples of quality losses and artifacts. (a) Quality loss due to miscalculation of timing, (b) artefact causing top part of the parenchyma to appear brighter, (c) high inter slice thickness, (d) transient hepatic attenuation, (e) non uniform attenuation, (f) beam hardening

This thesis consists of six chapters. First chapter describes the purpose of the study, clinical importance of liver vessel segmentation. A literature analysis is made on vessel segmentation methods and grand challenges are described in chapter two. Third chapter presents the various performance evaluation metrics for vessel segmentation methods. VEELA dataset is described in fourth chapter. The process of creation of the dataset, and encountered problems are presented. Results of four segmentation methods after tested on the abdominal CTA dataset and compared to the ground truth are described in chapter five. Sixth chapter describes the deep learning based semantic segmentation method DeepMedic, this method's training process for the vessel segmentation problem, and the developed approach for combining the DeepMedic results and increasing the segmentation performance. Seventh chapter is the conclusion.

CHAPTER TWO

LITERATURE ANALYSIS

Many methods were suggested for this task.

One of the most commonly used vessel segmentation methods was proposed by Frangi & Niessen (1998). This method focuses on detection light tubular structures on a dark background (or dark tubular structures on light background) and depends on the eigenvalues decomposition of the Hessian matrix of the 3D data. If three eigenvalues ($\lambda_1, \lambda_2, \lambda_3$) are assumed to follow $|\lambda_3| > |\lambda_2| > |\lambda_1|$ for each voxel, then the vesselness condition can be written as $|\lambda_1| \approx 0, |\lambda_1| < |\lambda_2|, \lambda_2 \approx \lambda_3$. Many vessel segmentation methods are based on Hessian matrix eigenvalues. However, these methods were not tested on liver vasculature due to its complex structure.

Many methods depend on the intensity difference between vessels and parenchyma, thus in cases where vessel intensity changes throughout the CTA slices due to before mentioned artefacts segmentation performances drop significantly.

A lot of new studies on biomedical segmentation field make use of deep learning algorithms. DeepMedic (Kamnitsas et al., 2017) is a 3D deep learning algorithm that was developed for brain lesion segmentation, and obtained 89.8% DSC value for BRATS2015 challenge. DeepVess (Haft-Javaherian et al., 2019) focuses on brain vasculature segmentation on 3D in vivo multiphoton microscopy images and managed to get 81.63% DSC scores with post-processing. 3D U-net (Çiçek, Abdulkadir, Lienkamp, Brox, & Ronneberger, 2016) is deep learning algorithm that can be trained with sparse annotation, the algorithm was tested on Xenopus kidney dataset and achieved an average of 0.863 Intersection over Union (IoU). Many other segmentation algorithms are based on 3D-Unet structure. Overall, deep learning algorithms show promising results. They are very good at learning intricate structures and determining the relations between several modalities. Furthermore, they are fed with the raw data, do not require the user to determine features of the data. However, liver vessel segmentation with deep learning has not been a subject in any recent publications.

Besides the complexity of the problem, a gold standard ground truth dataset does not exist. There are known vessel datasets. Peripheral Artery: Vein Enhanced Segmentation (PAVES) includes magnetic resonance angiography volumes of the lower leg where arterial and venous vasculature can be observed. Digital retinal images for vessel Extraction (DRIVE) dataset present 20 train and 20 test jpeg compressed images of the retinal vasculature. However, to the best of our knowledge, a publicly available liver vasculature dataset does not exist. Many liver-vessel segmentation papers test their methods on some kind of annotated data. (Lorigo et al., 2001) state that the performance of the developed liver-vessel segmentation method was qualitatively assessed due to the difficulty of obtaining ground truth for a dataset of this level complexity.

A lot of papers are getting published every year on various subjects in the biomedical field. Many of these papers' tasks were tackled by previous researches. However, obtaining a fair comparison of these methods is a difficult task. New papers often times include the performance results of older methods; however adaptation of the methods may differ and present unfair performance results for otherwise successful methods. Furthermore, evaluation of the developed methods are often times carried out on datasets that are not publically available and the credibility of these datasets cannot be assessed. Usage of the "bad" datasets lead to misleading achievements (Jimenez-del-Toro, Müller, & Krenn, 2016). Grand challenges provide public datasets for shared problems, thus allowing the comparison of many researchers' works on uniform fields. Many of the submitted methods are published, therefore creating novel solutions for long standing problems.

Gliomas are the most common brain tumors (Schwartzbaum, L Fisher, Aldape, & Wrensch, 2006). There is a lot of research targeting the disease; however, there is still need for accurate clinical diagnosis methods. Multimodal Brain Tumor Segmentation Challenge (BRATS), organized in MICCAI2012 (Menze et al., 2015) for this issue. Participants were provided with an annotated dataset consisting of 65 multi-contrast magnetic resonance (MRI) scans of glioma patients and 65 artificial scans generated by a simulation software. Challenge received multiple algorithms that perform better on different tumor sub-regions. An overall high performance was achieved by fusing

together several successful algorithms, and fused algorithm performed better on all sub regions compared to the individual methods. Annotated dataset is still publically available, allowing the subject to be tackled by many researchers.

Shape, size and appearance variations of the anatomical structures are often the results of some underlying diseases. Segmentation of the structures is done manually and there is need for an automated method. VISCERAL anatomy benchmark (Jimenez-del-Toro et al., 2016) included the results of state of art methods for segmentation of anatomical structures. All algorithms were implemented in a cloud and were tested by challenge administrators. A manually annotated dataset with 120 computed tomography (CT) and MRI volumes that contained 1295 anatomical structures in total was prepared for the challenge. Unlike the previous similar challenges, the test set was not available for the participants, furthermore challenge aimed the segmentation of the all anatomical structures as opposed to specific organs and a region of interests were not provided. Results of the participating algorithms were fused in order to create a Silver Corpus and the segmentation on test dataset performed better than all algorithms in the benchmark. The dataset, Silver Corpus and the evaluation framework is still publically available.

Breast cancer is the most frequent cancer type among women (Bándi, Geessink, & Manson, 2019). CAMELYON17 challenge was organized in order to provide a uniform comparison field for automated cancer metastases detection in lymph nodes algorithms. A train set consisting of 899 whole-side images (WSIs) was provided for the participants and submitted algorithms were tested on a test set with 500 WSIs. 37 algorithms were submitted, all methods were based on deep learning and post-processing. The kappa metric was used for the evaluation, participant performances varied between 0.89 to 0.13 kappa values and combination algorithms formed with top performing methods lead to the best performance results with up to 0.93 kappa values.

Overall, it can be concluded that apart from providing credible datasets and fair evaluation for the researchers, challenges often times useful in generating ensemble algorithms that perform better than individual methods. Many other challenges on various topics were organized throughout the years, (i.e. MRI cardiac multi-structure segmentation (Bernard, Lalande, & Zotti, 2018), retinal optical coherence tomography

fluid detection and segmentation (Bogunović, Venhuizen, & Klimesch, 2019), segmentation of white Matter hyperintensities (Kuijf, Biesbroek, & de Bresser, 2019), however to the best of our knowledge there have been no challenges on the vessel segmentation.



CHAPTER THREE

EVALUATION METRICS

The success of vessel segmentation methods is evaluated with respect to ground truth data that was manually annotated by an expert radiologist.

Evaluation metrics mainly use values true positive (TP), false positive (FP), true negative (TN) and false negative (FN). Positive refers to voxels that are classified as vessel and negative refers to voxels that are classified as background.

Most commonly used performance evaluation metrics for blood vessel segmentation are accuracy, sensitivity (true positive rate), specificity and dice similarity coefficient (DSC). Accuracy is the ratio of true predictions to all predictions.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.1)$$

Sensitivity measures the ratio of correctly identified vessels to all vessels.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (3.2)$$

DSC is an overlap measure where twice the number of voxels in the intersection of segmentation and ground truth data is divided by the total number of true voxels in both segmentation and ground truth data.

$$\text{DSC} = \frac{|X \cap Y|}{|X|+|Y|} = \frac{2TP}{FP+FN+2TP} \quad (3.3)$$

where X and Y are two binary datasets.

