

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

AUTOMATED ESTIMATION OF THE PROSTATE VOLUME
FROM ABDOMINAL ULTRASOUND IMAGES

NUR BANU ALBAYRAK
A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
DEPARTMENT OF COMPUTER ENGINEERING

GEBZE
2022

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

**AUTOMATED ESTIMATION OF THE
PROSTATE VOLUME FROM ABDOMINAL
ULTRASOUND IMAGES**

NUR BANU ALBAYRAK
**A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**
DEPARTMENT OF COMPUTER ENGINEERING

THESIS SUPERVISOR
PROF. DR. YUSUF SINAN AKGUL

GEBZE

2022

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ABDOMİNAL ULTRASON
GÖRÜNTÜLERİNDEN PROSTAT
BÜYÜKLÜĞÜNÜN OTOMATİK ÖLÇÜMÜ

NUR BANU ALBAYRAK
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

DANIŞMANI
PROF. DR. YUSUF SINAN AKGUL

GEBZE
2022



DOKTORA JÜRİ ONAY FORMU

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 03/02/2022 tarih ve 2022/06 sayılı kararıyla oluşturulan jüri tarafından 02/03/2022 tarihinde tez savunma sınavı yapılan Nur Banu Albayrak'ın tez çalışması Bilgisayar Mühendisliği Anabilim Dalında DOKTORA tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI)

: Prof. Dr. Yusuf Sinan Akgül

ÜYE

: Dr. Öğr. Üyesi Yakup Genç

ÜYE

: Doç. Dr. Ayşe Betül Oktay

ÜYE

: Prof. Dr. Songül Varlı

ÜYE

: Doç. Dr. Habil Kalkan

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun

...../...../..... tarih ve/..... sayılı kararı.

İMZA/MÜHÜR

SUMMARY

Prostate cancer (PCa) is one of the most frequently seen cancer types in Turkey and the World among males. Ultrasound (US) is a commonly preferred imaging technology in PCa diagnosis and treatment. There are two types of US methods for prostate imaging. Transrectal ultrasound (TRUS) provides better images for prostate size inference, but the procedure is very discomforting for the patients. The second type is the abdominal ultrasound (AUS) technique which is much more practical.

There is only a limited image analysis method in the literature on prostate volume estimation from AUS images. Existing work for AUS prostate segmentation requires initialization by experts, and it was not verified with an extensive image set.

In this study, novel methods were developed that use AUS images to automatically locate the prostate region and estimate the prostate volume. AUS and MR images were collected and manually marked by experts, to produce the ground truth data. The obtained measurement results were analyzed using scientific techniques to investigate the applicability of the proposed method for practical clinical work.

Due to the AUS-related problems, classical object localization and segmentation methods would not produce acceptable solutions for the volume estimation problem on AUS images. Hence, original localization and volume estimation methods needed to be developed. The methods proposed in this study can be explained under three headings. The main idea of the first method is to use the unrelated structures in the AUS images to help locate the prostate. Original techniques were developed for the second method similarly, which aims to estimate the volume of the prostate by contour extraction. In the third method, a novel deep learning architecture was proposed which detects four points on prostate contour which are used in clinical practice. Since the data set is too small for any deep learning training step, our second and third methods were patch-based. Also, using the local and the global information by using such a method was aimed. It was shown in our experiments that this patch-based methodology worked very well with our AUS images.

Key Words: Biomedical image analysis, Prostate, Deformable Part Models, Machine Learning, Active Shape Models, Deep Learning.

ÖZET

Prostat kanseri, Türkiye'de ve Dünya'da erkekler arasında en sık görülen kanser türlerinden biridir. US, prostat kanserinin tanı ve tedavisinde sıklıkla tercih edilen teknolojilerden biridir. Prostat görüntüleme için iki tip US yöntemi vardır. TRUS, prostat boyutu çıkarımı için daha iyi görüntüler sağlar, ancak uygulama prosedürü hastalar için çok rahatsız edicidir. İkinci tip, çok daha pratik olan AUS tekniğidir.

AUS görüntülerinden prostat hacmi tahmini konusunda literatürde sadece bir görüntü analiz yöntemi vardır. AUS prostat segmentasyonu için bahsedilen çalışma, uzmanlar tarafından başlatılmasını gerektirir ve kapsamlı bir görüntü seti ile doğrulanmamıştır.

Bu çalışmada, prostat bölgesini otomatik olarak konumlandırmak ve prostat hacmini tahmin etmek için AUS görüntülerini kullanan yeni yöntemler geliştirilmiştir. Kesin referans verilerini üretmek için AUS ve MR görüntüleri uzmanlar tarafından toplanmış ve manuel olarak işaretlenmiştir. Elde edilen ölçüm sonuçları, önerilen yöntemin pratik klinik çalışma için uygulanabilirliğini araştırmak için bilimsel teknikler kullanılarak analiz edilmiştir.

AUS ile ilgili problemler nedeniyle, klasik nesne konum tespiti ve bölütleme yöntemleri, AUS görüntülerinde hacim tahmini problemi için kabul edilebilir çözümler üretmeyecektir. Bu nedenle özgün konum ve hacim tahmin yöntemlerinin geliştirilmesine ihtiyaç duyulmuştur. Bu çalışmada önerilen yöntemler üç başlık altında açıklanabilir. İlk yöntemin ana fikri, prostatın yerini belirlemeye yardımcı olmak için AUS görüntülerindeki ilgisiz yapıları kullanmaktır. Kontur çıkarımı ile prostat hacmini tahmin etmeyi amaçlayan ikinci yöntem için de benzer şekilde orijinal teknikler geliştirilmiştir. Üçüncü yöntemde, klinik uygulamada kullanılan prostat konturu üzerinde dört noktayı tespit eden yeni bir derin öğrenme mimarisi önerilmiştir. Veri seti herhangi bir derin öğrenme eğitimi adımı için çok küçük olduğundan, ikinci ve üçüncü yöntemlerimiz yama tabanlıdır. Ayrıca böyle bir yöntem kullanılarak yerel ve global bilginin kullanılması hedeflenmiştir. Bu yama tabanlı metodolojinin AUS görüntülerimizle çok iyi çalıştığı deneylerimizde gösterildi.

Anahtar Kelimeler: Biyomedikal görüntü analizi, Prostat, Biçim Değiştirebilen Parça Modelleri, Makine öğrenmesi, Aktif Şekil Modelleri, Derin Öğrenme.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Prof. Dr. Yusuf Sinan Akgül for his invaluable advice, continuous support, and patience during my Ph.D. study. His guidance helped me in all the time of research and writing of this thesis. He will always be a shining example in my academic and personal life.

I also would like to thank my committee members, Dr. Yakup Genç, who always spared time when I asked for his advice and Doç. Dr. Ayşe Betül Oktay, who shared her insightful comments gently.

I would thank to Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) for their financial support and, Dr. Rahmi Çubuk and Dr. Orhun Sinanoğlu for their interdisciplinary support within the scope of the TÜBİTAK project 114E536.

I would like to express my gratitude to my parents, Remziye and Sedat Özalay, who raised me with a passion for learning and taught me to never give up. My biggest luck in life is having them as my parents.

I would like to thank my mother-in-law and father-in-law, Sevinç and Şemsettin Albayrak, for caring for my children during my Ph.D. study and always standing by me.

Last but not least, I would like to thank my husband, Resul Sezgin Albayrak, for his endless patience and encouragement. This Ph.D. study would not have been possible without his support.

TABLE OF CONTENTS

	<u>Page</u>
SUMMARY	v
ÖZET	vi
ACKNOWLEDGMENTS	vii
TABLE of CONTENTS	viii
LIST of ABBREVIATIONS and ACRONYMS	x
LIST of FIGURES	xii
LIST of TABLES	xiv
1. INTRODUCTION	1
1.1. Motivation	1
1.2. Problem Definition	3
1.3. Background	5
1.3.1. AUS Examination	5
1.4. Outline	6
2. LITERATURE REVIEW	7
2.1. Contour and Shape-Based Methods	7
2.2. Region-Based Methods	11
2.3. Classification-Based Methods	12
2.4. Deep Learning-Based Methods	14
2.5. Hybrid Methods	14
3. PROSTATE LOCALIZATION FROM AUS IMAGES BY USING AN SVM BASED GRAPHICAL MODEL	17
3.1. Motivation and Problem Definition	17
3.2. Proposed Method	17
3.2.1. Computation of the Appearance Energy	22
3.2.2. Computation of the Context Energy	26
3.3. Experiments	29
4. PROSTATE SIZE INFERENCE BY PATCH-BASED SEGMENTATION	32
4.1. Motivation and Problem Definition	32
4.2. Proposed Method	32

4.2.1. Segmentation	34
4.2.2. Size Inference	40
4.3. Experiments	40
5. STANDARD ELLIPSOID FORMULA BASED PROSTATE VOLUME ESTIMATION FROM AUS IMAGES BY USING A NOVEL DEEP MODEL	45
5.1. Motivation and Problem Definition	45
5.2. Proposed Method	45
5.2.1. Patch Extraction	46
5.2.2. Quadruplet Network	47
5.2.3. Prostate Volume Inference	51
5.3. Experiments	54
5.3.1. Dataset	54
5.3.2. QDCNN Results	55
5.3.3. Ablation Study	60
6. CONCLUSION AND FUTURE WORK	63
6.1. Conclusion	63
6.2. Future Work	65
REFERENCES	69
BIOGRAPHY	76
APPENDIX	77

LIST OF ABBREVIATIONS AND ACRONYMS

<u>Abbreviations and Acronyms</u>	<u>Explanations</u>
AUS	: Abdominal Ultrasound
BDPM	: Biçim Değiştirebilir Parça Modeli
BPH	: Benign Prostatic Hyperplasia
CT	: Computer Tomography
DCN	: Deformable Convolutional Network
DCNN	: Deep Convolutional Neural Network
DDC	: Discrete Dynamic Contour
DPM	: Deformable Part Model
GAN	: Generative Adversarial Network
GCN	: Graph Convolutional Network
GLOBOCAN	: Global Cancer Observatory
HOG	: Histogram of Oriented Gradients
HSGM	: Halk Sağlığı Genel Merkezi
KSVM	: Kernel Support Vector Machine
LMA	: Levenberg-Marquardt Algorithm
LUTS	: Lowe Urinary Tract Symptoms
MAVD	: Mean Absolute Value Difference
MR	: Magnetic Resonance
PCa	: Prostate Cancer
PCA	: Principal Component Analysis
PSA	: Prostate Specific Antigen
PSAD	: Prostate Specific Antigen Density
RBR	: Radial Bas-Relief
RNN	: Recurrent Neural Network
SDAVD	: Standard Deviation of Absolute Value Difference
SNR	: Signal to Noise Ratio
SVM	: Support Vector Machine
QDCNN	: Quadruplet Deep Convolutional Neural Network
TDCNN	: Twin Deep Convolutional Neural Network

TFCV	:	Ten-Fold Cross Validation
TRUS	:	Trans Rectal Ultrasound
t-SNE	:	t-distributed Stochastic Neighbor Embedding
US	:	Ultrasound

LIST OF FIGURES

<u>Figure No</u>	<u>Page</u>
1.1: Estimated number of new cases in different cancer types for 2020.	1
1.2: Estimated number of deaths in different cancer types for 2020.	2
1.3: Prevalence of different cancer types in Turkey for 2017.	2
1.4: Sample AUS and TRUS images from the transverse plane.	5
1.5: AUS image acquisition from different planes.	6
2.1: Snake energy minimization steps.	8
2.2: Representation of partial active shape model.	10
2.3: Representation of graph cut segmentation.	11
2.4: Feature extraction process for the classification.	12
2.5: Hybrid network structure.	13
2.6: Schema of Gabor-Support Vector Machine.	15
3.1: An early version of the DPM model called pictorial structure.	18
3.2: DPM samples on a car and a person model.	18
3.3: Representation of the proposed system.	19
3.4: Anatomical structures determined for the DPM-based method.	20
3.5: The DPM design for the proposed system.	21
3.6: Creation of HOG feature vectors.	23
3.7: Integral image representation.	24
3.8: Integral image region sum calculation.	25
3.9: Sample score maps for a given AUS image.	25
3.10: The initial and the final locations of the DPM.	27
3.11: DPM optimization on synthetic images.	28
3.12: Sample cluster maps of an AUS image.	28
3.13: Comparison of the SVM-based and DPM-based systems' results.	30
3.14: Sample results of the DPM-based localization method.	31
4.1: The schema of the patch-based segmentation system.	33
4.2: Creation of the training contours using expert markings.	34
4.3: Determination of class intervals for the distance task.	35
4.4: Determination of class intervals for the orientation task.	35
4.5: Determination of class intervals for the curvature task.	36

4.6:	The patch creation process.	37
4.7:	Representation of the training phase.	38
4.8:	Samples of a contour map and a contour.	39
4.9:	Architecture of AlexNet.	40
4.10:	Sample segmentation results for transverse and sagittal images.	41
5.1:	Patch extraction process.	46
5.2:	Locations chosen for the quadruplet patch centers.	47
5.3:	The proposed quadruplet model, QDCNN.	48
5.4:	Quantization of class label values.	49
5.5:	Training process with m transverse images.	50
5.6:	Creation of voting maps.	51
5.7:	Representation of the voting process.	53
5.8:	MAVD values between experts and models.	61
5.9:	t-SNE graph samples.	62
6.1:	Difference between 2D convolution and graph convolution.	67
6.2:	Representation of deformable convolution.	68
6.3:	Representation of the GAN model.	68

LIST OF TABLES

<u>Table No</u>	<u>Page</u>
1.1: Applications where prostate volume and/or boundaries are used.	3
1.2: Advantages and disadvantages of TRUS and AUS images.	4
4.1: Chamfer distance values between experts and the system.	42
4.2: Difference between volume estimations.	43
4.3: Mean values of the difference.	43
4.4: Ratio of volume estimations.	44
5.1: MAVD values.	57
5.2: SDAVD values for Table 5.1.	58
5.3: Ratio values.	59

1. INTRODUCTION

1.1. Motivation

According to GLOBOCAN [GLOBOCAN, 2020], PCa is the second most prevalent cancer in men worldwide in 2020 and the most prevalent cancer in men worldwide for the last five years. As shown in Figure 1.1 [GLOBOCAN, 2020], PCa is the second in the number of new cases for the year 2020 worldwide among men of all ages. Figure 1.2 [GLOBOCAN, 2020] shows that PCa is the fifth in the number of deaths for the year 2020 worldwide among men of all ages.

As shown in Figure 1.3 [HSGM, 2017], PCa is the second most prevalent cancer type in Turkey in 2017 among men of all ages, according to HSGM [HSGM 2017]. Benign prostatic hyperplasia (BPH) is another common prostatic disease that affects most of the male population, especially as they age [Liu et al., 2020]. BPH generally results in lower urinary tract symptoms (LUTS) [Chughtai et al., 2016] and reduces life quality [Thorpe et al., 2003] while increasing health costs [Barry et al., 2001].

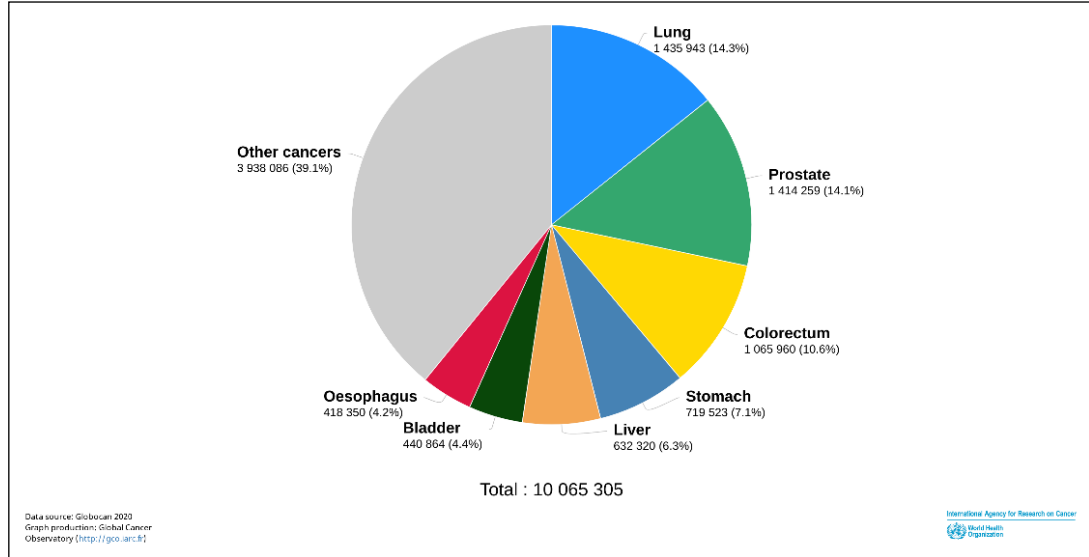


Figure 1.1: Estimated number of new cases in different cancer types for 2020. Values are worldwide, males, of all ages.

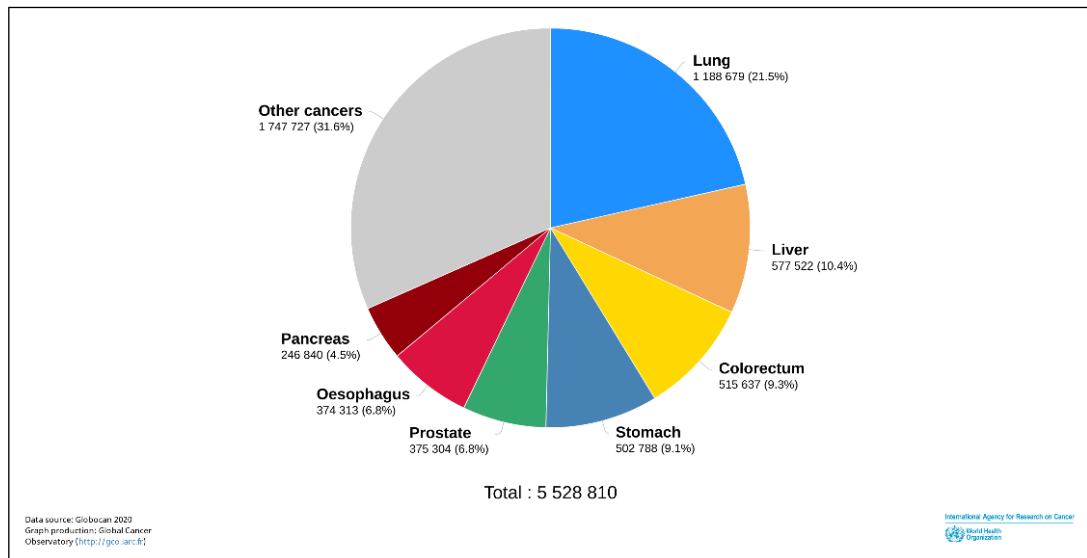


Figure 1.2: Estimated number of deaths in different cancer types for 2020. Values are worldwide, males, of all ages.