Specificity is another common measure that indicates the ratio of correctly labeled negative voxels. Low specificity value for a segmentation algorithm indicates a high false positive rate (FPR) (Moccia, De Momi, El Hadji, & Mattos, 2018).

$$\text{Specificity} = \frac{TN}{TN+FP} \quad \text{and} \quad \text{FPR} = \frac{FP}{TN+FP} \quad (3.4)$$

Positive predictive value (PPV) is the probability that a voxel that is identified as positive by the algorithm belongs to a vein. Negative predictive value (NPV) is the probability that a voxel that is identified as negative by the algorithm is a part of the background.

$$PPV = \frac{TP}{TP+FP} \quad \text{and} \quad NPV = \frac{TN}{TN+FN} \quad (3.5)$$

Apart from comparing voxels of the segmentation and ground truth, considering the structural constraint of a vessel is useful for the evaluation. Gegúndez-Arias, Aquino, Bravo, & Marín (2012) propose three evaluation metrics. It is known that the vessels are continuous structures, hence it is expected that the number connected components in the segmentation results to be very few, ideally one. Therefore the connectivity measure penalizes the algorithm according to the difference between connected components in the result and in the ground truth.

$$\text{Connectivity} = 1 - \min\left(1, \frac{|\#_C(S_G) - \#_C(S)|}{\#(S_g)}\right) \quad (3.6)$$

Where min is the minimum function, S_G is the ground truth data and S is the segmentation result. $\#_C(S_G)$ and $\#_C(S)$ are the number of connected components in ground truth and segmentation result respectively. $\#(S_g)$ is the total number of voxels of the ground truth.

Since ground truth images are often manually annotated vessel borders and width may show slight variations for each annotator. Therefore area measure evaluates the overlapping areas between the segmentation and ground truth while also providing a tolerance for slight width differences.

$$\text{Area} = \frac{\#((\delta_\alpha(S) \cap S_G) \cup (S \cap \delta_\alpha(S_G)))}{\#(S \cup S_G)} \quad (3.7)$$

δ_α is the morphological dilation operation with a disc of α radius. The value of α determined the level of tolerance to the width variations.

Length factor measures the level of similarity in terms of length.

$$\text{Length} = \frac{\#((\varphi(S) \cap \delta_\beta(S_G)) \cup (\delta_\beta(S) \cap \varphi(S_G)))}{\#(\varphi(S) \cup \varphi(S_G))} \quad (3.8)$$

where φ is the skeletonization operation and δ_β is the morphological dilation operation with a disc of β radius.

Cohen's kappa coefficient (κ) is the measure of agreement between two observers (Ben-David, 2008). If segmentation algorithm is accepted as the observer number one and the ground truth is accepted as observer number two, kappa will define the success of the algorithm while also taking into account the of agreement between two classes occurring by chance.

$$\kappa = \frac{\text{Accuracy} - p_e}{1 - p_e} \quad (3.9)$$

where p_e is the probability of agreement occurring by chance.

CHAPTER FOUR

VESSEL EXTRACTION AND EXTRICATION FOR LIVER ANALYSIS

DATASET – VEELA

4.1 VEELA

Our datasets were obtained randomly from Dokuz Eylül University Radiology Department's Picture Archiving and Communication System (PACS). 35 Abdominal CTA volumes consist of 12-bit DICOM images with 512 x 512 resolution. Slice thickness of datasets vary between 2 and 3.2 mm, with 90 slices per volume on average. Pixel spacing are around 0.68 x 0.68 mm.

Scans belong to 22 female and 13 male potential liver donors. The patient ages ranged from 18 to 57, with an average of 37 years.

4.2 Extraction of Liver Vascular Trees

For the manual segmentation we focused on hepatic and portal vessels. Annotations were done for 35 CTA volumes for each individual slice with an open source program 3D Slicer (Fedorov, Beichel, & Kalpathy-Cramer, 2012). Datasets were loaded to the program with the DICOM module, noise was filtered with curvature anisotropic filtering. By adjusting window level and width, vessel were made more visible. Figure 4.1 shows an example labelling process.

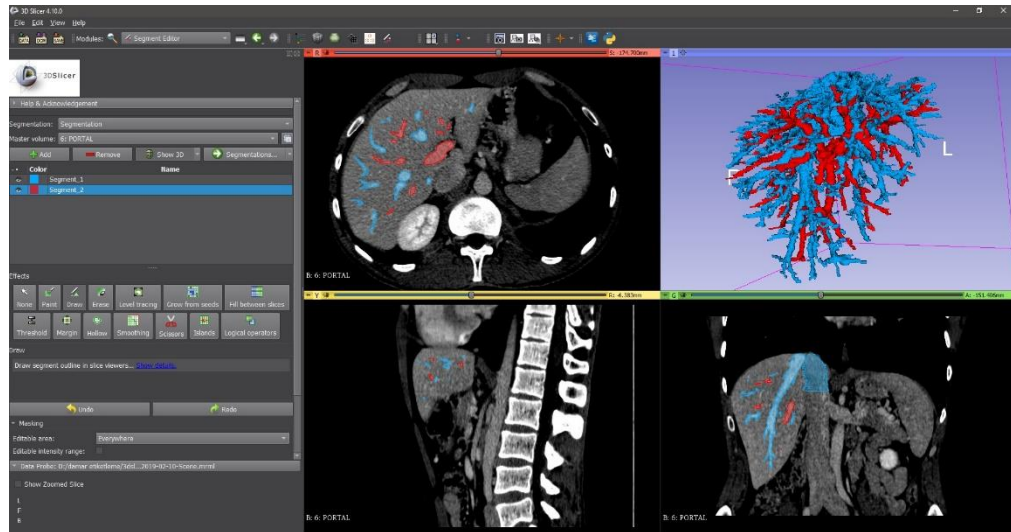


Figure 4.1 An example labeling process in 3D slicer

Hepatic vessels were labeled with blue and portal vessel were labeled with red color. Some of the labeled slices are given in Figure 4.2.

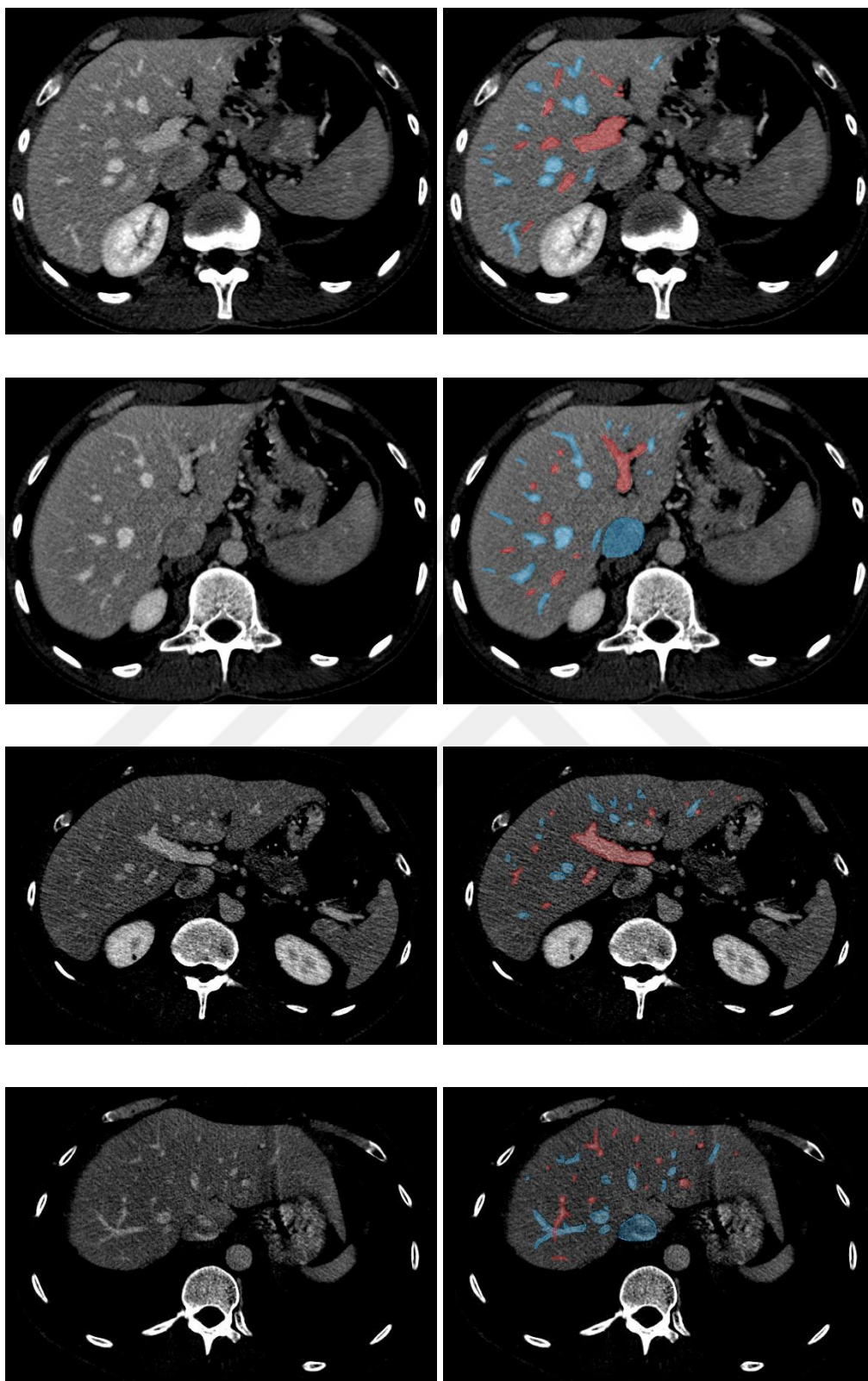


Figure 4.2 Labeled example slices

The program generated 3D models of the labeled vessels and examining these models allows us to control the labels for continuity. Examples of 3D models for portal and hepatic veins separately are given in Figure 4.3. Figure 4.4 demonstrates some of the computer generated example models.

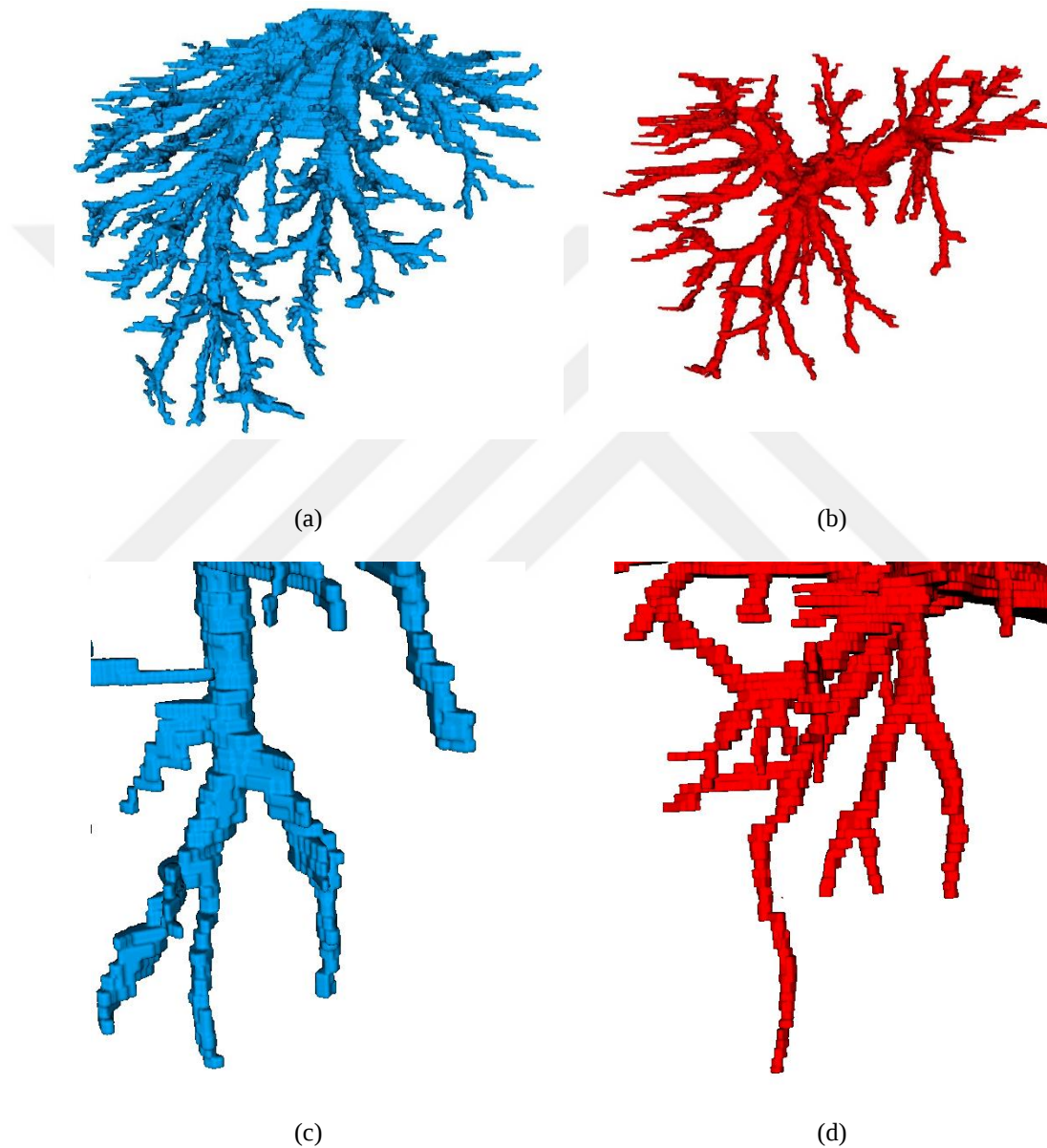
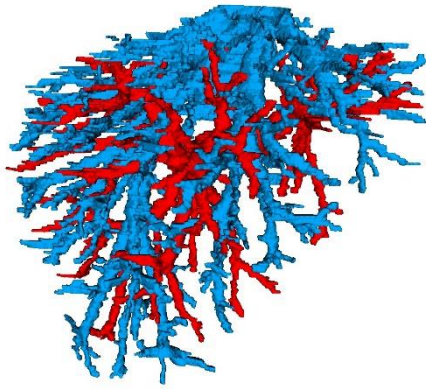
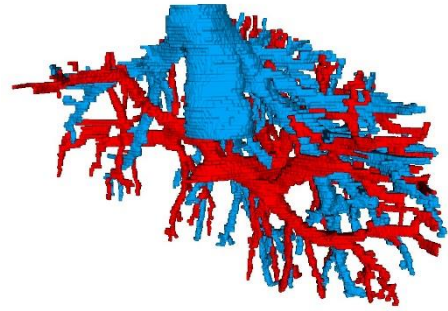


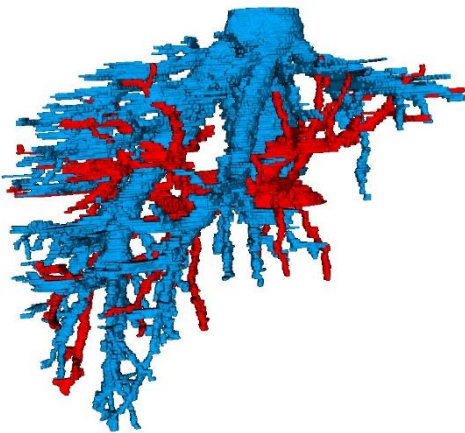
Figure 4.3 Example 3D models for separate vein systems for (a) a hepatic vein system, (b) a portal vein system, (c) close up of a hepatic vein, (d) close up of a portal vein



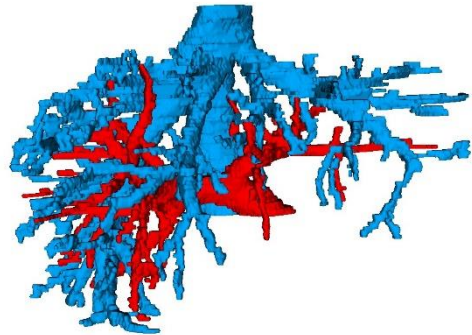
(a)



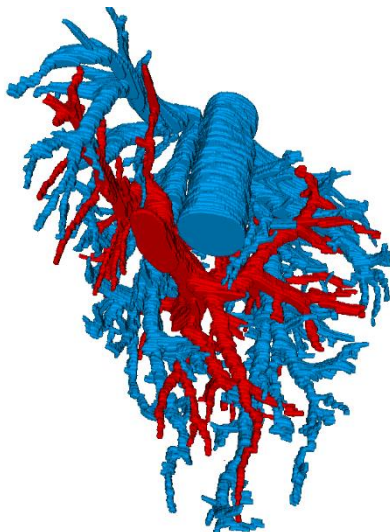
(b)



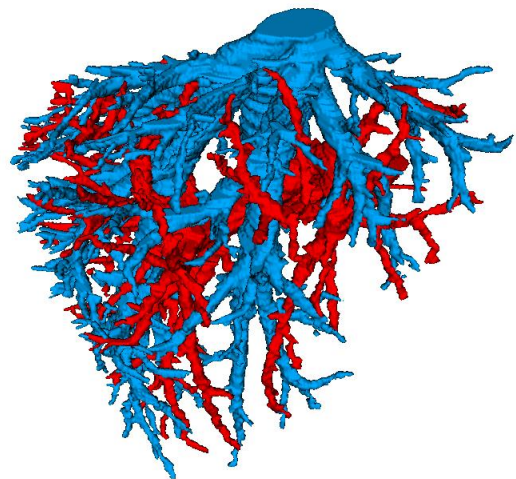
(c)



(d)



(e)



(f)

Figure 4.4 Example 3D models

All manual segmentations were controlled by an expert radiologist in axial, coronal and sagittal planes.

4.3 Challenges encountered during labeling

Manual segmentation process requires a lot of attention, it is very tedious and time consuming. Labeling of each dataset takes around 8-10 hours together with required additional time during which the labels are being controlled by another expert.