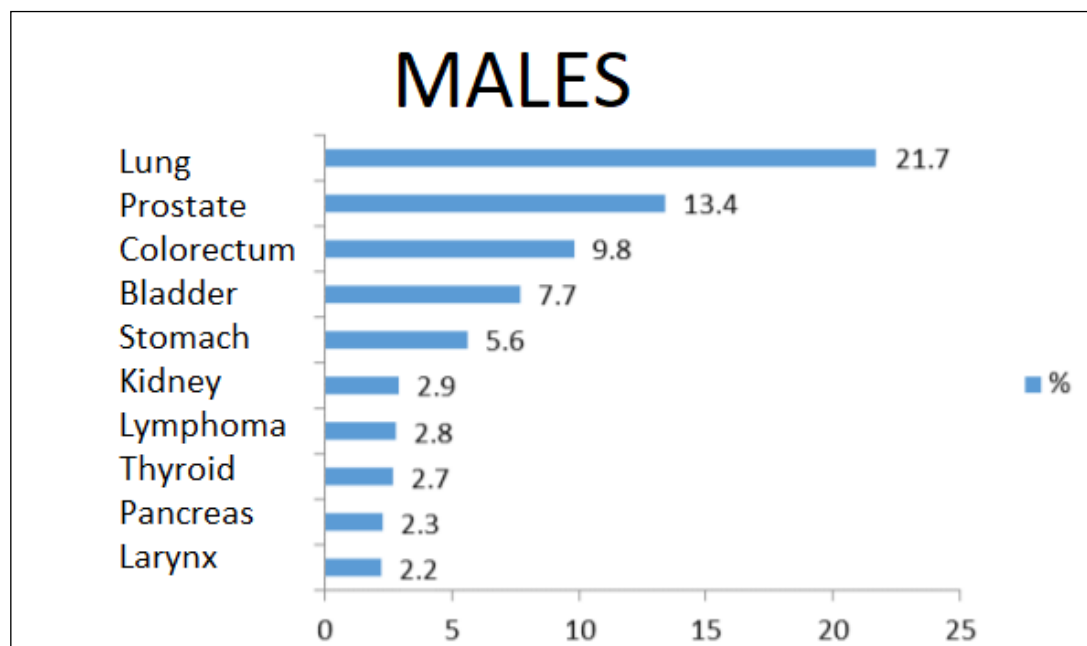


Figure 1.3: Prevalence of different cancer types in Turkey for 2017. Values are males, of all ages.

Prostate volume is a straightforward instrument during the diagnosing and management process of PCa and BPH. Literature shows an inverse relationship between prostate size and the incidence and aggressiveness of PCa [Yamashiro et al., 2021]. Prostate-Specific Antigen (PSA) value was used to detect and manage PCa but was not a proven reliable diagnostic tool. PSA Density (PSAD) appeared as a more

valuable tool by adjusting PSA with gland volume [Dunn et al., 2017], [Yu et al., 1998].

Table 1.1: Applications where prostate volume and/or boundaries are used.

Application	Parameter of Interest
<u>Diagnosis</u>	
PSA density	Volume
Assessment of benign disease	Volume
<u>Treatment</u>	
Prostate brachytherapy	Boundaries
High intensity focused ultrasound	Boundaries
Cryotherapy	Boundaries
Transurethral microwave therapy	Volume
Resection of benign tissue	Volume
<u>Follow up</u>	
Hormonal treatment	Volume

Accurate prostate volume measurement is essential for diagnosing cancer and determining the most effective treatment strategy [Wasserman et al., 2020]. It is also essential in diagnosing and managing BPH [Garvey et al., 2014], [Wasserman et al., 2015]. National American Urological Association guidelines for surgical management of BPH include examining prostate volume measurement prior to surgery [Foster et al., 2018]. Table 1.1 represents applications where prostate volume and/or boundaries are used [Pathak et al., 2000].

1.2. Problem Definition

The importance of accurate prostate volume measurement in diagnosing and managing prostatic diseases requires a more straightforward and reliable measurement process. There are many imaging techniques used in clinical practice to measure prostate volume. Commonly used techniques are US, MR, and CT [Turkbey et al., 2009]. Compared with others, US provides real-time imaging with lower costs. Also, US is portable and harmless, unlike other imaging techniques [Ghose et al., 2012].

Table 1.2: Advantages and disadvantages of TRUS and AUS images.

	Advantages	Disadvantages
TRUS	• Better contour representation • The prostate is in a fixed position • Shows only prostate	• Causes patient discomfort • Not for general purpose
AUS	• Easy and fast to apply • Can be used for general purpose • No discomfort for the patient	• Low SNR • Includes other anatomical structures • The prostate is represented smaller • The prostate is not in a fixed position

There are two frequently used US techniques in prostate imaging, TRUS and AUS. TRUS imaging is more commonly used due to its better imaging quality with a higher signal-to-noise ratio (SNR) and a larger view of the prostate with no other anatomical structures. Figure 1.4 shows examples of an AUS and a TRUS image, and one can see these differences. Besides these good sides of TRUS imaging, it is difficult in practice and causes patient discomfort. AUS is an easy-to-use alternative imaging technique where TRUS is not applicable. The advantages and disadvantages of TRUS and AUS images are represented in Table 1.2.

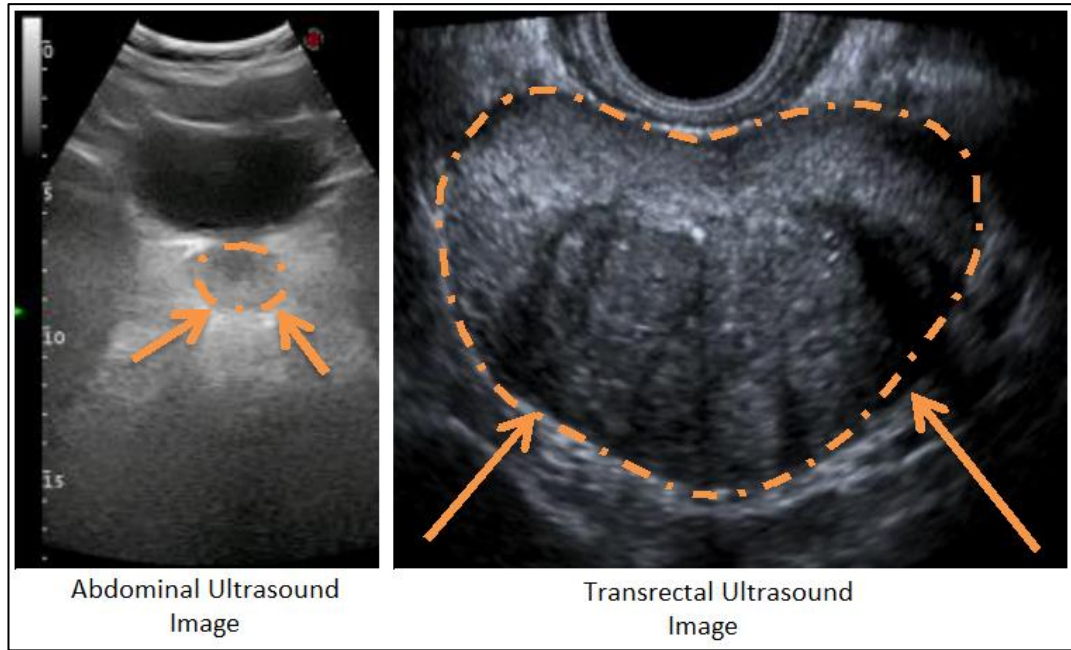


Figure 1.4: Sample AUS and TRUS images from the transverse plane. Prostate contours are given in orange dashed lines.

Generally, experts measure the prostate volume from medical images manually, but manual measurement results in high intra-expert and inter-expert differences caused by factors like imaging quality, personal experience, and human error [Choi et al., 2009]. Manual volume measurement is also time-consuming in the daily routines of clinicians.

Recent developments in automated prostate volume measurement systems show that the guidance of automated systems is possible where needed. As an imaging technique that did not get the attention it deserved in prostate volume measurement studies, AUS can be an excellent alternative to other imaging techniques with the guidance of automated systems.

1.3. Background

1.3.1. AUS Examination

Linear, sector, convex or wideband transducers with 2 to 5 MHz frequencies are used for adults' AUS imaging of the prostate gland. [Tyloc et al., 2016] The patient is positioned on the back during an examination. A well-filled urinary bladder is needed for a correctly performed examination. First, the transducer is applied transversely

above the pubic symphysis. The largest section of the prostate is searched to measure the width (W) and the height (H). Then the transducer is turned by 90° to obtain the sagittal section and measure the length (L) of the prostate. The acquisition methods from different planes are shown in Figure 1.5. To approximate the volume of the prostate, experts commonly use the standard ellipsoid formula,

$$V = \frac{W.H.L.\pi}{6} \quad (1.1)$$

where W, H, and L are the width, the height, and the length of the prostate.

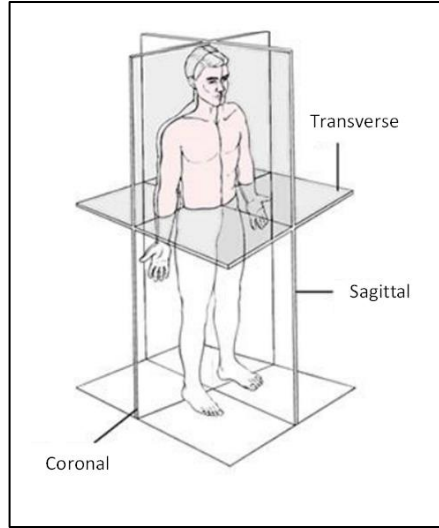


Figure 1.5: AUS image acquisition from different planes.

1.4. Outline

The rest of this thesis is structured as follows: Section 2 gives a literature review of AUS prostate segmentation studies. Section 3 proposes the DPM-based prostate localization method. Section 4 explains the size inference approach by using a patch-based segmentation method. In Section 5, a novel deep model is introduced, used in standard ellipsoid-based size inference. Section 6 concludes the study and gives future work.

2. LITERATURE REVIEW

Prostate volume measurement on US images is primarily performed by prostate segmentation on TRUS images. To the best of our knowledge, there is only one study on prostate segmentation on AUS images [Betrouni et al., 2005], apart from our previous work [Albayrak et al., 2015], [Albayrak et al., 2017a], [Albayrak et al., 2017b]. Therefore, this section will review prostate segmentation on TRUS images.

2.1. Contour and Shape-Based Methods

Automated prostate segmentation studies began with edge-based methods. Edge-based methods often use filters to extract edges from medical images. However, these algorithms needed to be supported by other methods and texture information due to the broken edges caused by low SNR values of US images.

Laplacian filtering was used in edge-based image segmentation methods. Aarnink et al. used a nonlinear Laplacian filtering-based method to determine the prostate contour from TRUS images [Aarnink et al., 1994]. They reflected the strength of an edge located by the Laplacian operator by the value of the gradient of the image at that location. Combining this information they constructed an edge intensity image. They performed edge enhancement techniques on the edge intensity image and then linked edge boundary edges by interpolation techniques.

[Liu et al., 1997] used a Radial Bas-Relief (RBR) technique to outline the prostate border. A bas-relief image is made by superimposing the positive image slightly onto the negative image and printing it as one image. The resulting image is simplified into a pattern of lines and flat tones. This work gives an incomplete border of the prostate where post-processing is needed. To obtain a smooth and complete border, [Kwoh et al., 1998] proposed to use the harmonics method on the incomplete border results of the RBR method. The harmonics method uses Fourier transforms to remove noise and encode a smooth boundary. Another problem with US images is the local differences in the visibility of the prostate border due to the orientation of the border sections, according to the US probe. [Aarnink et al., 1998] proposed to detect gray level transitions by using multiple scales of resolution in different stages to improve edge detection and localization. Using different filters in different stages, [Pathak et al., 2000] proposed a method to apply a sticks filter to enhance contrast and

reduce speckle noise. They smoothed the resulting image using an anisotropic diffusion filter in the second stage. They used some basic prior knowledge of the prostate at the final step, such as shape and echo pattern, to detect the most probable edges. Detected edges are manually linked, and by this means, patient-specific anatomical information is integrated.

Deformable models were a solution to the broken boundary problem of US images and were used in prostate-segmentation studies. Deformable models provide a complete contour. Its internal forces preserve the shape information while the external forces place in the best position representing the prostate border of the image. Active contour models (ACM) or snakes were proposed in [Kass et al., 1988] and they were used in literature commonly. Snakes were easy to implement and produced smooth contours. However, they needed reliable edge information and good initialization. Snakes change their shape towards the nearest salient contour through energy minimization. The energy consists of internal and external energy which forces the snake to salient image features like edges and puts the snake near the desired local minimum, respectively. Figure 2.1 represents snake energy minimization steps.

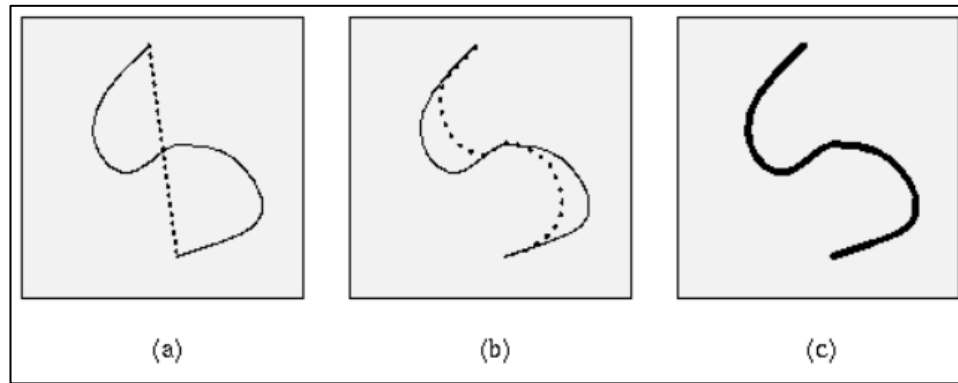


Figure 2.1: Snake energy minimization steps. (a) Initial snake position (dotted line) defined near the contour. (b) Iteration step of snake energy minimization. (c) Final snake position.

[Knoll et al., 1999] proposed a technique for elastic deformation restriction of snakes to particular object shapes. The technique used 1D dyadic wavelet transform-based localized multiscale contour parameterization as a multiscale boundary curve analysis tool. Since ACM needs initialization, many studies initialized the model with a few points on the prostate boundary which are selected by the user. 2D US slices were used in the literature to create a 3D volumetric image of the prostate and obtain

a better volume estimation. In [Ladak et al., 2000a] and [Wang et al., 2003], initialized a discrete dynamic contour (DDC) on a 2D TRUS slice from four points selected by the user. The DCC then refines the initial boundary to fit the prostate automatically. They propagated the estimated boundaries on 2D slices to obtain a 3D segmentation. [Nanayakkara et al., 2006] also initialized a DDC model from four seed points selected by the user. A fuzzy inference system is used to measure the reliability of points being on the boundary by using the information of small regions from inside and outside the boundary near the point. Zaim et al. proposed to replace the external energy of ACM with an energy map of a dot-pattern feature that detects dot-like patterns that normally appear within the prostate [Zaim et al., 2007a].

Similar to ACMs, deformable meshes deform towards the desired boundary starting from an initial location, under the influence of internal and external forces. Internal forces provide a smooth surface and external force drives the mesh towards the boundary. [Ghanei et al., 2001] creates the 3D volume image from 2D US slices using bilinear interpolation. The user initialized the 3D surface by drawing initial contours on a few 2D US images. The model is deformed by the movement of its vertices under internal and external forces. Internal forces were based on local curvature, while external forces were based on volumetric data by applying an appropriate edge filter. Using edge maps to drive the deformation of the model had weaknesses due to the characteristics of US images.

Level sets were introduced by Osher et al. and used in many medical image segmentation studies. A level set expands starting from a seed point in a direction normal to the desired curve surface. It stops where the intensity difference is highest in local. Kachouie et al. deformed an elliptic level set based on a velocity function derived using the first and second statistics of the TRUS image histogram [Kachouie et al., 2006].

Active shape models (ASM) were proposed to include prior shape information to the deformable model framework [Cootes et al., 1994]. That prevented the final segmentation results vary widely from the shape of the desired anatomical structures. Studies in literature extended ASM to overcome weaknesses like inaccuracy in large-scale variations and difficulty in extension to 3D. Using the Gabor filter bank to characterize prostate boundaries in a multiscale and multi-orientation fashion, [Shen et al., 2003] proposed a hierarchical shape deformation strategy. The Gaussian operation included in the Gabor filter bank removed the noise in US images while

multi-orientation filters provided edge directions. The hierarchical shape deformation strategy was enabled by the multiscale filters of the Gabor filter bank. That approach increased the robustness and accuracy of the system. Utilizing more than one model enables one to address the weaknesses of models and the characteristic challenges of US images. For this purpose, [Yan et al., 2010] proposed to use a partial active shape model (PASM) with a discrete deformable model (DDM). Their system estimated the shape prior by the PASM which uses some parts of the contour with salient boundary features. They used a point distribution model (PDM) and perpendicular normal vector profiles to obtain contrast variations to determine salient boundary points by discarding shadow regions. Then the DDM searched for the final boundary in a multi-resolution, energy minimization framework under the guidance of the shape prior. In Figure 2.2 (a) the normal vector profiles with two candidate points and (b) a contour with salient points in red are shown [Yan et al., 2010].

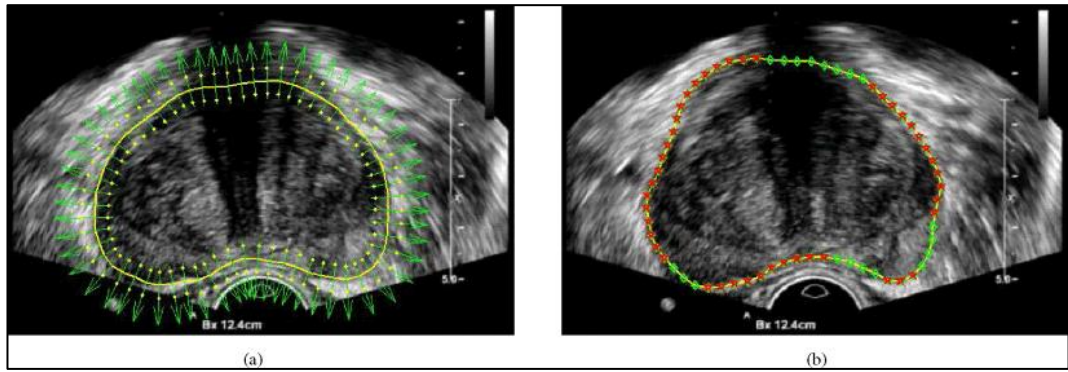


Figure 2.2: Representation of partial active shape model. (a) Normal vector profiles with two candidate points. (b) A contour with salient points in red.

Most of the model-based approaches were semiautomatic due to the model initialization problems. [Gong et al., 2004] proposed a solution for this problem with deformable superellipses. This model was fitted to manual prostate contours outlined by experts, and then with simple parametric deformations, it could fit the prostate shape efficiently. They developed a robust Bayesian segmentation algorithm based on the superellipses model.

2.2. Region-Based Methods

Intensity distributions of the prostate region are the focal point of region-based methods. Chan et al. proposed to use region-based statistics for the energy minimization during the propagation of the level set [Chen et al., 2001]. This method addressed the absence of strong edges problem presence of white noise and made the implementation of prostate segmentation in 3D easier.

Fan et al. applied the level set method to segment 3D US prostate images which consist of transverse, sagittal, and coronal images of the prostate [Fan et al., 2002]. They first used a fast discrimination method to extract the prostate region according to the intensity likelihood. Then they integrated the region information instead of the image gradient into the level set method. Another approach exploits small regions to create a reliability measure for a pixel on the boundary. Using shape and intensity priors, Yang et al. proposed to segment 3D ultrasound volumes in a level set framework [Yang et al., 2006]. They stacked 3D manual segmentations and used min/max flow to obtain a smooth 3D surface. In a Bayesian interface, they integrated the nonlinear shape model and the nonparametric intensity prior to defining a region-based energy function for the propagation of the level set.

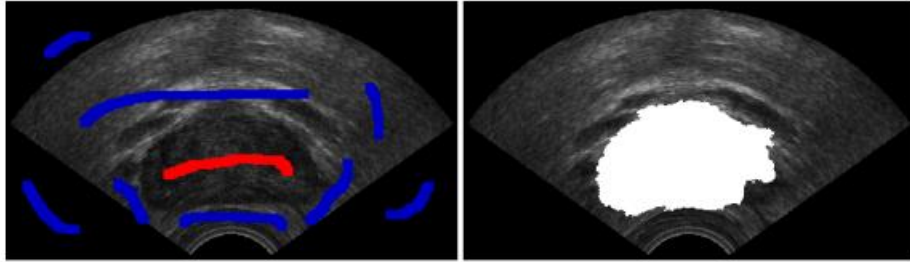


Figure 2.3: Representation of graph cut segmentation. On the left, user input of constraints, and on the right, graph cut segmentation result.

Another solution to the absence of strong edges problem is the graph-partition algorithm. In graph-partition algorithms, pixels or pixel groups are defined as nodes while the gradients between pixels are defined as costs. The graph is then partitioned by minimizing the cost function. Zouqi et al. proposed a graph-partition framework where pixels were the nodes and the edges were the edge discontinuity penalties [Zouqi et al., 2008]. Two special nodes were built from the object and the background by the selection of the user: the source and the sink. The max-flow algorithm increased

the flow from the source to the sink. The edges corresponding to the minimum cost graph were the optimal segmentation. Figure 2.3 represents user input of constraints on the left and graph cut segmentation result on the right [Zouki et al., 2008].

2.3. Classification-Based Methods

Classifier-based methods create a training set by assigning a class label to each object, such as pixel or region, with supervision. The classifier is then trained with features created from the training set and learns to assign a class label to an unseen object. Many different classifiers were used in the literature like neural networks and support vector machines (SVM). Besides determining the classifier, the main challenge of classification-based studies is the extraction of the best features for the task.

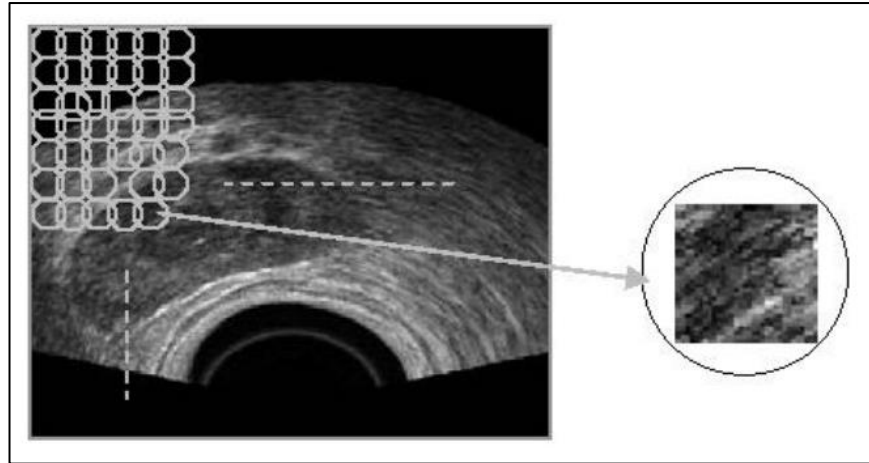


Figure 2.4: Feature extraction process for the classification.

Zaim employed spatial, texture, and gray-level information to train a self-organizing map network [Zaim, 2005]. Zaim et al. then used an SVM for the classification of prostate and non-prostate tissue [Zaim et al., 2007b]. They proposed using multi-wavelet decomposition for feature extraction. They extracted overlapping windows and used their entropy and energy of symmetric, orthonormal, and second-order wavelet coefficients. The feature extraction process is shown in Figure 2.4 [Zaim et al., 2007b].

[Yang et al., 2016] also used a patch-based feature learning framework. Patient-specific anatomical features were extracted for each voxel, and a kernel support vector

machine (KSVM) was trained. The trained KSVM was used to localize the prostate of the new patient.

Hybrid network structures were used in many studies for better classification performance and less computational time. Dokur et al. proposed a hybrid neural network that consists of three hidden layers for the segmentation of TRUS images [Dokur et al., 2002]. They called the proposed network intersecting spheres (InS). They extracted features by using the discrete cosine transform (DCT) of pixel intensities in the ROI. The hybrid network structure is shown in Figure 2.5 [Dokur et al., 2002]. The first hidden layer consists of genetic algorithms (GAs) where each node forms a closed region which is represented by a hypersphere (HS).

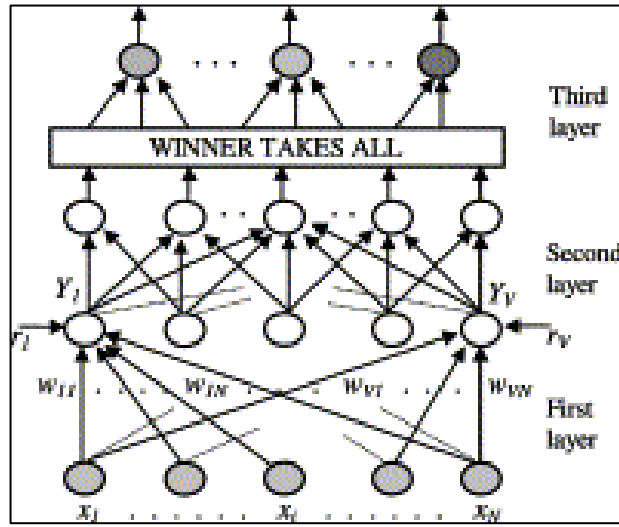


Figure 2.5: Hybrid network structure.

The similarity of textures in prostate and non-prostate regions, and the variation of textures of the same tissue at different zones of the prostate, challenge many classification algorithms. As a solution, Akbari et al. Proposed to use a wavelet support vector machine (W-SVM) and a static shape model together [Akbari and Fei, 2012]. A set of W-SVMs was employed to adaptively collect texture priors of the prostate. The textures of different tissues were then classified by the W-SVM set using wavelet transform features. After the feature extraction step, shape statistics were incorporated into a shape model to modify the segmentation. Several aspects of the features were extracted by several wavelet transforms. Textures that are ambiguous in some regions or planes are classified by feature extraction from three planes.

2.4. Deep Learning-Based Methods

Deep learning methods moved the feature extraction process from the human side to the algorithm side. Promising results in deep learning studies inspired experts in many areas to use deep learning in their studies. [Yang et al., 2017] exploited recurrent neural networks (RNNs) to infer prostate shape by exploring shape priors sequentially. For this method, they previously created dynamic sequential ultrasound images from static ultrasound images. They also adopted a multi-view fusion strategy to merge shape predictions observed from different perspectives. To segment the prostate during a real-time ultrasound scan, [Anas et al., 2018] deployed an RNN that received current and previous frames as input. The proposed network structure used both immediate and distant past frames to improve the detection of the prostate shape while increasing the robustness to imaging artifacts. [Ghavami et al., 2018] used an adapted U-Net [Ronneberger et al., 2015] architecture with interoperative TRUS images for 2D slice segmentation and 3D volume segmentation. Most studies on medical images needed solutions to overcome the data limitation problem, like integrating supervision mechanisms into deep models. [Lei et al., 2019] proposed a deep supervision strategy with a hybrid loss in different stages to overcome the data limitation problem. 3D original TRUS and multiple filtered images were used as multichannel samples to train the proposed network to enhance performance. Although deep learning structures are powerful tools to learn multi-level features, directly applying them causes poor localization of prostate boundaries due to redundant features extracted because of the characteristics of TRUS images. To overcome the occasion, [Wang et al., 2019] proposed deep attentive features generated at each layer by learning the complementary information of the low-level detail and high-level semantics in multi-layer features. They claimed that deep attentive features represent prostate characteristics better.

2.5. Hybrid Methods

Hybrid methods combine different segmentation methods to exploit different advantages of different methods. Combining contour, shape, region, or classification information makes these methods robust to imaging problems such as artifacts or noise and improves segmentation accuracy.

Shape and texture features were modeled and used together in literature by many researchers. The creation of the texture and shape models varied among different studies. Zhan et al. created a texture space by using 2D rotation invariant Gabor filters for each of the two panes of 3D US prostate images [Zhan et al., 2003]. They classified the texture space into the prostate and non-prostate regions via a kernel support vector machine (KSVM). They used the classified features as an external force in a deformable model framework to segment the prostate. In their consequent work [Zhan et al., 2006], they proposed to train a Gabor SVM (G-SVM) which consists of a Gabor filter bank and a KSVM, for each local patch of the surface. The schematic representation of the G-SVM is shown in Figure 2.6 [Zhan et al., 2006]. The Gabor filter bank of the G-SVM extracted Gabor texture features around the corresponding voxel. Then the pretrained KSVM gives a label indicating the likelihood of the voxel belonging to the prostate. Subsequently, the deformable model is forced to the prostate boundary by the label of G-SVM. The process is carried out until convergence.

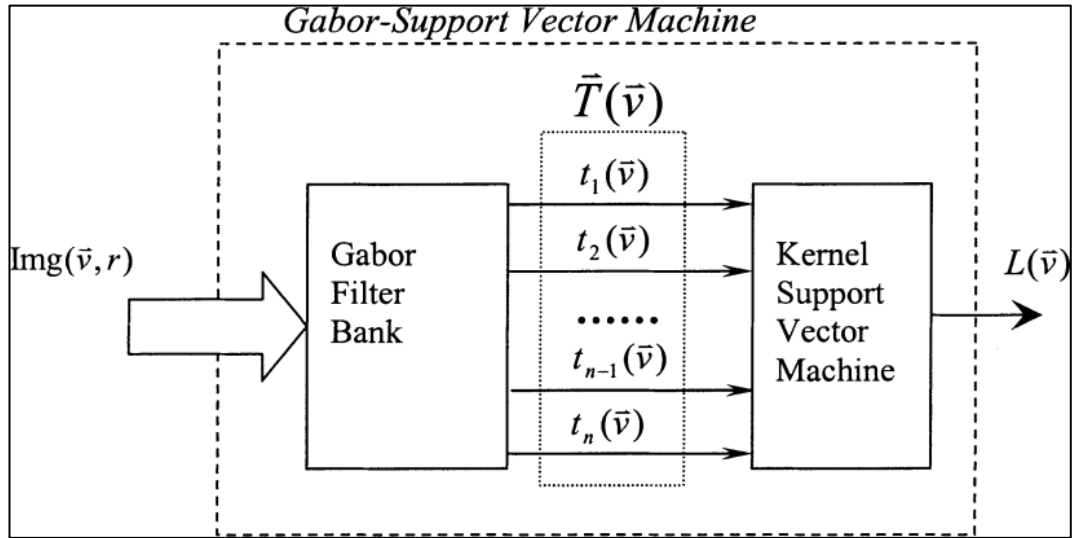


Figure 2.6: Schema of Gabor-Support Vector Machine.

Cosio proposed a robust initialization of an ASM with position and grayscale information of TRUS images followed by optimization with a multi-population genetic algorithm [Cosio, 2008]. The initialization of the ASM was done with a binary image created by a Bayes classifier from a Gaussian mixture model of position and grayscale including three Gaussians, prostate, non-prostate, and halo around the

prostate. The optimization was done by finding the optimum combination of pose and shape parameters by a multi-population genetic algorithm.

[Ghose et al., 2013] proposed an approach based on multiple mean parametric models obtained from principal component analysis (PCA) of shape and posterior probabilities in a multiresolution framework. The model parameters were modified with prior knowledge to achieve optimal prostate segmentation. The posterior probabilities of the prostate region were determined from random forest classification and used for building, initialization, and propagation of the appearance model.

The uncertainty of the predictions in deep learning-based methods can be reduced by using prior knowledge. [Karimi et al., 2019] propose to train a CNN ensemble to identify uncertain segmentations and create a segmentation uncertainty map. They then utilize the shape-prior to improve uncertain segmentations.

3. PROSTATE LOCALIZATION FROM AUS IMAGES BY USING AN SVM BASED GRAPHICAL MODEL

3.1. Motivation and Problem Definition

Automated prostate segmentation studies showed that prostate segmentation on TRUS images needed an initialization, even though the prostate is the only organ shown. Edge-based studies filtered falsely detected edges after the initialization of the edges roughly [Pathak et al., 2000]. Deformable model-based studies initialized the contour from manually outlined samples [Shen et al., 2003] or user selection of several points on the contour [Ladak et al., 2000b].

To study AUS images, localization of the prostate is even more crucial because it involves other anatomical structures, and it has many characteristic challenges. Detection of the prostate location provides prior knowledge for prostate segmentation studies. Besides, utilizing anatomical structures other than the prostate would enhance the localization results.

3.2. Proposed Method

In this study, a prostate localization method is proposed that utilizes the organs and tissues other than the prostate on AUS images. The proposed method is based on DPM, which allows modeling the group of anatomical parts included in an AUS image. An early version of DPM called pictorial structures is shown in Figure 3.1 where a face is modeled [Fischler et al., 1973]. DPM yields global results for a given image while working locally. DPM optimizes its location on a given image according to the appearance and the context energy starting from an initial location. The appearance energy of the model is obtained from the image information at the locations of each part, while the context energy is obtained from the spatial relations between the parts of the model. DPM was used in literature in many object detection frameworks. Figure 3.2 shows an example of DPM on a car and a person model [Felzenszwalb et al., 2010].

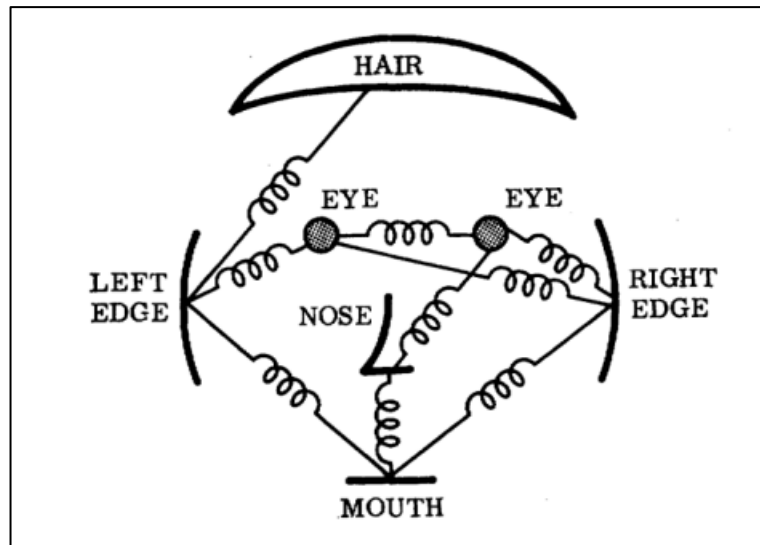


Figure 3.1: An early version of the DPM model called pictorial structure. This is an example of a face model.



Figure 3.2: DPM samples on a car and a person model. Their corresponding appearance energy maps are given on the lower part.

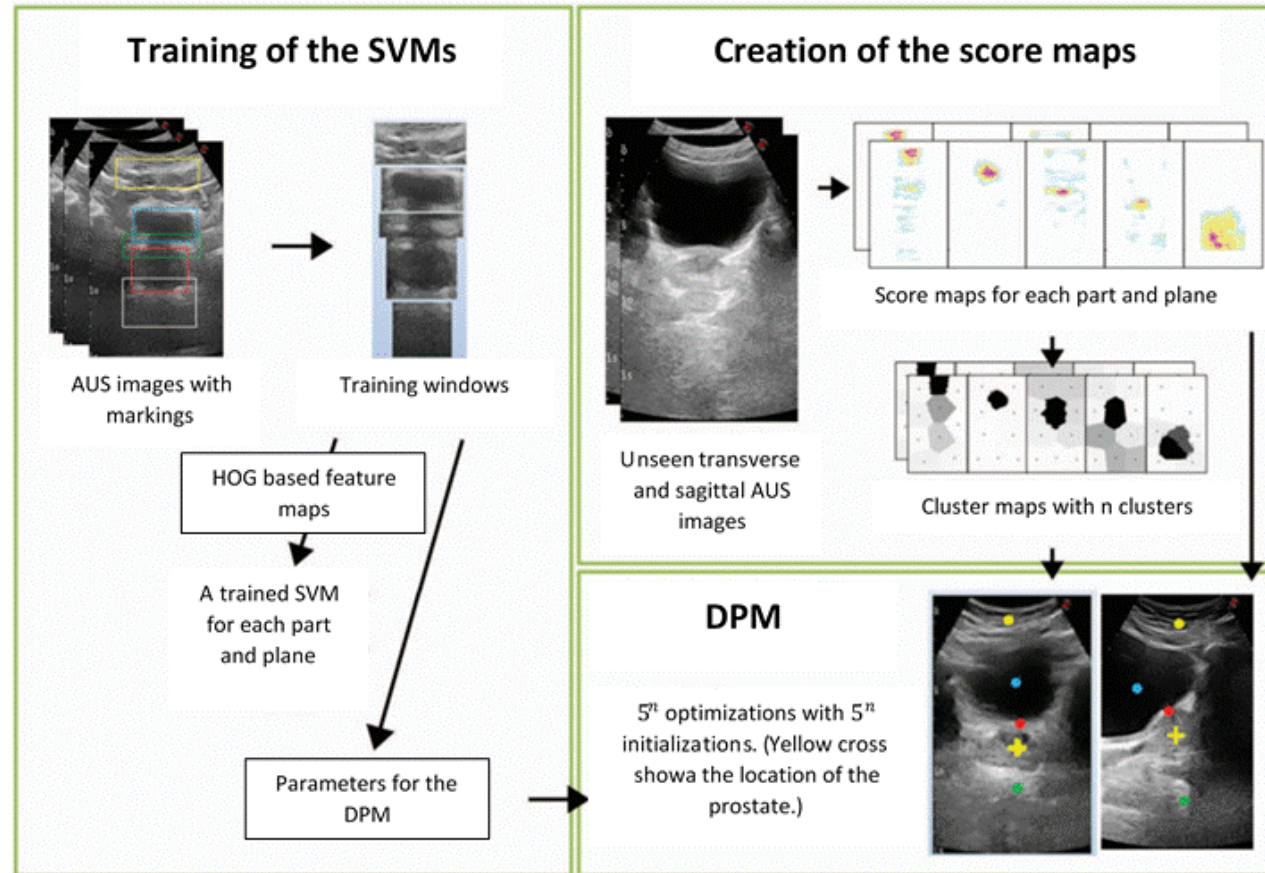


Figure 3.3: Representation of the proposed system. An SVM was trained for each of the anatomical structures. Score maps and corresponding cluster maps were created and used for the DPM initialization and optimization.