Main issues encountered during segmentation are caused due to the unavoidable reasons affecting the imaging quality. Due to the routine clinical standards of acquisition, some visual quality degradations occur because of the high slice thickness. For some datasets, the time elapsed between the injection of the contrast agent and beginning of the scan was not correctly adjusted. This situation led to the inability to achieve the desired intensity differences between the vessel tissue and the liver parenchyma. Imaging quality loss not only prevents vessel border to be clearly identified, but also some visible vessel structures may appear different than their actual size and shape. This makes the images unsuitable as the extracted anatomical structure of the vessels will be erroneous.

There are also some CTA artifacts that may cause incorrect labeling. Beam hardening is an artifact that occurs when the beam passing through the tissues with different density causes dark density lines (Boas & Fleischmann, 2012). It may cause distortions in the image, causing the vessel shape to be misinterpreted.

Transient hepatic attenuations (Chen & Chen, 1999) cause some parts of the parenchyma to be brighter than the remaining tissue and bright areas may be misinterpreted as vessels.

CHAPTER FIVE

APPLICATION OF COMMON VESSEL SEGMENTATION METHODS AND RESULTS

Four vessel segmentation methods were tested for 20 CTA datasets and results were evaluated with respect to the manually annotated ground truth data.

5.1 Used Methods

5.1.1 Hessian Eigenvalues – Frangi Method

Frangi et al. (Frangi & Niessen, 1998) proposed a vessel segmentation method based on eigenvalue decomposition of the Hessian matrix that detects tubular structures. Eigen values are calculated for each voxel in the volume and are sorted according to $|\lambda_3| > |\lambda_2| > |\lambda_1|$.

For a voxel to have a high *vesselness* value:

- (i) $|\lambda_1|$ value needs to be as small as possible, ideally zero,
- (ii) λ_3 and λ_2 values need to be high in magnitude and of the same polarity.

Polarity of eigenvalues are determined according to the color of the vessels. For bright vessel structures on dark background λ_3 and λ_2 values need to be negative, for dark vessel structures on light background λ_3 and λ_2 values need to be positive. Vessels on CTA volumes are known to be bright.

Each eigenvalue indicates the amount of change in its direction. Since vessels are tubular structures, the eigenvalue that follows the vessel's direction is expected to be the smallest. At the same time other two eigenvalues are expected to be high at vessel borders due to the intensity difference. Hence, Frangi's vesselness criteria can be summarized as given in (5.1):

$$|\lambda_1| \approx 0, |\lambda_1| < |\lambda_2|, \lambda_2 \approx \lambda_3 \quad (5.1)$$

Three different measures are constructed using obtained eigenvalues. First measure $\mathcal{R}_B$ calculates the deviation from the blob-like structure, however is not able to differentiate line-like and plate-like structures.

$$\mathcal{R}_B = \frac{|\lambda_1|}{\sqrt{|\lambda_2\lambda_3|}} \quad (5.2)$$

Second measure $\mathcal{R}_A$ is the ratio of two highest eigenvalues, and distinguishes between plate-like and line-like structures.

$$\mathcal{R}_A = \frac{|\lambda_2|}{|\lambda_3|} \quad (5.3)$$

First two measures make use of the geometric structure of the vessels. However it is known that in CTA images vessel are brighter than parenchyma. Third measure is calculated with the norm of Hessian matrix.

$$S = \sqrt{\sum_{j \leq D} \lambda_j^2} \quad (5.4)$$

where D is the number of eigenvalues.

This value is low for the background voxels and larger for the voxels that belong to the vessel, since at least one of the eigenvalues is going to be large.

Finally, vesselness function can be written as:

$$\mathcal{V}_o(s) = \begin{cases} 0 & \text{if } \lambda_2 > 0 \text{ or } \lambda_3 > 0, \\ \left(1 - \exp\left(-\frac{\mathcal{R}_A^2}{2\alpha^2}\right)\right) \exp\left(-\frac{\mathcal{R}_B^2}{2\beta^2}\right) \left(1 - \exp\left(-\frac{S^2}{2c^2}\right)\right) & \text{otherwise} \end{cases} \quad (5.5)$$

where α, β and c are tuning parameters and need to be adjusted for datasets.

5.1.2 Hessian Eigenvalues – Jerman Method

Jerman et al. (Jerman, Pernus, Likar, & Spiclin, 2016) proposed another method based on eigenvalue decomposition of the Hessian matrix. Eigenvalues are calculated similar to the Frangi et al. method. First, a volumetric ratio that is given in (5.6) is considered in order to determine spherical structures.

$$VR = |\lambda_1 \lambda_2 \lambda_3| \left[\frac{3}{|\lambda_1| + |\lambda_2| + |\lambda_3|} \right]^3 \quad (5.6)$$

Herein by including the spherical structures, bifurcations, lesions and tumors are also considered. In order to enhance elongates structures, λ_1 is substituted with $\lambda_2 - \lambda_1$ since $\lambda_1 \ll \lambda_2$. Rounded structures where $\lambda_1 \approx \lambda_2 \approx \lambda_3$ are repressed since $\lambda_2 - \lambda_1$ will be equal to zeros. Hence λ_1 is removed from the equation.

In order to enhance the vesselness value for eigenvalues with lower magnitudes, λ_3 are regularized with the scale s :

$$\lambda_\rho(s) = \begin{cases} \lambda_3 & \text{if } \lambda_3 > \tau \max_x \lambda_3(x, s), \\ \tau \max_x \lambda_3(x, s) & \text{if } 0 < \lambda_3 \leq \max_x \lambda_3(x, s), \\ 0 & \text{otherwise} \end{cases} \quad (5.7)$$

where τ is a regularization parameter between zero and one. Lower τ values lead to more intense outputs.

In order to be able to enhance the structures with elliptic cross-sections, λ_ρ is substituted with $\lambda_\rho - \lambda_2$ and the response for structures with $\lambda_2 \geq \lambda_\rho/2$ is set to one. Finally Jerman's vesselness function can be summarized as given in (5.8).

$$\mathcal{V}_\rho = \begin{cases} 0 & \text{if } \lambda_2 \leq 0 \vee \lambda_\rho \leq 0, \\ 1 & \text{if } \lambda_2 \geq \lambda_\rho/2 > 0, \\ \lambda_2^2 (\lambda_\rho - \lambda_2) \left[\frac{3}{\lambda_2 + \lambda_\rho} \right]^3 & \text{otherwise} \end{cases} \quad (5.8)$$

5.1.3 Multilevel Thresholding – Otsu Method

Otsu method (Otsu, 1979) allows for an automatic threshold selection for segmentation of gray level images. Method searches for a threshold value that minimizes inter-class variability for the segmented image.

An image that was segmented with the k value consists of background and foreground classes, which are represented with C_0 and C_1 respectively. Assuming that

image has L gray level values, C_0 includes pixels with $[1, \dots, k]$ levels and C_1 includes pixels with levels $[k + 1, \dots, L]$. The probabilities of class occurrences can be written as $\omega_0 = \sum_{i=1}^k p_i$ and $\omega_1 = \sum_{i=k+1}^L p_i$, where p_i is the probability of each gray level. The mean class levels are μ_0 and μ_1 . Class variances are σ_0 and σ_1 . The validity of the threshold level is determined by the measures given in (4.9).

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \quad \kappa = \frac{\sigma_T^2}{\sigma_W^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2} \quad (5.9)$$

where σ_T is the variance of the whole picture and rest of the variables are given in (5.10).

$$\sigma_W^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2, \quad \sigma_B^2 = \omega_0 \omega_1 (\mu_1 - \mu_0)^2 \quad (5.10)$$

A k value that maximizes one of the measures given in (4.9) can be chosen as threshold. Multi-level thresholding is obtained by choosing more k values. For instance, image is segmented into the three classes C_0, C_1 and C_2 with two threshold values $0 < k_1 < k_2 < L$.

5.1.3 K-Means Clustering

K-means clustering method (Dhanachandra, Manglem, & Chanu, 2015) divides an image into K clusters. First K numbers of cluster centers are selected randomly, all points in the dataset are appointed into clusters according to their Euclidean distance to the cluster centers. After assigning all the points new cluster centers are calculated and process is repeated until change in points' positions is less than pre-determined threshold value.

5.2 Results

Methods mentioned in previous section were tested with 20 CTA datasets and results were compared with the manually annotated ground truth data. Visualization of segmentation results for one dataset are given in Figure 4.1. Performance evaluation results are presented value in Table 4.1 in terms of mean DSC and standard deviation.