It was observed that a machine learning system trained with image windows of a single anatomical structure is not adequate to detect the location of the structure. The low SNR and uncertainty of AUS images cause this problem. Therefore, performing the localization process with the help of neighboring structures is more appropriate and it produces better results than a system that will work by considering only the prostate. The DPM method is quite suitable for solving this multi-part localization problem. Figure 3.3 shows the overall representation of the proposed system.

Five structures were determined including the prostate to apply the DPM-based method as a solution to the problem. Figure 3.4 shows these structures for the transverse (a) and the sagittal (b) AUS images. The first structure is shown in a yellow window and appears as bright waves. The second structure is the bladder which is shown in a green window. The third structure is the shiny tissue between the prostate and bladder and is shown with a blue window. The fourth structure is the prostate, and the fifth structure is shown in a white window, and it is the shiny tissue under the prostate.

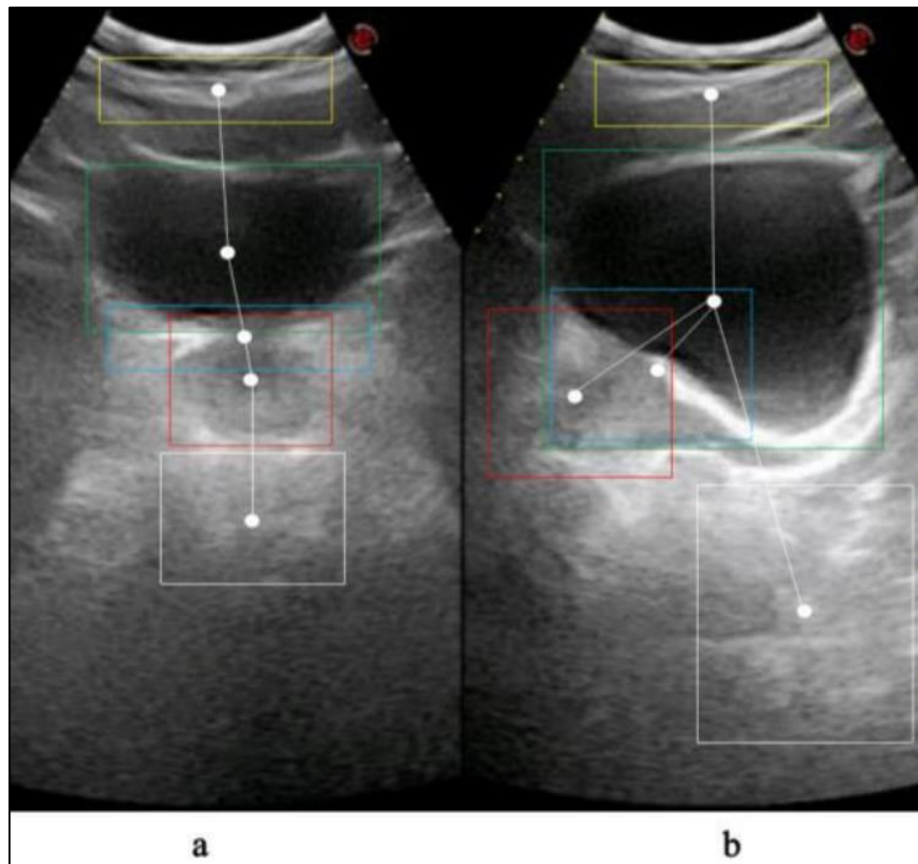


Figure 3.4: Anatomical structures determined for the DPM-based method.

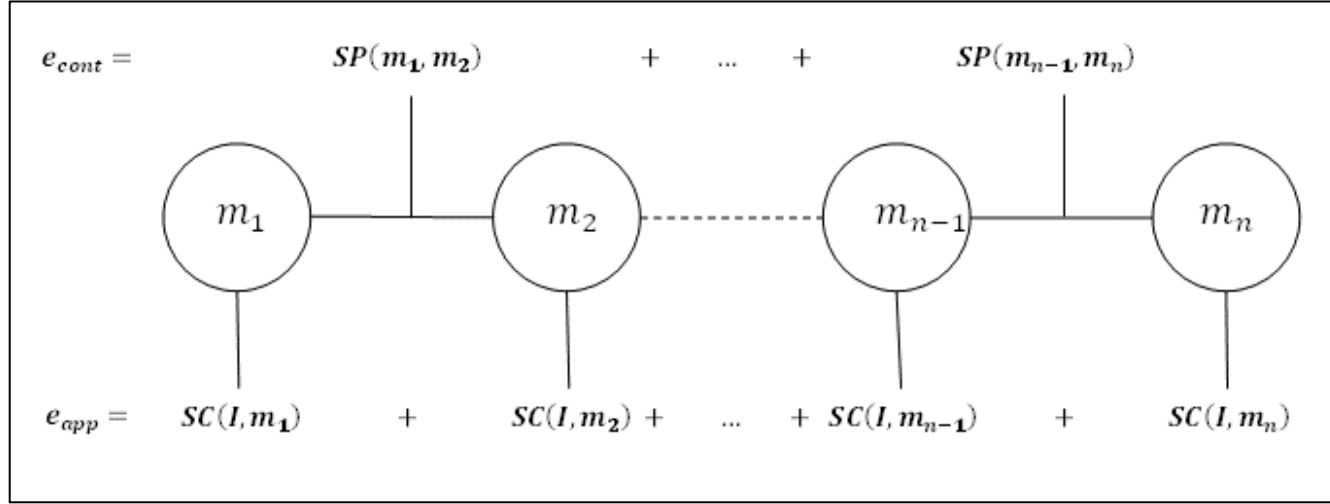


Figure 3.5: The DPM design for the proposed system. $m_1, \dots, m_n$ represent the parts of the DPM. The function $SP(m_i, m_j)$ gives the spatial relation value between parts m_i and m_j . Context energy e_{cont} is the sum of spatial relation values between neighboring parts. $SC(I, m_i)$ function gives the score value obtained by a machine learning system for the location of m_i on image I . Appearance energy e_{app} is the sum of score values.

Assume $M = \{m_1, m_2, \dots, m_n\}$ is a localization of the parts of the DPM where m_i represents the location of the i^{th} part and $\hat{M}$ is the optimal localization. The proposed DPM optimizes its location from M to $\hat{M}$ on an image I , by using (3.1) where, $\emptyset$ represents the parameters learned from the training data.

$$\hat{M} = \underset{M}{\operatorname{argmax}} P(M|I, \emptyset) \quad (3.1)$$

The function, $P(M|I, \emptyset)$ consists of the appearance e_{app} and the context e_{cont} energies modeled as a Gibbs distribution which is given in (3.2). In this formula, z is a normalizing constant.

$$P(M|I, \emptyset) = \frac{e^{-(e_{app} + e_{cont})}}{z} \quad (3.2)$$

Figure 3.5 shows the DPM design for the proposed system. $m_1, \dots, m_n$ represent the parts of the DPM. The function $SP(m_i, m_j)$ gives the spatial relation value between parts m_i and m_j . Context energy e_{cont} is the sum of spatial relation values between neighboring parts. $SC(I, m_i)$ function gives the score value obtained by a machine learning system for the location of m_i on image I . Appearance energy e_{app} is the sum of score values. Computation of the appearance and the context energy, e_{app} and e_{cont} is explained in the following sections.

Levenberg-Marquardt algorithm [Marquardt, 1963] (LMA) was used to optimize (3.1). LMA is a gradient descent algorithm that works locally, and it is more successful than commonly used gradient descent algorithms with complex models.

3.2.1. Computation of the Appearance Energy

A score map was created for each structure and it was used to compute the appearance energy, e_{app} of a localization, M of the model. For this purpose, a support vector machine (SVM) was trained for each structure. The training was done with feature vectors extracted from image windows of the training data. Feature vectors of image windows, including entirely one of the structures, were considered positive samples of that structure, while feature vectors of image windows with only a part of the structure or no structure were considered negative samples.

Given an unseen image, a score map for each structure was created using each of the trained SVMs. For this purpose, feature vectors from image windows at successive locations were extracted using the sliding window method. Each SVM outputted a score value for each of the feature vectors. The score values were added to the location of the center of the image window in corresponding score maps. Scores of more than one window size around the mean values of the training data are created, and the highest score is chosen. We saved these window sizes to use later while calculating the context energy, which is explained in the following section.

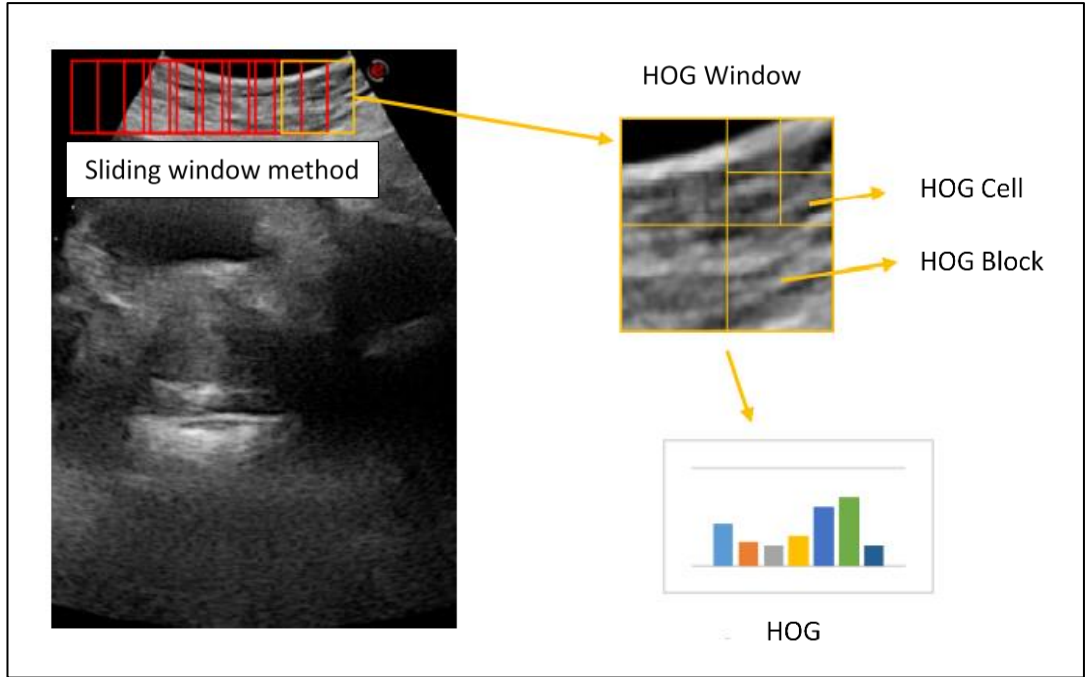


Figure 3.6: Creation of HOG feature vectors. Implementation of the sliding window method.

The score maps were then used to compute the appearance energy, e_{app} . The computation of e_{app} is explained in (3.3), and it is the total of the score map values at the locations of each structure.

$$e_{app} = \sum_i SC(I, m_i) \quad (3.3)$$

The histogram of oriented gradients (HOG) method was chosen as the feature vector extraction method. HOG was created by summing the gradients of the image into bins according to the orientations of the gradients. The image window was masked

with an operator which computes the gradient. The Sobel operator was used for this purpose. As shown in Figure 3.6, the image window was divided into blocks and cells, and the gradient values were summed and normalized considering the cells. The histogram was created according to the orientations of the gradients with these normalized values. The feature vector of the image window was obtained by concatenating the histograms of each block.

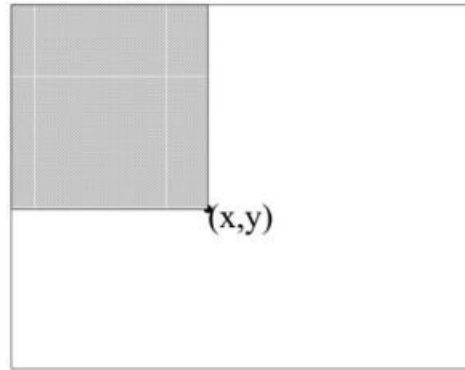


Figure 3.7: Integral image representation. Each location on an integral image is the sum of the top and the left locations.

While creating the score maps, many overlapping image windows were created with the sliding window method. We used the integral image method to avoid the recurring calculation of gradient values [Viola and Jones, 2001]. As shown in Figure 3.7, the value of each location is the sum of the top and the left locations. We can quickly obtain the sum of a region by using integral images, as explained in Figure 3.8.

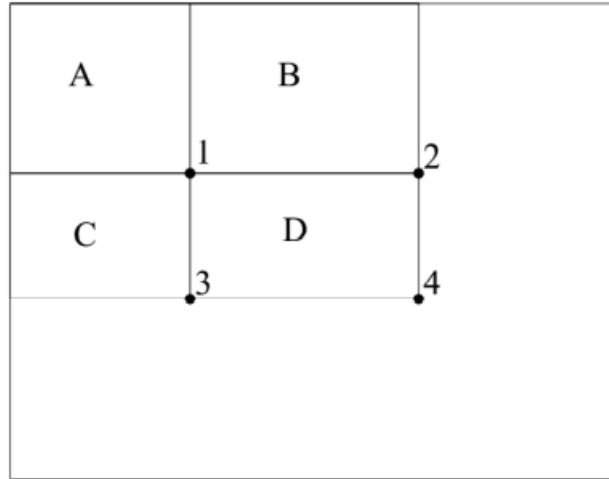


Figure 3.8: Integral image region sum calculation. Given an integral image, the sum of region D is easily obtained. Here, point 1 is the sum of A, point 2 is the sum of A and B, point 3 is the sum of A and C, and point 4 is the sum of A, B, C, and D. Hence, we can obtain the sum of D as point 4 + point 1 – (point 2 + point 3).

Due to AUS images' structural challenges, a model that learns only one structure would not detect locations as accurately as a model with more than one cooperating structure. Considering independent score maps of 5 different structures shown in Figure 3.9, one can see many maximums, including actual and false locations. Especially structures with similar characteristics like bladder and prostate may be represented as false maximums in each other's maps.

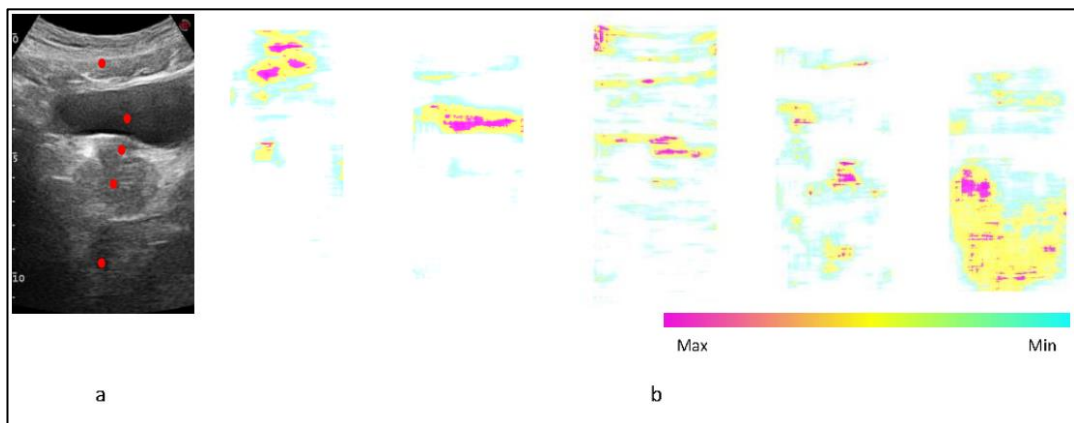


Figure 3.9: Sample score maps for a given AUS image. a) Locations of five structures from manual annotations. b) Score maps for the structures from top to down.

DPM overcame the challenges caused by the false maximum locations on the score maps by gathering different structures in a model. Since DPM works locally, the

initialization of the model was essential. Methods like initializing the parts from the maximums of the score maps or clustered score maps do not work properly because of the false maximum locations problem mentioned above.

3.2.2. Computation of the Context Energy

The context energy of the DPM was calculated considering the distance, orientation, and size relationships between its parts, as shown in (3.4) and (3.5).

$$e_{cont} = \sum_{i=1}^{n-1} SP(m_i, m_{i+1}) \quad (3.4)$$

$$SP(m_i, m_{i+1}) = \alpha \cdot f_d(m_i, m_{i+1}) + \beta \cdot f_o(m_i, m_{i+1}) + \delta \cdot f_s(m_i, m_{i+1}) \quad (3.5)$$

In (3.5), α , β , and δ are the weights of the distance, the orientation, and the size terms, respectively. (3.6), (3.7), and (3.8) show the computation of the distance, the orientation, and the size terms of the context energy.

In (3.6), $d_{i,i+1}$ is the ratio between the distance between the i^{th} and the $(i + 1)^{th}$ parts, and the sum of the distances between all parts. Using this ratio instead of using only the distance value made the system robust to distance values varying from sample to sample. μ_i^d and σ_i^d represent the mean and the standard deviation values of the distance term according to the training data.

$$f_d(m_i, m_{i+1}) = \frac{(d_{i,i+1} - \mu^d)^2}{(\sigma^d)^2} \quad (3.6)$$

$$f_o(m_i, m_{i+1}) = \frac{(o_{i,i+1} - \mu^o)^2}{(\sigma^o)^2} \quad (3.7)$$

$$f_s(m_i, m_{i+1}) = \frac{(s_{i,i+1} - \mu^s)^2}{(\sigma^s)^2} \quad (3.8)$$

In (3.7), $o_{i,i+1}$ is the orientation between the i^{th} and the $(i + 1)^{th}$ parts. μ^a and σ^a represent the mean and the standard deviation values of the orientation term according to the training data.

In (3.8), $s_{i,i+1}$ is the ratio between the size of the i^{th} and the $(i + 1)^{th}$ parts. μ^s and σ^s represent the mean and the standard deviation values of the size term according to the training data. As mentioned in the previous section, each part's size at each location was determined while creating the score maps.

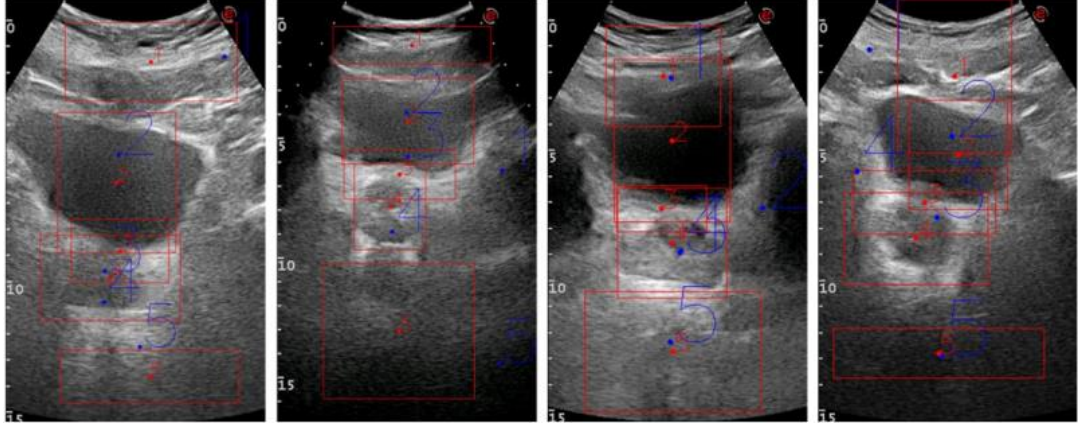


Figure 3.10: The initial and the final locations of the DPM. The blue dots and numbers represent the initial locations of the parts, while the red dots and numbers represent the final locations of the parts. The windows show the final estimated size value for the anatomical structures.

Figure 3.10 shows sample initial and final locations of the DPM. Numbered blue dots represent the initial locations of DPM parts. Numbered red dots and red windows represent the final locations and estimated sizes, respectively. Here one can see that the initial locations determined according to the score maps are faulty because the score maps were created considering only single parts of the DPM. On the contrary, with the enhancement of the DPM, one can see that the final locations are more accurate, and the window sizes are proper.

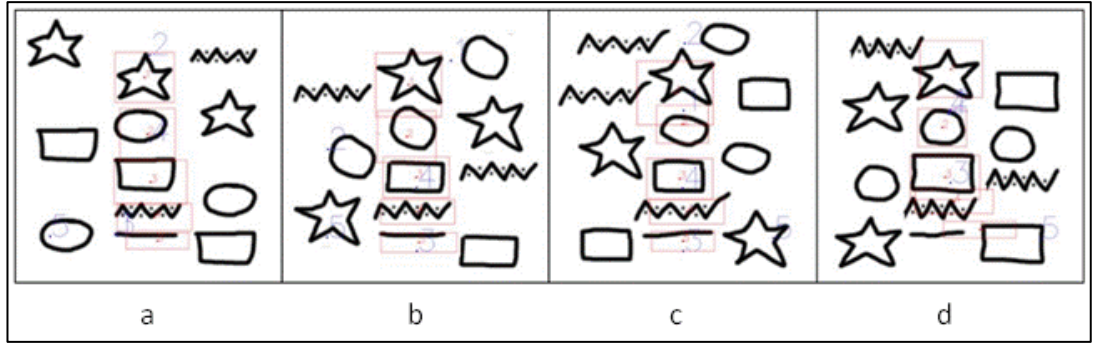


Figure 3.11: DPM optimization on synthetic images. Blue dots and numbers represent the initial locations, while red dots and numbers represent the final locations for the parts. The red windows represent the final estimated sizes.

Experiments on synthetic images showed the effect of the context energy of the DPM. As shown in Figure 3.11, synthetic images with more than one similar structure were created as in AUS images. Blue dots and numbers represent the initial locations of the DPM, while red dots and numbers represent the final locations. Though score maps of the parts include high scores for similar structures, the context energy of DPM moved them to accurate locations.

Since the system works locally, the initialization of the locations of the parts was essential. Initializing parts from the wrong locations affected the model's energy and caused the model to move to the wrong locations. In some cases, the optimization needed more loops to move all parts to their optimal location. In Figure 3.11-d, one can see that part number 5 needed more loops to move its optimal location.

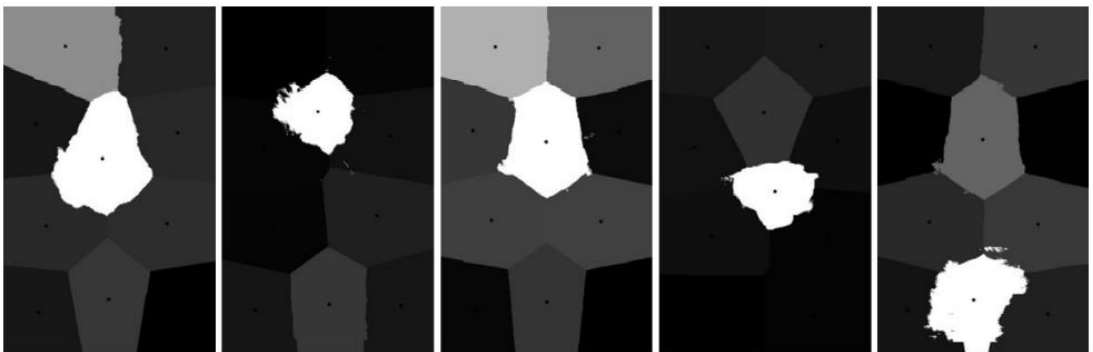


Figure 3.12: Sample cluster maps of an AUS image. Each of these cluster maps was created by clustering the score map of the corresponding part.

To solve the initialization problem, multiple initializations of the parts of the DPM were proposed. For this purpose, cluster maps with k clusters were created from the score maps, and each of the n parts of the DPM was initialized from the centers of

the k cluster. This way, n^k initializations and optimizations were done, and the localization with the best score was chosen as the final localization. Although n initialization seems like a very high number for the optimization method, it could be completed quickly because the LMA algorithm is efficient. As mentioned before, n (i.e., the number of parts) was 5. k was chosen as 10 due to the experiments. Figure 3.12 shows the clustering maps obtained from the score maps of an AUS image.

3.3. Experiments

140 transverse AUS images from different patients were used for the experiments. The size of the images was between 295x470 and 380x610 pixels. Our dataset consisted of healthy and diseased patient samples. Experts outlined the anatomical structures determined as the DPM parts on all images. We used ten-fold cross-validation to validate the system.

The proposed DPM-based system achieved a higher accuracy by utilizing more than one anatomical structure instead of only an anatomical structure. As we mentioned in our previous study [Albayrak et al., 2015], only 55.7% of the detected prostate centers were inside actual prostate contours when only feature maps of an individual structure were used. With the use of DPM, this value achieved 83.5%.

When only a feature map of the prostate was used, the average distance between the detected and marked centers was 2.1 cm. The value decreased to 1.08 cm with the use of the DPM-based system. Figure 3.13 compares the SVM-based system that detects only the prostate and the proposed DPM-based system. A boxplot is shown for each of the systems with the distance values in cm between the detected center and the actual prostate center. The red mark inside the box represents the median, the edges of the box are the 25th and the 75th percentiles, and the plus signs are the outliers. One can see that the proposed system has smaller distance values than the SVM-based system and the SVM-based system's results are not consistent.

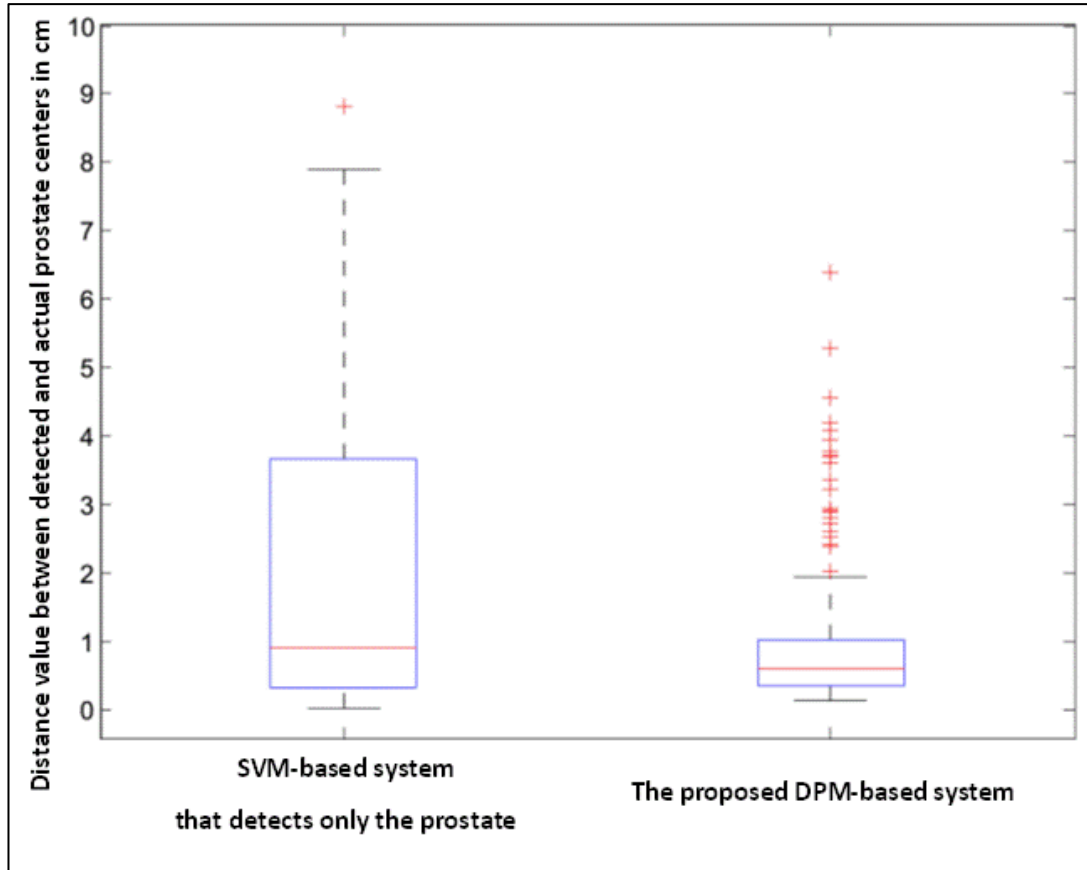


Figure 3.13: Comparison of the SVM-based and DPM-based systems' results. The red mark inside the box represents the median, the edges of the box are the 25th and the 75th percentiles, and the plus signs are the outliers.

Figure 3.14 shows sample results from the DPM-based localization method. In the figure, blue dots and numbers represent the initial locations of the part centers, while red dots and numbers represent the final locations of the part centers. The red windows represent the estimated size values. Considering the initial part locations determined using the SVM-based score maps and the final part locations determined by the DPM-based system, one can see the effect of the DPM-based system on better localization of the parts.

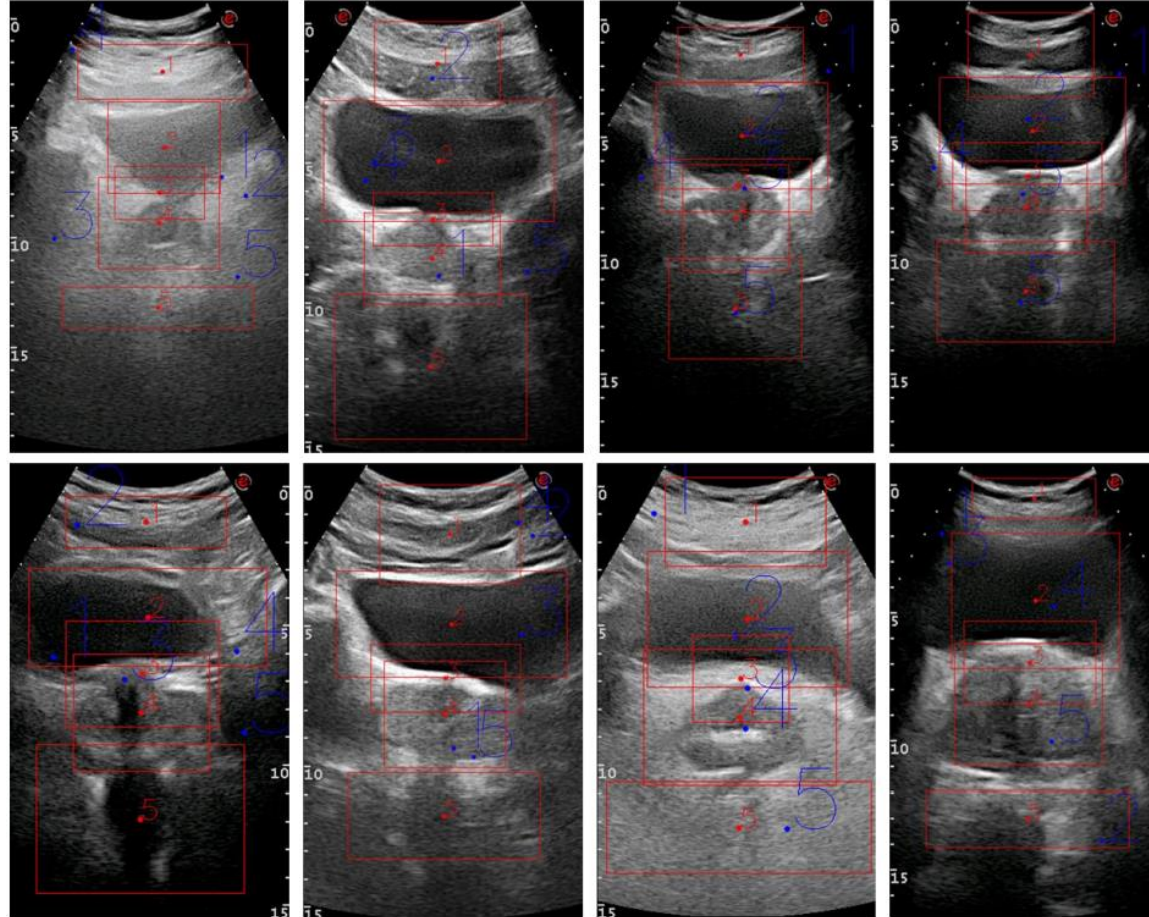


Figure 3.14: Sample results of the DPM-based localization method. Red dots and numbers represent the detected prostate centers while the red windows represent the estimated size values.

4. PROSTATE SIZE INFERENCE BY PATCH-BASED SEGMENTATION

4.1. Motivation and Problem Definition

Building a fully automated AUS prostate size inference system requires working with raw images that include several anatomical structures and the prostate. Thus, such a size inference system should be able to learn both object-level and lower-level features. By this means, it should detect both prostate location and boundary using global and local information.

A deep convolutional neural network (DCNN) is a promising tool for simultaneously learning intermediate features and final tasks from raw AUS images. The need for massive data makes using DCNN compelling in complex tasks like contour extraction or segmentation. Insufficient data may result in over-fitting during training due to the significant number of parameters. This problem arises primarily in domains where collecting data or hand labeling is very expensive or complicated, such as medical image analysis [Careniro et al., 2011].

4.2. Proposed Method

A patch-based deep learning system was proposed for prostate segmentation and size inference. A DCNN for each plane was trained with patch-based training data extracted from corresponding AUS images. For an unseen image, trained DCNNs were used to create contour maps. Then, the proposed system extracted the final contours from the contour maps, yielding the segmentation results. The final contours were used for the size inference. The general schema of the system is represented in Figure 4.1.

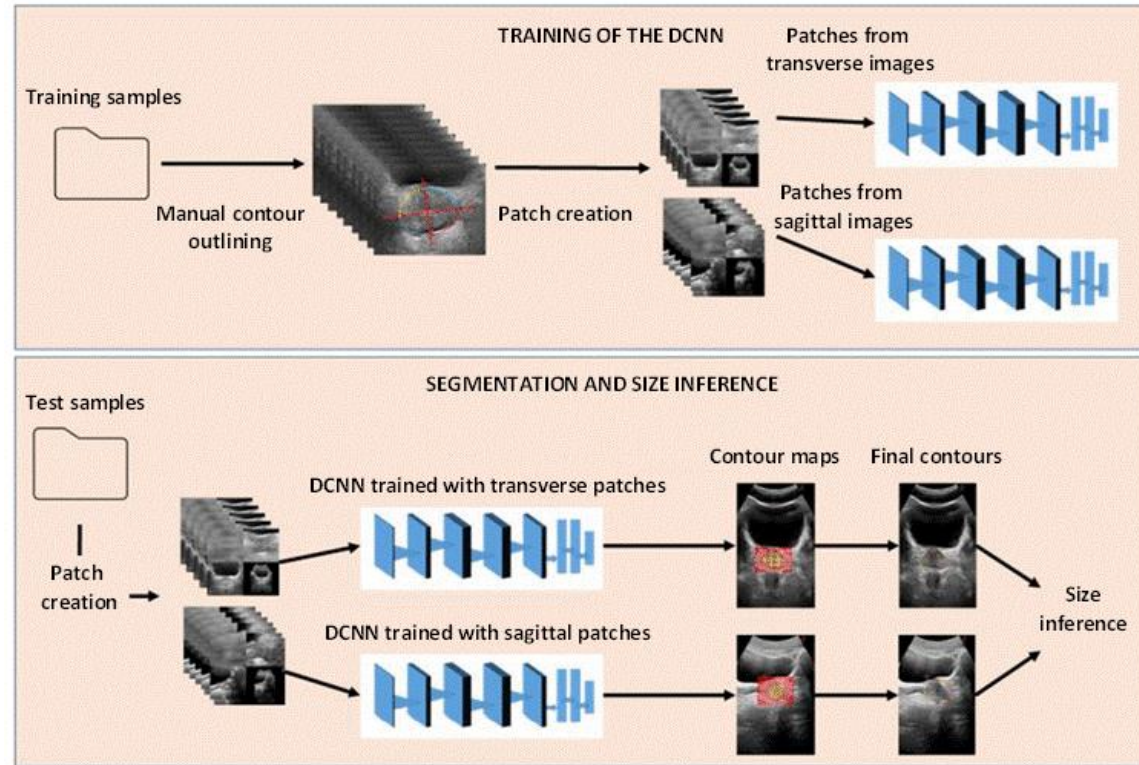


Figure 4.1: The schema of the patch-based segmentation system. Complete contours were obtained from manually annotated four points on the prostate contour. Training image patches were extracted from centering locations on different layers around or inside the contour. A DCNN for each plane was trained. During segmentation, image patches were extracted from several locations in a sliding window manner. Trained DCNNs classified these patches and contour maps were created according to the classification results. Final contours obtained from the contour maps and size inference were done from the final contours.

4.2.1. Segmentation

To train the segmentation system, complete contours corresponding to the training images were needed. Since the training data included expert markings of only the four points which represent the width and the height of the prostate, a method to create complete contours was proposed. Complete contours were created with the help of the expert markings by assuming that the prostate contour is a joint of four different ellipses between the four expert markings.

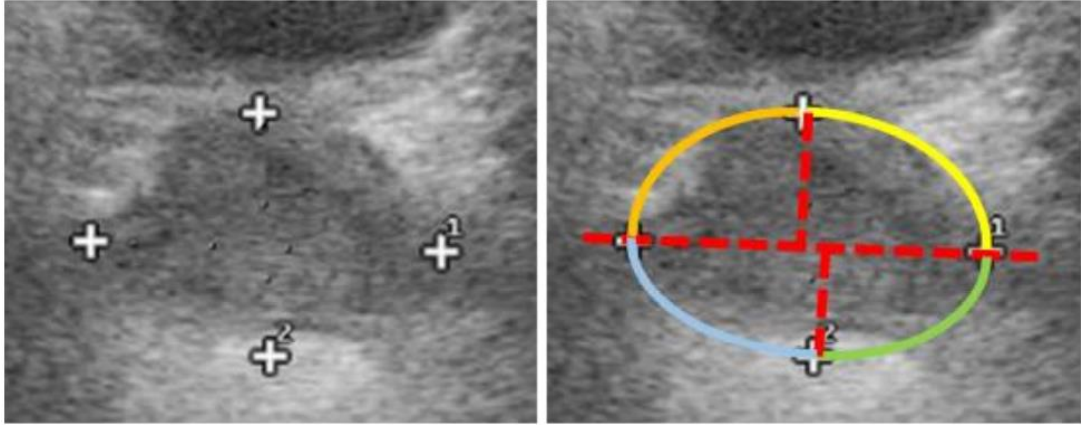


Figure 4.2: Creation of the training contours using expert markings. The horizontal axis of the prostate was assumed as the shared horizontal axis of four different ellipses.