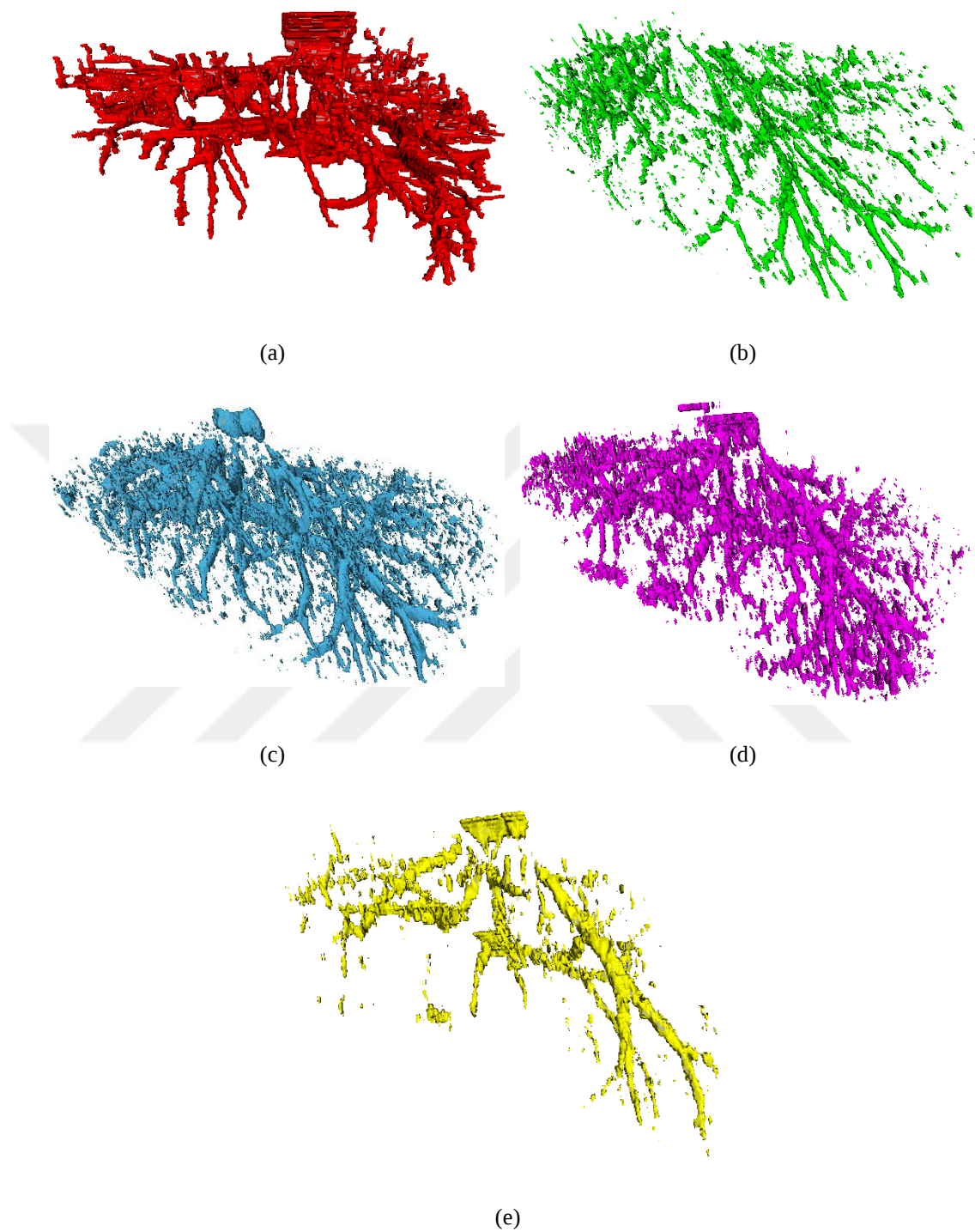


Figure 5.1 Results of the segmentation algorithms. (a) Ground truth, (b) Frangi, (c) Jerman, (d) multilevel Otsu, (e) K-means clustering

Table 5.1 Mean DSC results and standard deviations for vessel segmentation methods

Method	Frangi	German	Otsu	K-means
Mean DSC	0.28	0.5	0.25	0.31
Standard deviation	0.07	0.04	0.15	0.19

All methods were unsuccessful in segmenting inferior vena kava, which has a lower intensity value compared to the other vessels. Furthermore, all of the methods failed in establishing continuous shapes.



CHAPTER SIX

PROPOSED DEEP LEARNING STRATEGY

6.1 Training DeepMedic

DeepMedic (Kamnitsas, et al., 2017) is a 3D convolutional neural network (CNN) that was developed brain tumor segmentation in multimodal MRI scans for BRATS2015 challenge. CNN architecture is coupled with a 3D conditional random field (CRF).

DeepMedic is a dual path, 11 layers deep 3D convolutional neural network. Medical scans are scanned with two different scales simultaneously, in order to include both local and larger contextual information. During post-processing 3D conditional random fields are used to decrease false positives. The convolutional model is given in Figure 6.1.

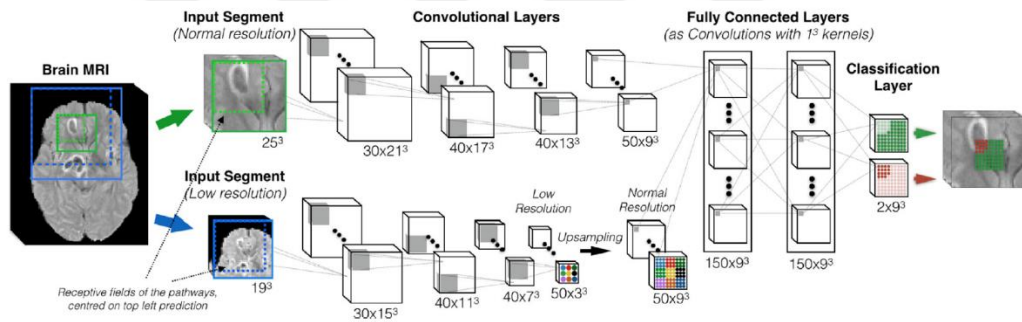


Figure 6.1 Model of DeepMedic The kernel sizes of both pathways are given as 5^3 for the purpose of reducing figure's size and actual kernels are of 3^3 size (Kamnitsas et al., 2017)

We trained used liver vasculature CTA datasets, with manual annotations as ground truth data. In addition, liver segmentations that were generated for previous studies were used as region of interest (ROI) maps in order to decrease the computational workload. Datasets were also reconstructed for coronal and sagittal planes and training process was repeated. Two PCs with 4 gb Nvidia 960M gpu and and 8 gb Nvidia Quadro 4000 gpu. Trainings were done in 35 epochs each with 20 subepochs.

10 of the 35 total datasets were used as for training. Training process took around 32 hours for axial plane and 20 hours for coronal and sagittal planes. Training process

in terms of accuracy, sensitivity, precision (positive predictive value) and DSC for training in with three orthogonal planes in Figure 6.2.

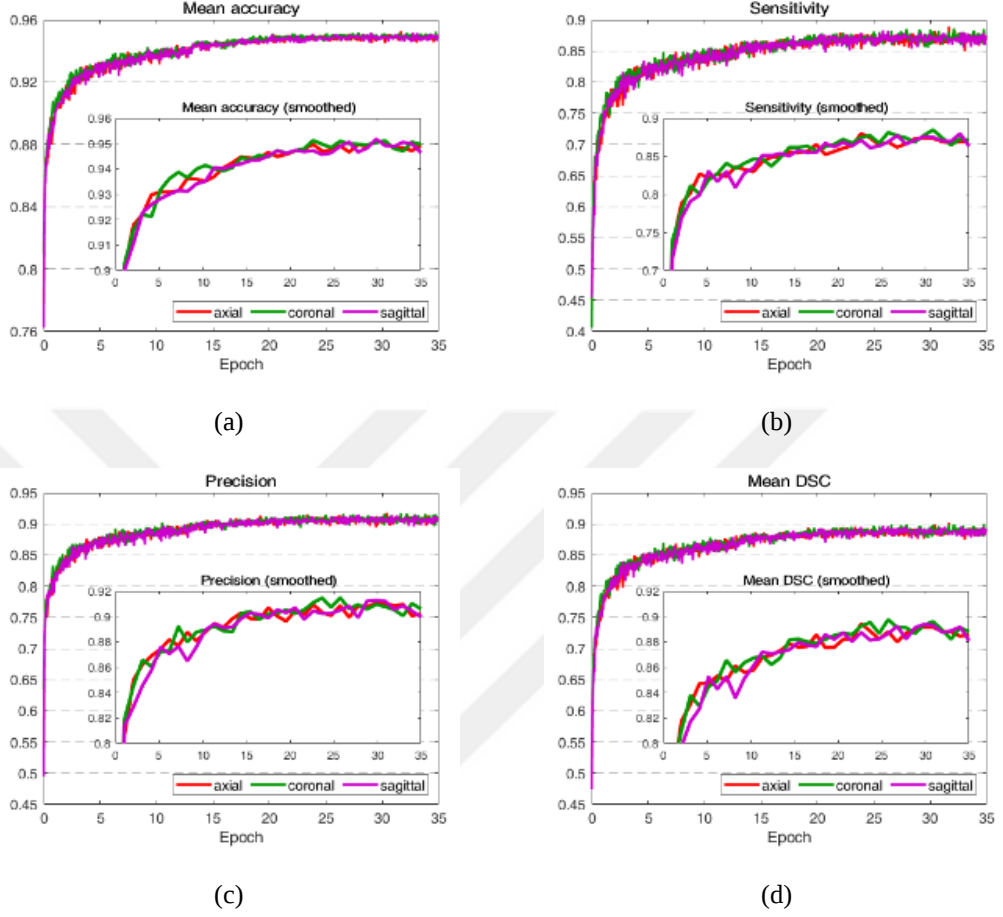


Figure 6.2 Training process for datasets in three orthogonal planes: (a) Mean accuracy, (b) sensitivity, (c) precision and (d) mean dice values over 35 epochs. Smaller figures only show mean results for each epochs and ignores individual subepoch results

Boxplots for training and test sets' final DSC values for axial, coronal and sagittal planes are given in Figure 6.3.