As shown in Figure 4.2, four different ellipse sections were fitted to the four expert markings. The horizontal line between the expert markings of prostate width was assumed as the intersection of the horizontal axes of the four ellipses. Vertical axes of the ellipses were normal to the horizontal axes from the expert markings of prostate height.

A multi-task learning process was proposed which uses image patches to learn the complete contours created from the expert markings. Three different tasks were chosen to learn distance, orientation, and curvature values for each image patch location according to the prostate contour or center.

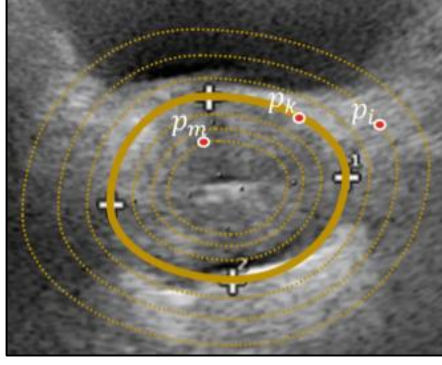


Figure 4.3: Determination of class intervals for the distance task. Different layers were created by scaling the contour down or up. Training data was created using patches centering many different locations on the layers and the distance class labels were determined according to the layer a patch center was on.

The distance task learned the distance of a patch center to the closest contour point. Several concentric contour layers were created by scaling up or down the prostate contour. Distance values were quantized into several classes according to being on any of these layers. Figure 4.3 shows three sample points, p_i , p_k , and p_m on three different layers. The labels of these three points, l_i^1 , l_k^1 , and l_m^1 were determined according to the layer the corresponding point is on. For the given example in Figure 4.3, assuming the innermost layer's label is 1, the respective values for l_i^1 , l_k^1 , and l_m^1 will be 6, 4, and 1.

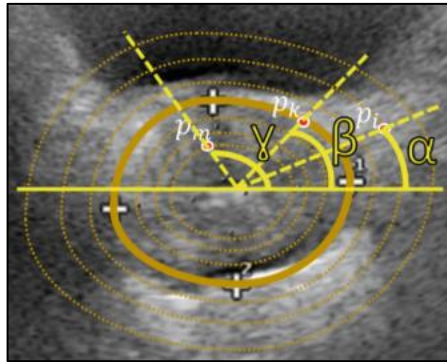


Figure 4.4: Determination of class intervals for the orientation task. 360 degrees were quantized into several classes equally to determine class intervals. The orientation value of a patch center was obtained as the angle between the line segment from the patch center to the prostate center and the horizontal axis.

Given P as the line segment between the prostate center and a patch center point, p_i , the orientation task learned the angle between P and the x-axis. Figure 4.4 shows three sample points, p_i , p_k , and p_m with different angle values, α , β , and γ ,

respectively. Angle values were quantized into several classes for the determination of the corresponding labels l_i^2 , l_k^2 , and l_m^2 .

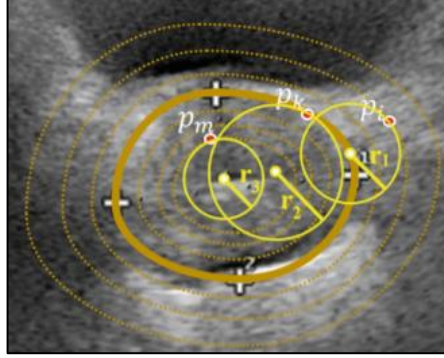


Figure 4.5: Determination of class intervals for the curvature task. The curvature value of a patch center was determined as the radius of the circle on which the patch center and two neighboring locations on the same layer. Curvature values were quantized into several classes according to the training data.

The curvature value of three points was calculated as the circle's radius on which these three points are. The curvature task learned the curvature of a given patch center point. Figure 4.5 shows the curvature values, r_1 , r_2 , and r_3 , of 3 different points, p_i , p_k , and p_m , on three different layers. Two neighbor points on the same layer with the patch center point were used to calculate the curvature value for the patch center. Curvature values, r_1 , r_2 , and r_3 were quantized into several classes to obtain the corresponding orientation task labels, l_i^3 , l_k^3 , and l_m^3 .

We proposed a patch-based design to overcome AUS images' data scarcity and structural problems. Because of the structural problems of AUS images, using single small patches that do not completely contain any structure did not provide enough information about the location of the patch center. Using patches bigger than the prostate provided similar information from different locations, which is redundant. Combining image patches of smaller and larger sizes was a solution to the problems that arise while using image patches of the same size. Thus, this study proposes to use image patches which are concatenations of sub-patches of different sizes. For this purpose, sub-patches with different sizes were created where smaller sizes do not contain the whole prostate, and bigger sizes contain the whole prostate. These sub-patches were then resized to equal size and concatenated. Figure 4.6 shows the patch creation process.

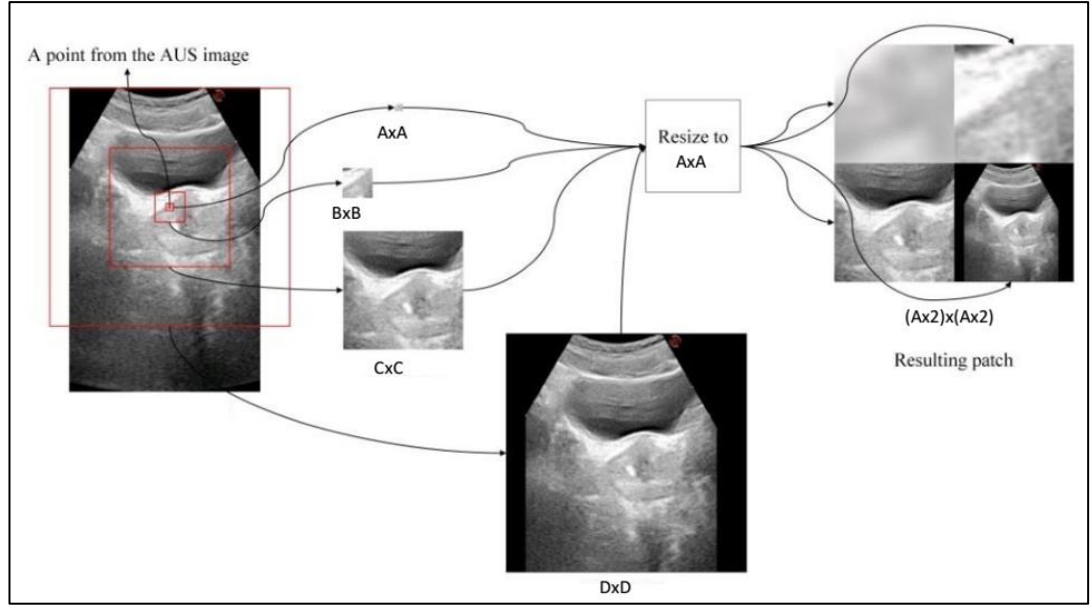


Figure 4.6: The patch creation process. For a given location on an AUS image, patches of different sizes centering the location were created. Each of the patches was resized to the same size and concatenated.

For the training phase, image patches from transverse and sagittal planes centering locations on previously determined layers were extracted, and two different DCNNs were trained for each of the planes. Figure 4.7 shows the creation of k training patches centering several locations on each of m layers and training of DCNN. Here, $P^i = \{p_1^i, \dots, p_k^i\}$ is a group of patches on i^{th} layer and it contains the training data with label i .

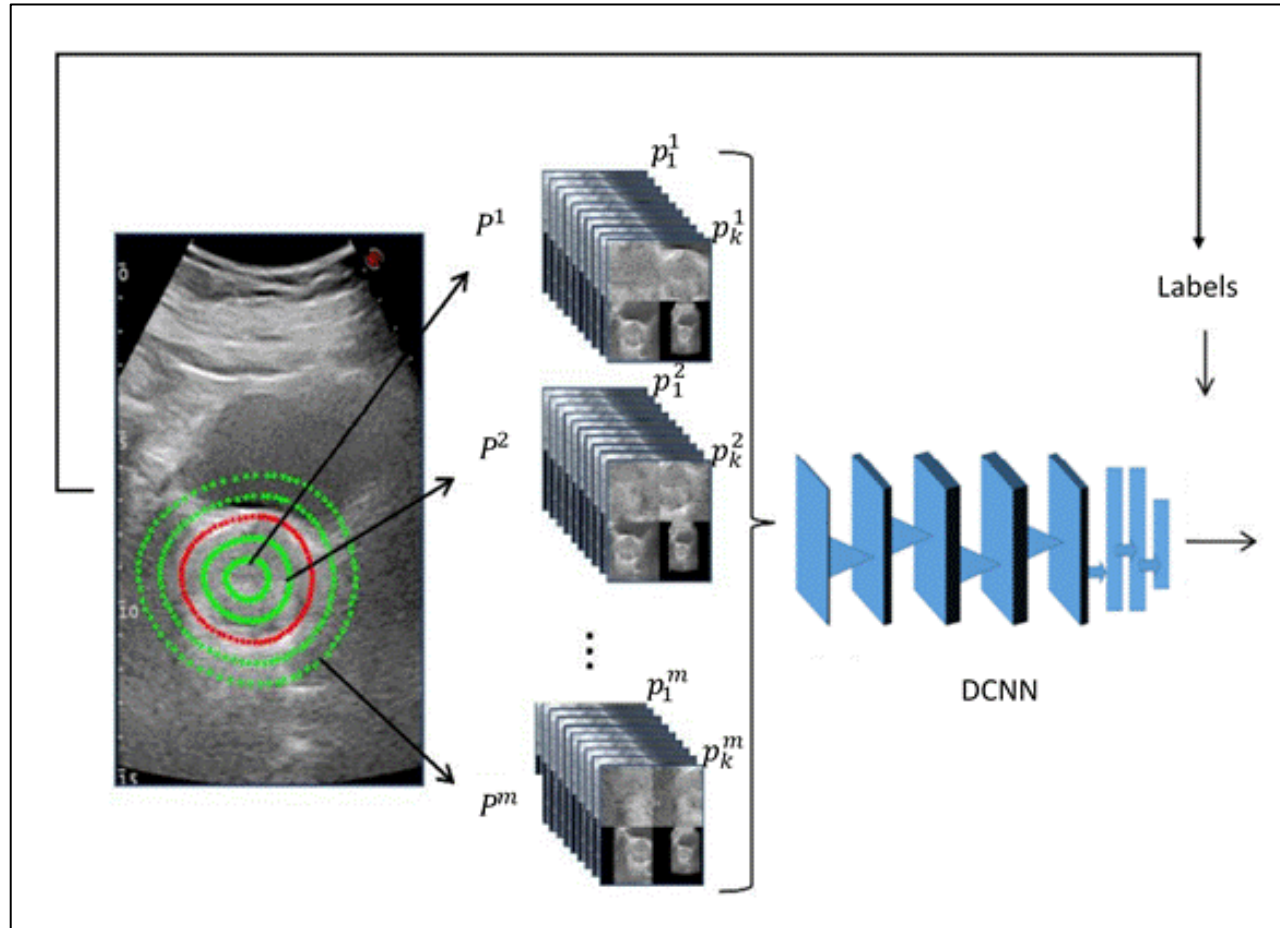


Figure 4.7: Representation of the training phase. Represents extraction of k training patches centering several locations on each of m layers and training of the DCNN.

For the test phase, windows that contain the whole prostate were manually extracted from transverse and sagittal AUS images of the test data. From these windows, patches were created, centering locations in a sliding window manner. Each of the patches was classified by the corresponding trained DCNN, and a contour map was created according to the classification results of the DCNN. Figure 4.8-a shows an example of a contour map on a transverse AUS image.

To obtain an initial contour from the contour map, the locations classified on the contour were used. Several ellipses were fitted to 10 degrees slices of these locations. These ellipse slices were concatenated and formed the initial contour. The final contour was obtained by fitting a new ellipse on the initial contour. The final contour was used to make size inference. Figure 4.8-b shows an example of the initial and the final contours in yellow and red, respectively.

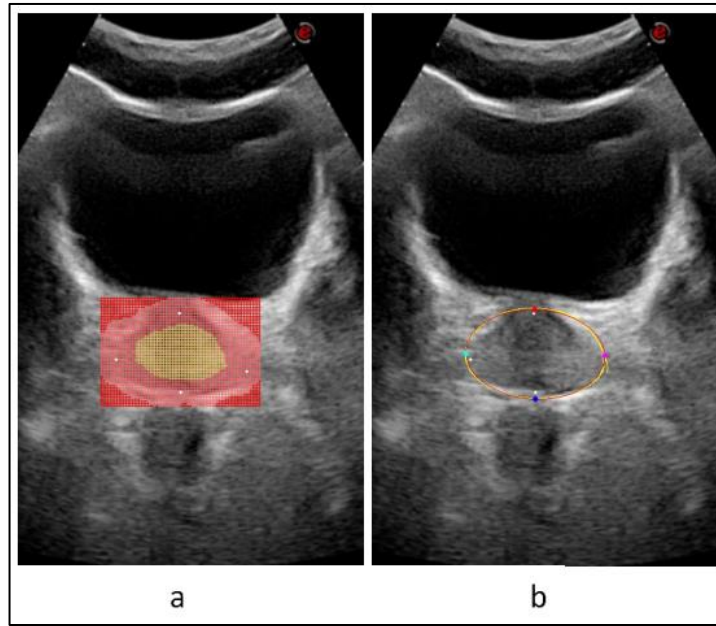


Figure 4.8: Samples of a contour map and a contour. a) A contour map that was created by the classifications of the DCNN. White dots show the expert markings. b) The yellow partite contour is the contour obtained considering the contour map, and the red ellipse is the final contour.

For the learning process, we used Caffe [Jia et al., 2014] implementation of AlexNet [Krizhevsky et al., 2012]. The architecture of AlexNet is shown in Figure 4.9. DCNN optimization was done by Stochastic Gradient Descent. We used the web-based user interface, Digits of NVIDIA, during training, and we used the command line interface of Caffe during the test phase.

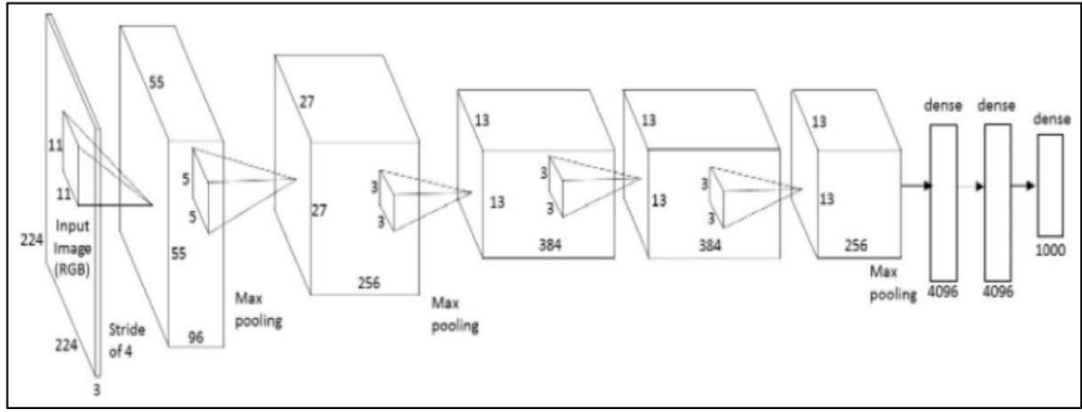


Figure 4.9: Architecture of AlexNet.

4.2.2. Size Inference

The standard ellipsoid formula which was given in (1.1) is commonly used in clinical practice by experts to measure the size of the prostate by using the width, height, and length values of the prostate. The proposed system also used the standard ellipsoid formula given in (1.1) for size inference. For each patient sample with a transverse and a sagittal AUS image, the two final contours which are ellipses from two planes were used to obtain the values of the ellipsoid that represents the prostate. Width and height values were obtained from the final contour of the transverse images, and length values were obtained from the final contour of the sagittal images. To obtain these three values, the endpoints of the axis of the ellipses were used.

4.3. Experiments

90 AUS patient samples with transverse and sagittal images were used for the experiments. Ten-fold cross-validation (TFCV) was applied where each fold has 81 training and 9 test patient samples. In each fold, training data was created for the 81 transverse and 81 sagittal images, and two different DCNNs were trained for the transverse and the sagittal data. Test data was created for the 9 transverse and the 9 sagittal images of the test phase in a similar way.

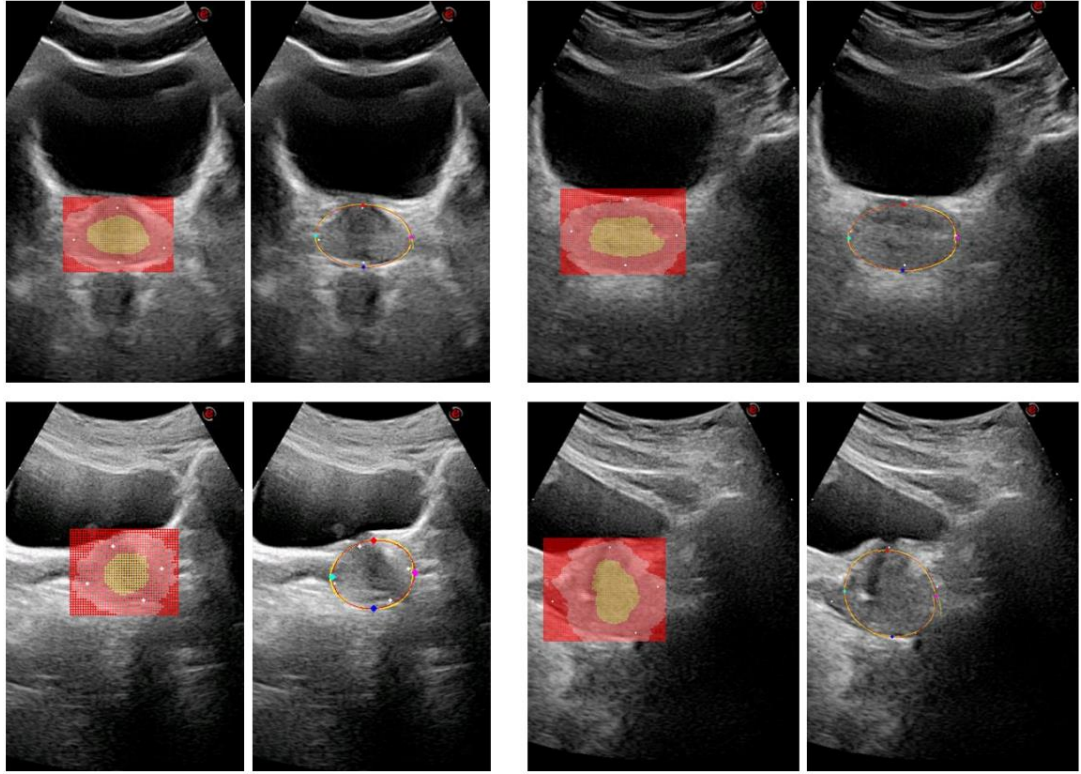


Figure 4.10: Sample segmentation results for transverse and sagittal images. The top row shows contour maps and segmentation results for two transverse images, while the bottom row shows contour maps and segmentation results for two sagittal images. White dots are the expert markings. The yellow partite contour is obtained considering the contour map, and the red ellipse is the final contour.