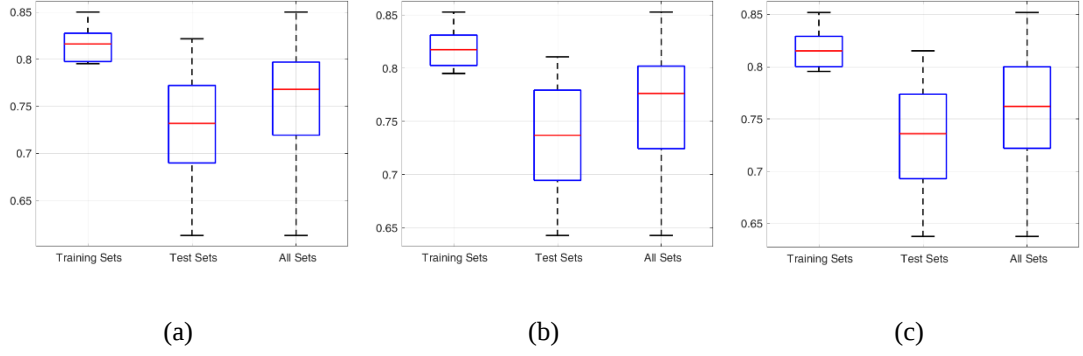


Figure 6.3 Final DSC values for training sets, test sets and all sets in (a) axial, (b) coronal and (c) sagittal planes

Since ground truth images were manually labeled, vessels that were found by the algorithm and vessels that were manually labeled may have varying vessel borders. In fact, same annotator may label same vessels with slightly different borders each time the labelling process is repeated. The difference images obtained by subtracting ground truth image and the result are given in Figure 6.4. Thin lines show the disagreement in vessel border. These borders are not necessarily indicator of false segmentation, however they cause drop in DSC result. Therefore, the area metric that was described in Evaluation Metrics section was applied to all datasets with $\alpha = 1$ and $\alpha = 2$. Boxplot of comparison between DSC and area values for all datasets is given in Figure 6.5.



Figure 6.4 Difference between ground truth and results. Thin lines in images indicate disagreement in vessel borders between algorithm results and ground truth labels

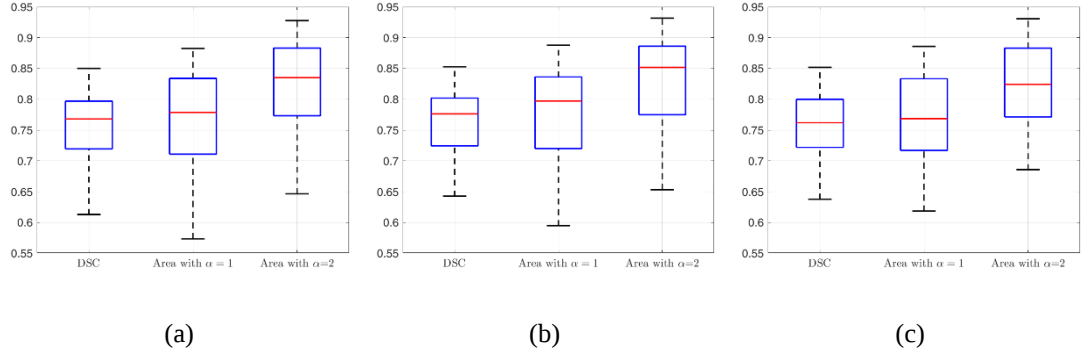


Figure 6.5 Comparison between DSC value and Area values with $\alpha = 1$ and $\alpha = 2$. (a) Results on axial plane, (b) coronal plane and (c) sagittal plane

Training with different planes, helps the network to learn different parts of the vessels. Since vessel structure is more visible in coronal plane, network training was shorter in comparison and mean dice and area values were higher. Inhomogeneity in intensity difference may lead to false positives; however reconstructing the volume may help the training process. Looking at Figure 6.6 where three example results are shown, it can be seen that mistakes that were done by one training process were not repeated by the others. Hence, false positive rate can be lowered by taking the intersection volume of three results.

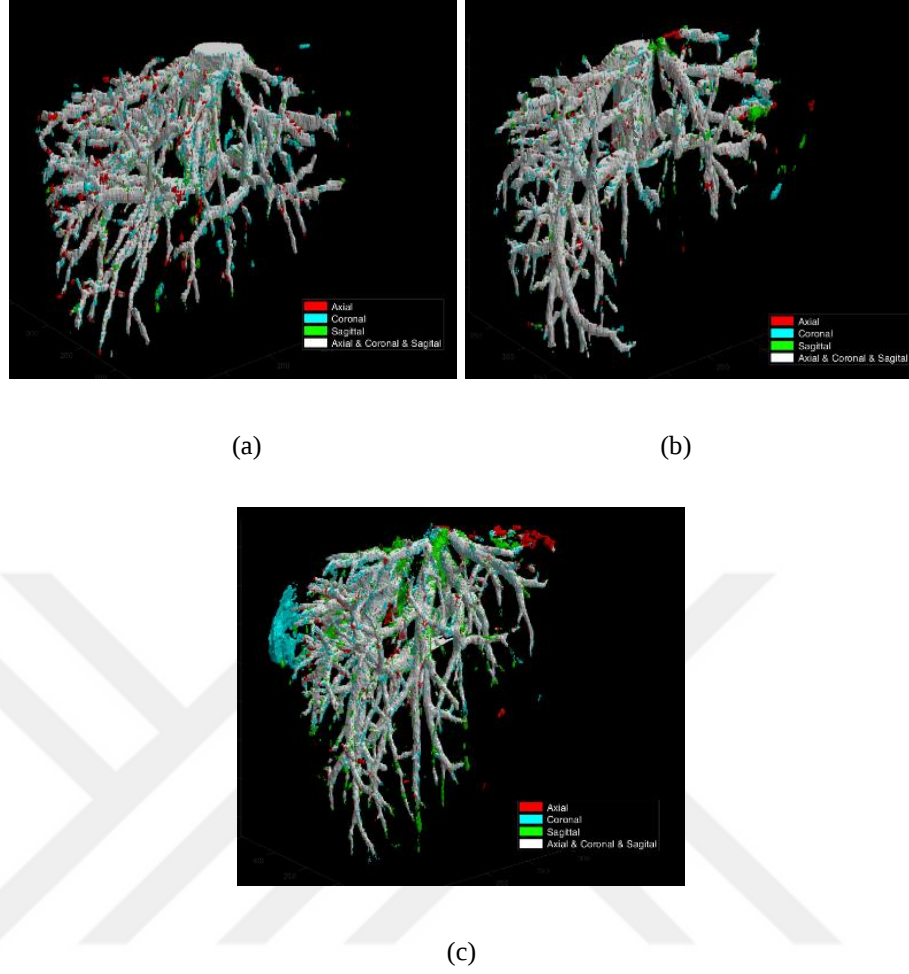


Figure 6.6 Example results of training process for axial, coronal and sagittal planes, and volumes that were obtained from the intersection of these three results

Figure 6.7 shows the DSC results for every dataset in axial, coronal and sagittal planes, and DSC results of the volumes that were obtained by taking the intersections and unions of these results. Additionally DSC results where FPs were removed from the datasets are given in Figure 6.8.

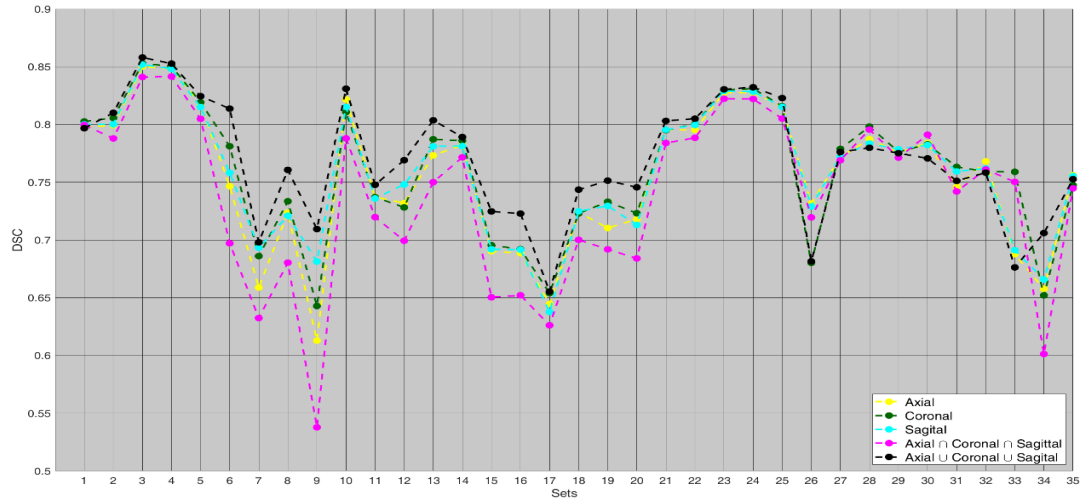


Figure 6.7 DSC results of all training that were performed in axial, coronal and sagittal planes, and DSC results of the volumes that were obtained with intersections and unions of these results

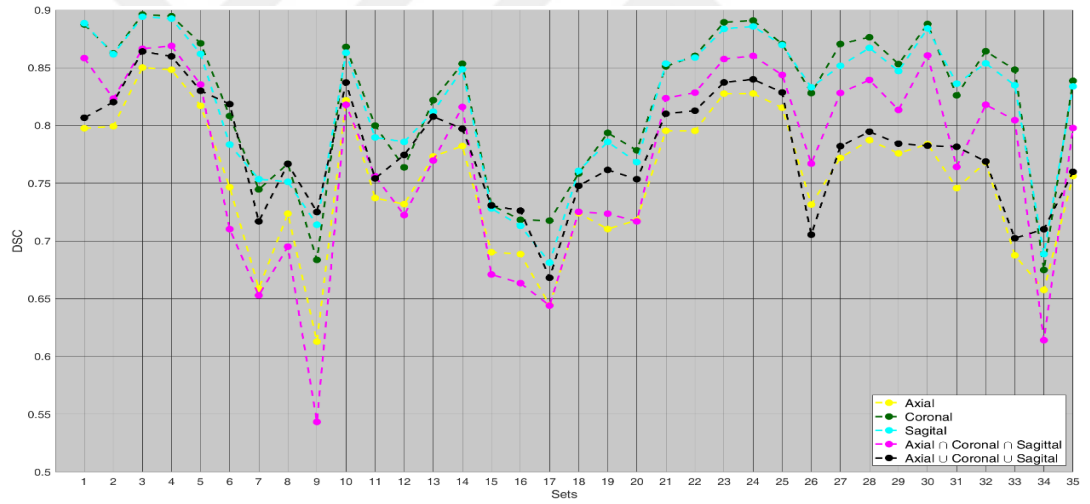


Figure 6.8 DSC results of axial, coronal, sagittal planes, intersection volumes and union volumes where FPs are removed

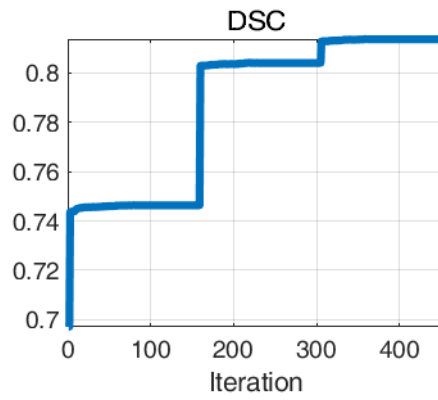
6.2 Combining DeepMedic Results

Developed approach that makes use of the continuity feature of the vascular trees for combining deep medic results for three orthogonal planes and increasing dice coefficient is summarized below:

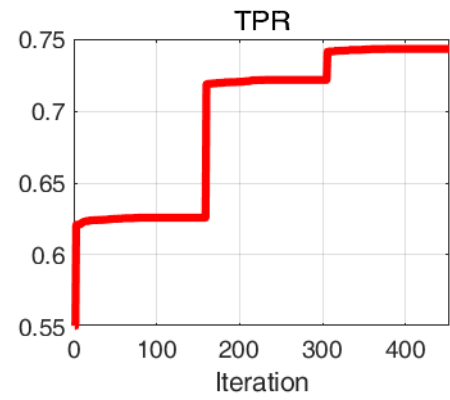
1. A main frame was constructed by taking the intersection of training results in axial, coronal and sagittal planes.
2. Then remaining volumes were obtained by subtracting the main frame from all of the results in axial, coronal and sagittal planes.
3. Connected component analysis was performed for the main frame and for the remaining volumes.
4. Each component of the remaining frames was combined with the main frame if the number of connected components in resulting volume was equal or lower than the previous state of the main frame.
5. The combined parts were removed from the results and the process was repeated until number of components were not decreasing anymore.

Change in DSC, false positive rate, true positive rate and true negative rate during the process for an example dataset is given in Figure 6.9, change of mean values for all sets is given in Figure 6.10.

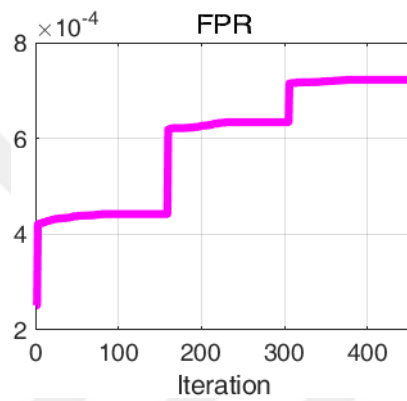
Increase in Dice coefficient and true positive and decrease in false negative rate are very desirable effects. And while the increase in false positive rate presents itself as an issue, range of FPR is very low (10^{-4}).



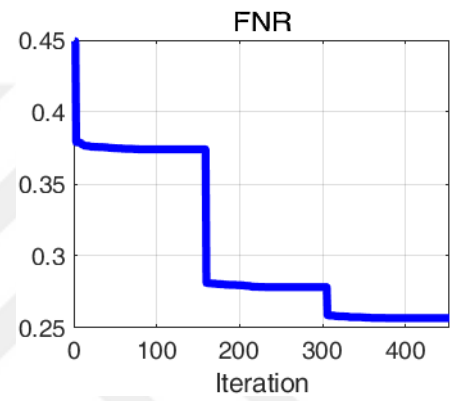
(a)



(b)

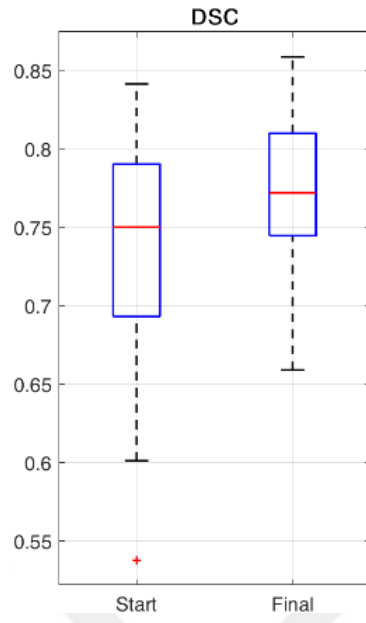


(c)

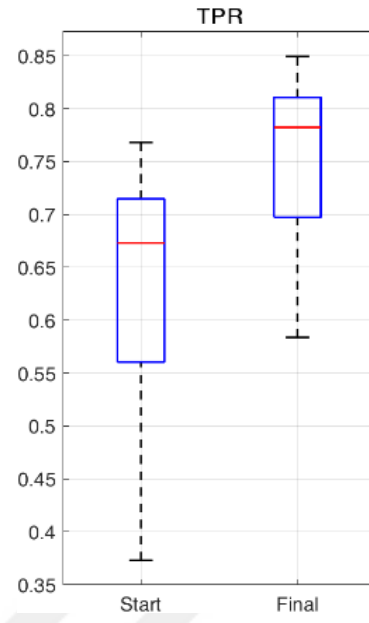


(d)

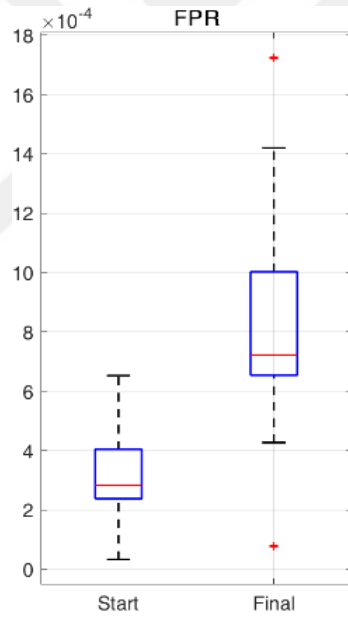
Figure 6.9 Change in parameters in terms of (a) Dice coefficient, (b) true positive rate, (c) false positive rate and (d) false negative rate



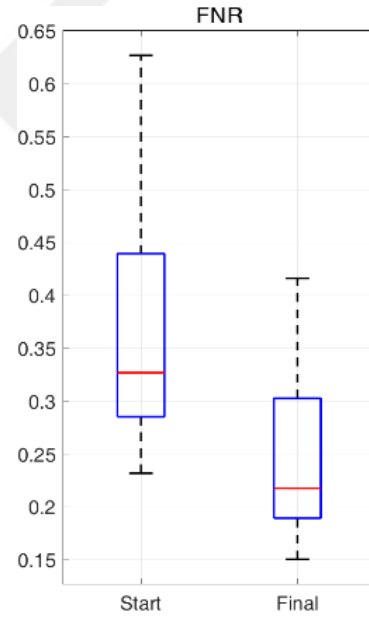
(a)



(b)



(c)



(d)

Figure 6.10 Change in parameters in terms of (a) Dice coefficient, (b) true positive rate, (c) false positive rate and (d) false negative rate for all datasets

CHAPTER SEVEN

CONCLUSION

In this thesis, a liver vascular tree ground truth dataset was generated for 35 abdominal CTA scans. For the best of our knowledge, this is the first dataset for liver vasculature.