The test data extraction process was done in a sliding window manner. Since the proposed algorithm aimed not to localize the prostate but infer its size, we used a window smaller than the whole AUS image, which includes the entire prostate, to extract test data. The contour maps shown in Figure 4.10 were created in this way.

We obtained the final contour using the method explained in Section 4.2.1. Sample segmentation results are represented in Figure 4.10 where the top row shows examples from transverse and the bottom row shows examples from sagittal images. Table 4.1 represents the chamfer distance values between expert markings and the location estimations of the proposed system for each fold of the TFCV. The chamfer distance values were obtained according to the four points representing the prostate's height and width on transverse images.

Table 4.1: Chamfer distance values between experts and the system. The chamfer distance values were obtained according to the four points representing the prostate's height and width on transverse images.

Fold	Chamfer Distance for transverse images cm (pixel)	Chamfer Distance for sagittal images cm (pixel)
1	2.22 (8.7)	0.39 (15.7)
2	0.28 (10.1)	0.34 (12.1)
3	0.25 (8.7)	0.45 (15.2)
4	0.26 (10.6)	0.39 (15.8)
5	0.18 (6.1)	0.29 (10.3)
6	0.26 (7.5)	0.47 (13.3)
7	0.28 (10.4)	0.54 (20.1)
8	0.25 (8.4)	0.33 (11.5)
9	0.16 (7.8)	0.50 (24.5)
10	0.20 (7.8)	0.44 (17.7)
Mean	0.23 (8.61)	0.41 (15.62)

Size inference was made according to the volume of the prostate in cm^3 . The standard ellipsoid formula in (1.1) was used for the volume calculation. The difference values between volumes calculated from expert markings on AUS and MR images and contour estimations of the proposed system are given in Table 4.2. Each row shows the mean of the comparison results for the corresponding fold. It is supposed that the difference between folds is due to the data scarcity problem in our experiments.

The ratio of volume estimations of expert markings on AUS images – system, expert markings on MR images – system, and expert markings on AUS images – expert markings on MR images are represented in Table 4.4. Each row shows the average of a fold.

Table 4.2: Difference between volume estimations. Values are of expert markings on AUS images – system, expert markings on MR images – system, and expert markings on AUS images – expert markings on MR images in cm^3 . Each row shows the average of a fold.

Fold	$AUS_{expert} - AUS_{system} (cm^3)$	$MR_{expert} - AUS_{system} (cm^3)$	$AUS_{expert} - MR_{expert} (cm^3)$
1	11.1	16.1	12.6
2	9.7	7.3	7.0
3	8.8	7.7	6.9
4	14.3	15.1	9.7
5	3.6	5.3	4.2
6	8.0	6.1	6.4
7	11.9	12.5	3.7
8	7.5	8.3	4.4
9	9.8	9.2	4.6
10	5.2	6.7	6.9
Mean	8.9	9.4	6.7

Table 4.3: Mean values of the difference. Values are between the volume estimation values of expert markings on AUS images, expert markings on MR images, and system estimations on AUS images. Red represents the highest value while green represents the lower value.

	$AUS_{expert} (cm^3)$	$MR_{expert} (cm^3)$	$AUS_{system} (cm^3)$
$AUS_{expert} (cm^3)$		6.7	8.9
$MR_{expert} (cm^3)$	6.7		9.4
$AUS_{system} (cm^3)$	8.9	9.4	

Table 4.4: Ratio of volume estimations. Values are of expert markings on AUS images – system, expert markings on MR images – system, and expert markings on AUS images – expert markings on MR images. Each row shows the average of a fold.

Fold	$\frac{\text{AUS}_{\text{expert}}}{\text{AUS}_{\text{system}}} (\text{cm}^3)$	$\frac{\text{MR}_{\text{expert}}}{\text{AUS}_{\text{system}}} (\text{cm}^3)$	$\frac{\text{AUS}_{\text{expert}}}{\text{MR}_{\text{expert}}} (\text{cm}^3)$
1	1.44	1.49	1.06
2	1.34	1.01	1.30
3	1.53	1.34	1.12
4	1.31	1.26	1.26
5	0.89	0.91	1.04
6	1.12	0.99	1.13
7	1.68	1.65	1.19
8	1.24	1.21	1.06
9	1.66	1.85	1.03
10	1.09	1.08	1.12
Mean	1.33	1.28	1.13

5. STANDARD ELLIPSOID FORMULA BASED PROSTATE VOLUME ESTIMATION FROM AUS IMAGES BY USING A NOVEL DEEP MODEL

5.1. Motivation and Problem Definition

From surveying literature, one can see that automatic prostate volume estimation is mainly done by using segmentation algorithms. However, in the case of this study where AUS images are used, there are problems such as low SNR values, artifacts, and incomplete contours, making complete contour detection on AUS images challenging for even state-of-the-art techniques. These problems can be tackled by using a more straightforward method similar to manual estimation methods.

Most prostate volume estimation studies present an end-to-end solution that is not compatible with experts' everyday workflow. To trust an automated system's results, clinicians need to know how they are produced [Wang et al., 2021]. It is not easy for everyone to examine the results of an end-to-end system due to its complexness [Murphy et al., 2021]. These systems should be transparent and explainable to experts [Sunarti et al., 2021]. Combining human and artificial intelligence would outperform fully automated systems or humans [Rundo et al., 2020].

Experts use the standard ellipsoid formula given in (1.1) to calculate prostate volume after manually marking transverse and sagittal images. Thus, a system that allows standard ellipsoid-based calculation with manual intervention would benefit clinical applications.

5.2. Proposed Method

In this study, a prostate volume estimation system was proposed that obtains the parameters of the standard ellipsoid formula (1.1), which are the width, the height, and the length of the prostate, and uses it to estimate the volume of the prostate. The system detected four ellipsoid points from the transverse and two ellipsoid points from sagittal images and obtained the parameters. The proposed method allowed both manual intervention and fully automatic use by producing intermediate and final results.

Intermediate results were the diameter endpoints, and the final result was the volume of the prostate. We proposed an image patch voting-based system like our previous work in Section 4, with some fundamental differences in creating patches and learning the model.

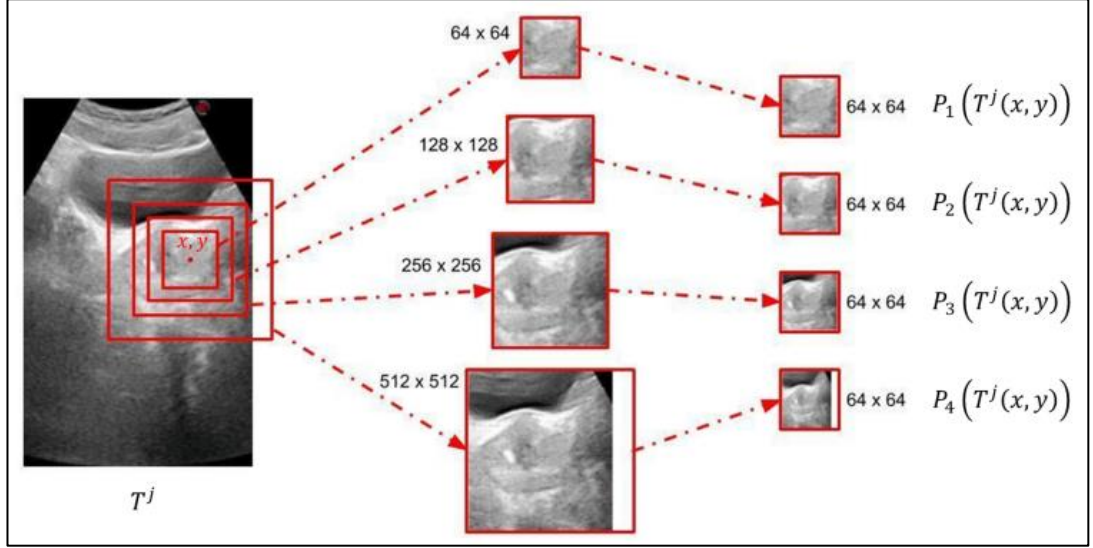


Figure 5.1: Patch extraction process. An example of a given point $T^j(x, y)$ on a given transverse image T^j . System extracts four patches from different scales centering (x, y) . These patches are downsized to the smallest scale and the quadruplet patch, $PQ(T^j(x, y))$ is created.

5.2.1. Patch Extraction

The patch extraction process in this work was different from the patch extraction process of the previous work with the use of sub-patches. The patches extracted from different scales were called quadruplet patches, and instead of being combined, each of them was inputted individually to a model explained in the following section. Patch extraction process is represented in Figure 5.1. The utilization of the patches was also different from the previous system. In this case, each quadruplet patch voted for each of the diameter endpoints of the prostate.

Quadruplet patches of the training data were created, centering randomly chosen locations from a normal distribution around diameter endpoints. For the test data, evenly spaced locations in a sliding window manner were chosen as the centers of the quadruplet patches. Sample patch centers for training and test patches on transverse

and sagittal images are shown in Figure 5.2. Many quadruplet patches were created from each image which means the data was augmented.

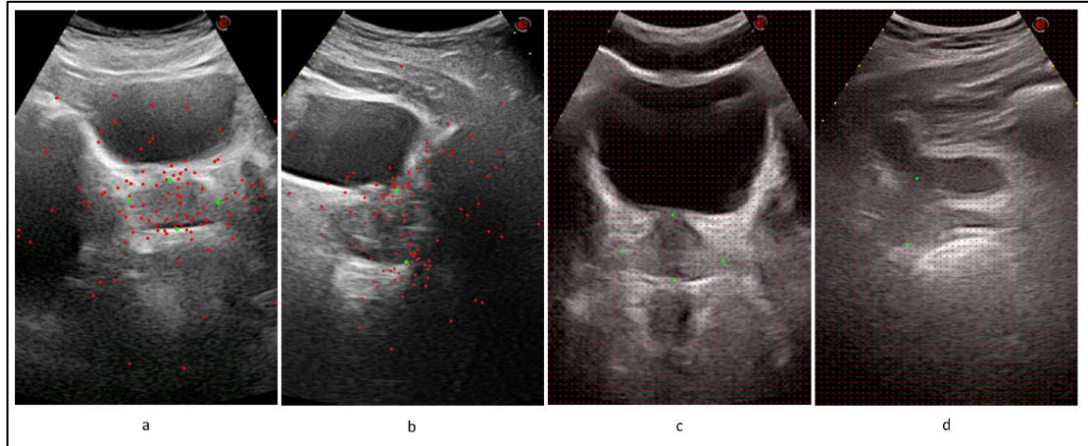


Figure 5.2: Locations chosen for the quadruplet patch centers. Green dots are expert markings, and red dots are patch centers. a) Training patch centers on a transverse AUS image. b) Training patch centers on a sagittal AUS image. c) Test patch centers on a transverse AUS image. d) Test patch centers on a sagittal AUS image.

5.2.2. Quadruplet Network

Several studies in the literature with different types of quadruplet networks have been popular in re-identification or similarity learning studies [Chen et al., 2017]. In these studies, quadruplet networks were trained with four images where two of them were from the same class, and others were from different classes. Pairs in the same class were learned as positive, while pairs in different classes were learned as negative. Greater inter-class and less intra-class variation were achieved by calculating a quadruplet loss while identifying images.

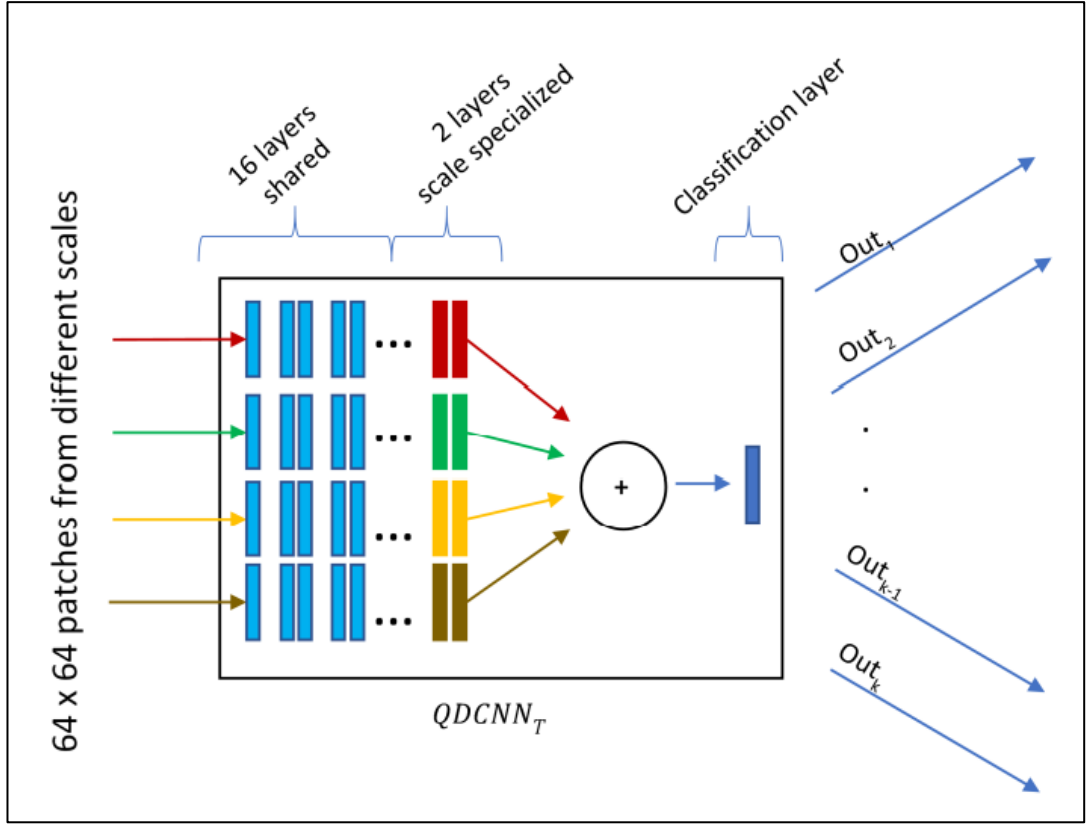


Figure 5.3: The proposed quadruplet model, QDCNN. The model comprises four DCNNs with a joint classifier layer and a joint loss. k is the number of outputs which is 8 for the $QDCNN_T$, which is for transverse images, and 4 for the $QDCNN_S$, which is for sagittal images.

In this study, a novel quadruplet model called Quadruplet Deep Convolutional Neural Network (QDCNN) was proposed, through which quadruplet patches vote for the diameter endpoints [Albayrak and Akgul, 2022]. The model is represented in Figure 5.3. Four ResNet18 DCNNs comprised a QDCNN, with a joint classification layer and a joint loss. Each quadruplet network got each of the corresponding quadruplet patches as input. The QDCNN learned to predict distance and orientation classes relative to each diameter endpoint for a given quadruplet patch center by multi-task classification. Two different QDCNN structures were trained for this purpose. $QDCNN_T$ for the transverse plane had 8 outputs, and $QDCNN_S$ for the sagittal plane had 4 outputs.

Assuming the transverse image, T^j of the j^{th} patient with four diameter endpoints e_1, e_2, e_3 , and e_4 and sagittal image, S^j with two diameter endpoints, e_5 and e_6 , where $e_i \in R^2$. A quadruplet patch $PQ(T^j(x, y))$, centering a given point $T^j(x, y)$ on a transverse image, T^j is composed of four equal-size patches $P_1(T^j(x, y))$,

$P_2(T^j(x, y))$, $P_3(T^j(x, y))$, and $P_4(T^j(x, y))$ of different scales. A quadruplet patch $PQ(S^j(x, y))$, centering a given point $S^j(x, y)$ on a sagittal image, S^j is composed of four equal-size patches $P_1(S^j(x, y))$, $P_2(S^j(x, y))$, $P_3(S^j(x, y))$, and $P_4(S^j(x, y))$ of different scales, similarly.

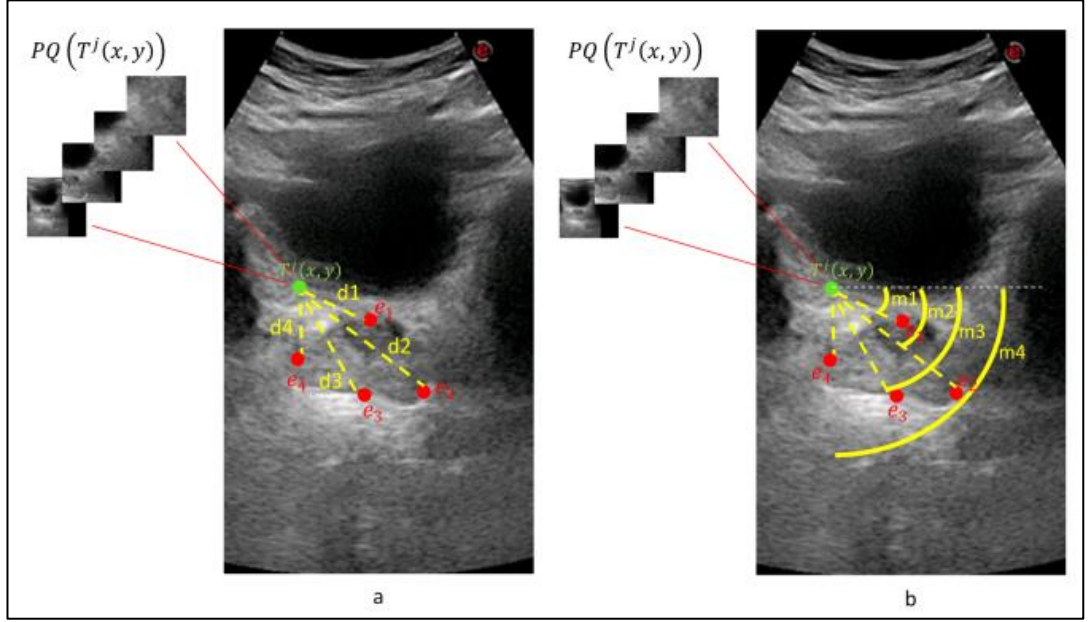


Figure 5.4: Quantization of class label values. For a given point $T^j(x, y)$, on a transverse image T^j where four diameter points (e_1, e_2, e_3, e_4) exist, measurement of distance (a) and orientation (b) class values are shown. Classes or eight different tasks are predicted for this transverse image case.