Liver vasculatures are very complex structures, therefore development of an accurate segmentation method is a challenging task. However deep learning methods show promising results. A deep learning based method DeepMedic was trained from scratch with the dataset that was sliced in axial, coronal and sagittal planes. Each training allowed the network to learn different aspects of the vasculature. A connected component based method was used to combine all these results. Mean performance was increased in terms of Dice coefficient, true positive rate and true negative rate. However false positive rate was increased, therefore an additional method for decreasing FPR needs to be developed.

A region growing was applied on ground truth images where obtained vessel segmentation results were used as seed points. Results had over 90% DSC scores. This tells us that the vast majority of vessel branches were segmented correctly, however vessels in proximal regions or vessel that have unexpected intensity values or vessel whose appearance were obstructed by the presence of noise cause a decrease in overall DSC scores. Hence, it can be said that current method shows very promising results and, very high performance results may be obtained with some additional steps.

REFERENCES

- Abdel-Misih, S.R., & Bloomston, M. (2010). Liver anatomy. *Surgical Clinics*, 90 (4), 643-653.
- Balogh, J., Victor, D., Asham, E., Burroughs, S., Boktour, M., Saharia, A., et al. (2016). Hepatocellular carcinoma: a review. *Journal of Hepatocellular Carcinoma*, 3, 41-53.
- Bándi, P., Geessink, O., & Manson, Q. (2019). From detection of individual metastases to classification of lymph node status at the patient level: The CAMELYON17 challenge. *IEEE Transactions On Medical Imaging*, 38 (2), 550-560.
- Barrett, J. F., & Keat, N. (2004). Artifacts in CT: recognition and avoidance. *Radiographics*, 24 (6), 1679-1691.
- Ben-David, A. (2008). About the relationship between ROC curves and Cohen's kappa. *Engineering Applications of Artificial Intelligence*, 21, 874-882.
- Bernard, O., Lalande, A., & Zotti, C. (2018). Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Transactions On Medical Imaging*, 37 (11), 2514-2525.
- Boas, F., & Fleischmann, D. (2012). CT artifacts: Causes and reduction techniques. *Imaging in Medicine*, 4 (2), 229-240.
- Bogunović, H., Venhuizen, F., & Klimescha, S. (2019). RETOUCH - The retinal OCT fluid detection and segmentation benchmark and challenge. *IEEE Transactions On Medical Imaging*.
- Chen, W.P., & Chen, J.H. (1999). Spectrum of transient hepatic attenuation differences in biphasic helical CT. *American Journal of Roentgenology*, 172 (2), 419-424.
- Conversano, F., Franchini, R., Demitri, C., Massoptier, L., Montagna, F., Maffezzoli, A., et al. (2011). Hepatic vessel segmentation for 3D planning of liver surgery:

- Experimental evaluation of a new fully automatic algorithm. *Academic Radiology*, 18 (4), 461-470.
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *MICCAI 2016 Medical Image Computing and Computer-Assisted Intervention*, 424-432.
- Dhanachandra, N., Manglem, K., & Chanu, Y. (2015). Image segmentation using k-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, 54, 764-771.
- Esneault, S., Lafon, C., & Dillenseger, J.L. (2010). Liver vessels segmentation using a hybrid geometrical moments / graph cuts method. *IEEE Transactions On Biomedical Engineering*, 57(2), 276-283.
- Fasel, J., Selle, D., Evertsz, C. J., Terrier, F., Peitgen, H., & Gailloud, P. (1988). Segmental anatomy of the liver: poor correlation with CT. *Radiology*, 206 (1), 151-156.
- Fedorov, A., Beichel, R., & Kalpathy-Cramer, J. (2012). 3D Slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging*, 30, 1323–1341.
- Frangi, A., & Niessen, W. (1998). Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention*, 130-137.
- Gegúndez-Arias, M., Aquino, A., Bravo, J., & Marín, D. (2012). A function for quality evaluation of retinal vessel segmentations. *IEEE Transactions On Medical Imaging*, 31 (2), 231-239.
- Haft-Javaherian, M., Fang, L., Muse, V., Schaffer, C. B., Nishimura, N., & Sabuncu, M. R. (2019). Deep convolutional neural networks for segmenting 3D in vivo multiphoton images of vasculature in alzheimer disease mouse models. *Plos One*, 14 (3).

- Jerman, T., Pernus, F., Likar, B., & Spiclin, Z. (2016). Enhancement of vascular structures in 3D and 2D angiographic images. *IEEE Transactions Medical Imaging*, 35 (9), 2107-2118.
- Jiang, H., He, B., Fang, D., Ma, Z., Yang, B., & Fang, L. (2013). A region growing vessel segmentation algorithm based on spectrum information. *Computational and Mathematical Methods in Medicine*.
- Jimenez-del-Toro, O., Müller, H., & Krenn, M. (2016). Cloud-based evaluation of anatomical structure segmentation and landmark detection algorithms: VISCERAL anatomy benchmarks. *IEEE Transactions On Medical Imaging*, 35 (11), 2459-2475.
- Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., et al. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, 36, 61-78.
- Kuijf, H., Biesbroek, J., & de Bresser, J. (2019). Standardized assessment of automatic segmentation of white matter hyperintensities; results of the WMH segmentation challenge. *IEEE Transactions on Medical Imaging*.
- Lehmann, K. S., Ritz, J.P., Valdeig, S., Schenk, A., Holmer, C., Peitgen, et al. (2008). Portal vein segmentation of a 3D-planning system for liver surgery—In vivo evaluation in a porcine model. *Annals of Surgical Oncology*, 15 (7), 1899–1907.
- Lesage, D, Angelini, E.D., Bloch, I., & Funka-Lea, Gareth. (2009). A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. *Medical Image Analysis*, 13, 819-845.
- Lorigo, L., Faugeras, O., Grimson, W., Keriven, R., Kikinis, R., Nabavi, A., et al. (2001). CURVES: Curve evolution for vessel segmentation. *Medical Image Analysis*, 5, 195-206.
- Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., et al. (2015). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions On Medical Imaging*, 34 (10), 1993-2024.

- Moccia, S., De Momi, E., El Hadji, S., & Mattos, L. (2018). Blood vessel segmentation algorithms — Review of methods, datasets and evaluation metrics. *Computer Methods and Programs in Biomedicine*, 158, 71-91.
- Murphy, D. J., Aghayev, A., & Steigner, M. L. (2018). Vascular CT and MRI: a practical guide to imaging protocols. *Insights into Imaging*, 9 (2), 215–236.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9 (1), 62-66.
- Reitinger, B., Bornik, A., Beichel, R., & Schmalstieg, D. (2006). Liver surgery planning using virtual reality. *IEEE Computer Graphics and Applications*, 26 (6), 36-47.
- Schwartzbaum, J., L Fisher, J., Aldape, K., & Wrensch, M. (2006). Epidemiology and molecular pathology of glioma. *Nature Clinical Practice Neurology*, 494-503.
- Selle, D., Preim, B., Schenk, A., & Peitgen, H.O. (2002). Analysis of vasculature for liver surgical planning. *IEEE Transactions On Medical Imaging*, 21 (11), 1344-1357.
- Selver, M. A., Kocaoğlu, A., Demir, G.K., Doğan, H., Dicle, O., & Güzeliş, C. (2008). Patient oriented and robust automatic liver segmentation for pre-evaluation of liver transplantation. *Computers in Biology and Medicine*, 38 (7), 765-784.
- Wong, H., Desser, T. S., & Jeffrey, R. B. (2018). Transient hepatic attenuation differences in computed tomography from extrahepatic portal vein compression. *Radiology Case Reports*, 3 (1), 113.