Two functions, c_d and c_o were defined to obtain distance (cd) and orientation (co) classes, respectively. For a given point $T^j(x, y)$ on a given transverse image T^j , $cd_i^j(x, y) = c_d(T^j(x, y), e_i)$, and $co_i^j(x, y) = c_o(T^j(x, y), e_i)$ where $i = 1, \dots, 4$. Similarly, for a given point $S^j(x, y)$ on a given sagittal image S^j , $cd_i^j(x, y) = c_d(S^j(x, y), e_i)$, and $co_i^j(x, y) = c_o(S^j(x, y), e_i)$ where $i = 5, 6$. That is, QDCNN_T had four classification tasks and QDCNN_S had two classification tasks. Figure 5.4 represents the calculation of the orientation and distance class values.

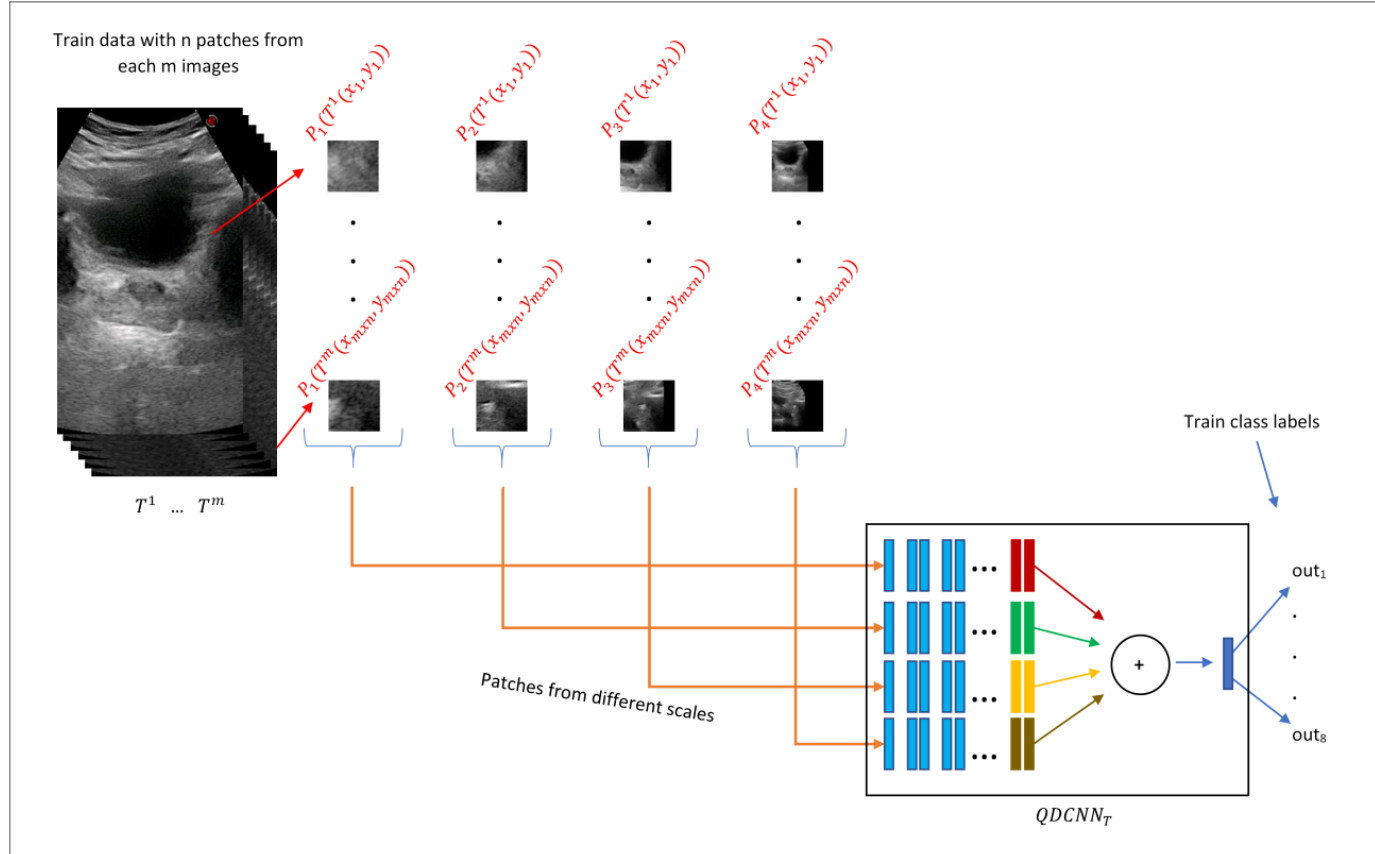


Figure 5.5: Training process with m transverse images. For each of the m transverse images, n quadruplet patches $PQ(T^1(x_1, y_1)), \dots, PQ(T^m(x_{m \times n}, y_{m \times n}))$ are extracted randomly. QDCNN_T is trained to predict eight classes where eight is $number_of_tasks \times number_of_diameter_points$.

Patch extraction for the training phase was done from normally distributed random locations around diameter endpoints of each transverse and sagittal AUS image of training patient samples. During the training of QDCNN_T , n quadruplet patches, $PQ(T^j(x_1, y_1)), \dots, PQ(T^j(x_n, y_n))$ were extracted from each of the m transverse training images, $T^1, \dots, T^m$. These quadruplet patches were then input to the QDCNN_T for the training process. Figure 5.5 represents this process. Training of the QDCNN_S was done in a similar way.

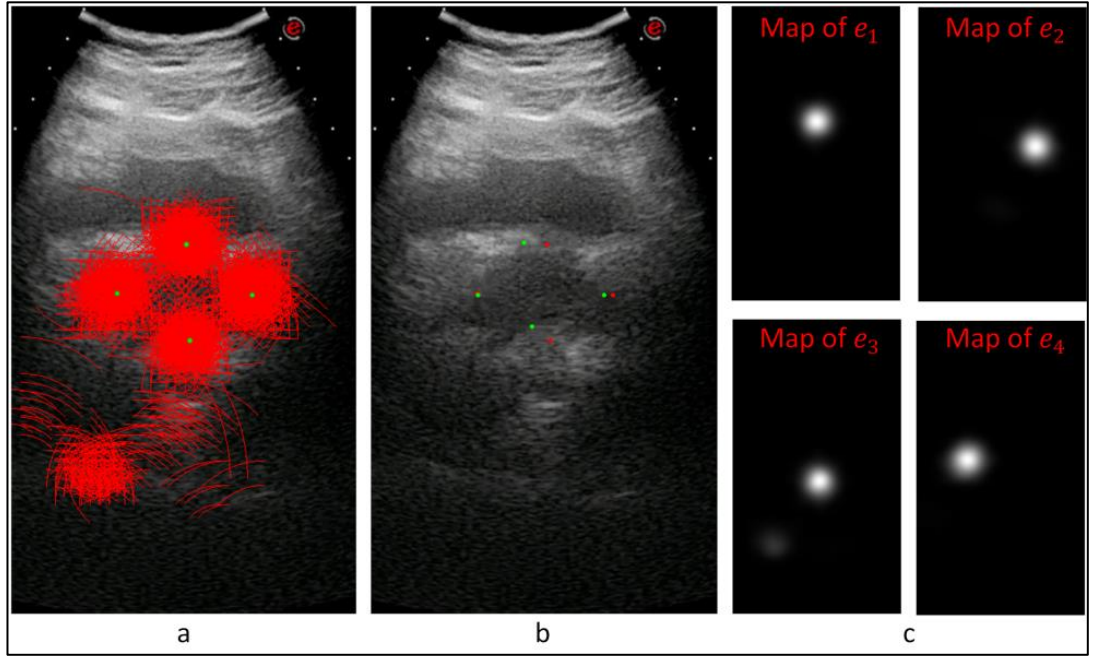


Figure 5.6: Creation of voting maps. Voting arcs (without thickness) are shown for a sample transverse image. b) Red dots are locations of expert markings, and green dots are the detected locations of the diameter endpoints. c) The voting maps for the diameter endpoints, e_1, e_2, e_3, e_4 .

5.2.3. Prostate Volume Inference

Prostate volume inference was made by the image patch voting method. A voting space with the same resolution as the input image was created for this purpose. Each quadruplet patch voted for each diameter endpoint at the voting space. During the voting process, cd and co values were created for each quadruplet patch which goes through QDCNN_T or QDCNN_S depending on whether it was extracted from a transverse or a sagittal AUS image. Voting was done along an arc where the arc center was the patch center. The radius of the arc was given by cd , and the orientation of the

arc was given as co . The arc's thickness and length were given by the median and the range of the distance and orientation class intervals. This voting process ended with a voting map for each of the diameter endpoints of a given AUS image, where the peak was the location estimation of the diameter endpoint.

In Figure 5.6-a, voting arcs are shown red, and the detected diameter endpoints are green. The thickness of the arcs is not represented in this sample to show the arc locations better. Figure 5.6-b shows the detected locations with red dots and the expert markings with green dots. The voting maps for each of the diameter endpoints are shown in Figure 5.6-c. These maps were convolved with a Gaussian smoothing map to suppress the noise.

For a given test patient sample with a transverse image T^j and a sagittal image S^j , quadruplet patches were extracted centering evenly spaced locations in a sliding window manner. Each quadruplet patch was classified by the corresponding QDCNN and by this means votes for the distance and the orientation class values. The voting mechanism is shown in Figure 5.7 on a transverse AUS image T^j . For the voting process of a location (x, y) , the trained QDCNN_T got the quadruplet patch $PQ(T^j(x, y))$ which centers the location and produced eight outputs. These outputs were the $cd - co$ pairs for each of the endpoints. Voting was done for each location, and voting maps for each endpoint were created. The peaks of the voting maps were determined as the final locations of the corresponding endpoints. Endpoint locations for the sagittal image S^j was obtained similarly and the volume was calculated by using the standard ellipsoid formula.

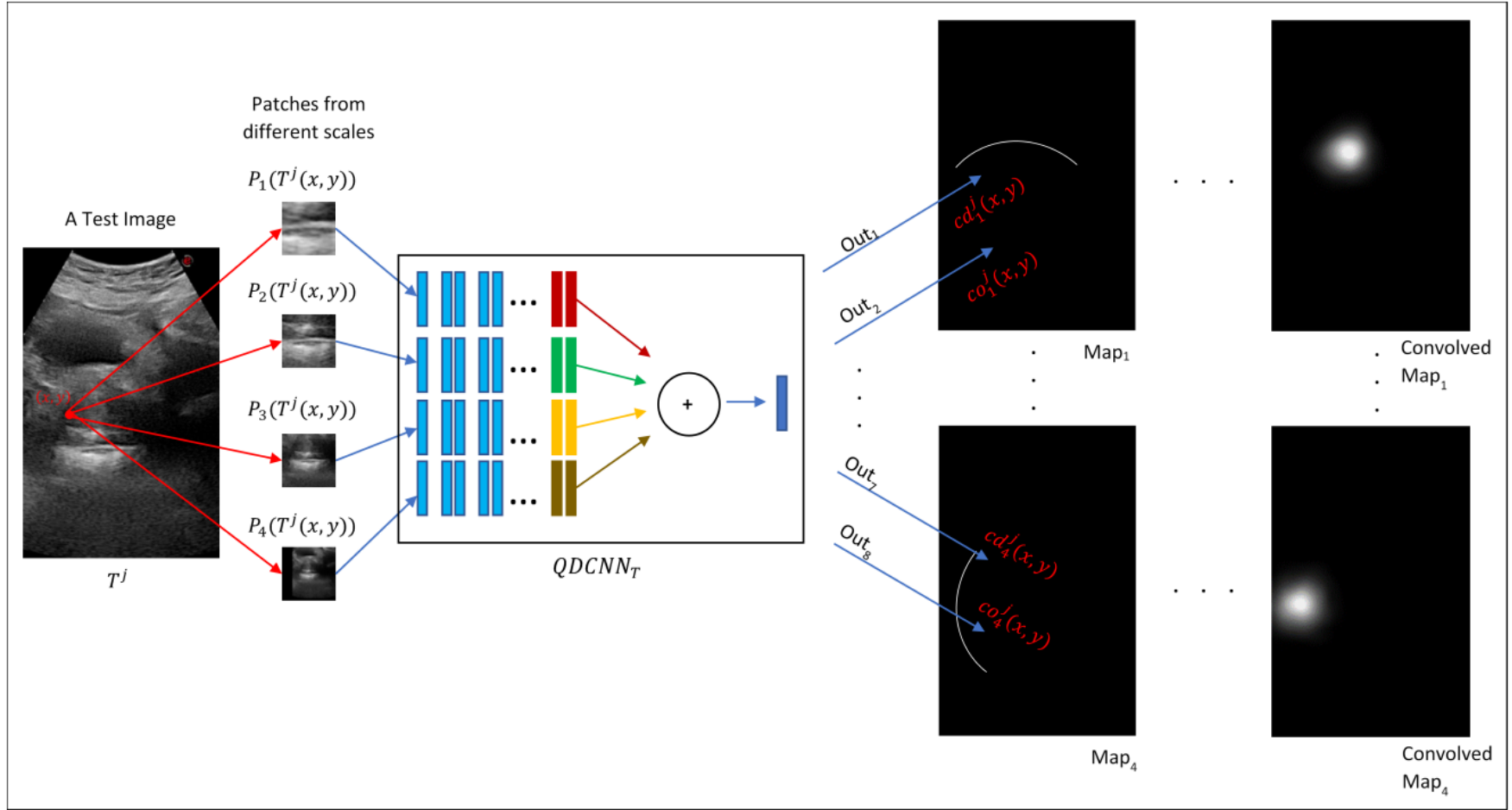


Figure 5.7: Representation of the voting process. Extraction of the patches and the creation of the voting maps for a location $T^j(x, y)$ on a transverse image T^j are represented.

5.3. Experiments

5.3.1. Dataset

The dataset consisted of 305 AUS patient samples, where each sample had a transverse and a sagittal AUS image. 75 of these also had MR images of transverse and sagittal planes. Several experts annotated these AUS and MR images during medical treatments. To train and test the system, we used the AUS annotations. We used the MR annotations as the gold standard.

251 of 305 AUS images and 73 of 75 MR images were annotated by two different experts at two different sessions to compare the proposed system's estimations with expert estimations and obtain intra-expert and inter-expert volume estimation differences. To represent the width, the height, and the length of the ellipsoid, we defined three functions, $D_W(M(T^j))$, $D_H(M(T^j))$, and $D_L(M(S^j))$, respectively. M represents a measure on a transverse (T^j) or a sagittal (S^j) image. To compare the differences, we calculated the mean absolute value difference (MAVD) between two different measurements M_1 and M_2 as given in (5.1). When both measurements are from AUS images, $N = 251$, and when one is from MR images, $N = 73$.

$$MAVD(M_1, M_2) = \frac{1}{N} \sum_{j=1}^N |V_1^j - V_2^j|$$

where

$$V_{k=1,2}^j = V \left(D_W(M_k(T^j)), D_H(M_k(T^j)), D_L(S_k(T^j)) \right) \quad (5.1)$$

and

$$V = width.height.length.\pi/6$$

Intra-expert and inter-expert MAVD values of manual volume estimations from AUS and MR images are shown in the top eight rows of Table 5.1. The average prostate volumes considering the annotations on MR images was 25.9 cm^3 Table 5.2 shows the standard deviation of absolute value differences (SDAVD) values for Table 5.1. The image type ($IT = AUS, MR$), the expert ID ($e = exp1, exp2$), and the

annotation session ID ($k = mark1, mark2$) are represented by the column and the row headings as IT_e^k .

The average intra-expert MAVD value for MR images was $2.81 \left(\frac{2.63+3.00}{2} \right) \text{ cm}^3$, while it was $3.73 \left(\frac{3.46+4.01}{2} \right) \text{ cm}^3$ for AUS images from Table 5.1. Human experts' volume estimations on the same modality images can vary at different marking sessions, even for the MR modality, which is the gold standard. Due to low SNR values and structural problems, a greater intra-expert MAVD value is expected.

Between different modalities, MR and AUS, the average intra-expert MAVD was $7.63 \left(\frac{7.70+7.20+8.67+7.96+7.15+8.05+6.93+7.38}{8} \right) \text{ cm}^3$. We can see that manual annotation on AUS images cause greater MAVD values, considering MR as the gold standard.

Greater values for intra-modality and inter-modality comparisons occurred when inter-expert MAVD values were examined. For MR-MR and AUS-AUS comparisons, average values of $5.03 \left(\frac{4.96+5.48+4.79+4.92}{4} \right)$ and $5.07 \left(\frac{5.19+5.44+4.76+4.89}{4} \right) \text{ cm}^3$ inter-expert MAVD values were obtained, respectively, from Table 5.1. Thus, for both MR and AUS modalities, manual annotations by different experts cause greater MAVD values. One can infer from Table 5.1 that manual annotations of different experts have a high MAVD between AUS and the gold standard, where we obtained an average of $8.06 \left(\frac{6.36+7.12+6.43+7.06+9.53+9.08+9.68+9.26}{8} \right) \text{ cm}^3$.

One can see the difference between different expert volume estimations from the comparisons of estimations from manually marked images. Same experts also marked different positions at different marking sessions. It was inferred from these results that the guidance of the automated system is expected to enhance the experts' consistency and stability of the volume estimation results.

5.3.2. QDCNN Results

The proposed system was evaluated by TFCV. Each image was resampled to a 40 pixel/cm scale using the pixel sizes available from the US device to eliminate scale differences between images.

In Table 5.1, the first row of the second part shows the MAVD values between the QDCNN system and the experts. An average of $4.95 \left(\frac{4.86+4.52+5.08+5.37}{4} \right)$ MAVD

was obtained which was smaller than the inter-expert MAVD (5.09 cm³) when the system's volume estimations were compared with the expert volume estimations on AUS images.

The proposed system's volume estimations ranked among inter-expert comparisons on AUS images. Considering Table 5.2, one can generally see small standard deviation values for small absolute value differences for both manual and automated measurements. Thus, the proposed system's estimations can be considered as stable as the manual estimations.

When the proposed system's volume estimations are compared with expert volume estimations on MR images, to evaluate the system's volume estimations concerning the gold standard, one can obtain $6.22 \left(\frac{7.10+6.55+5.37+5.86}{4} \right)$ cm³ average MAVD, which is less than both the intra-expert (7.63 cm³) and inter-expert (8.06 cm³) average MAVD between different modalities (AUS-MR). It was inferred that, under the guidance of the proposed system, experts could possibly produce more consistent and accurate volume estimations.

A baseline system was implemented that accepted inputs of 300 × 600 pixels AUS images and produced the diameter endpoint locations as output. By this means, the proposed system's performance was compared against traditional deep learning systems. A modified version of the state-of-the-art DenseNet121 was used to produce the endpoint locations. The baseline system had a single-network and non-voting structure in contrast with the proposed system. Data augmentation was done by random cropping during training of the baseline system. Table 5.1's last row compares the proposed system with the baseline system. One can numerically see the proposed system's image-patch voting mechanism's advantage. The average MAVD between the baseline system and experts on MR images was $6.73 \left(\frac{7.92+7.27+5.97+5.77}{4} \right)$ cm³ and the average MAVD between the baseline system and experts on AUS images was $6.25 \left(\frac{6.62+6.13+5.79+6.46}{4} \right)$ cm³. It was concluded that the image-patch voting technique improved the overall results.

Table 5.1: MAVD values. Intra-expert and inter-expert MAVD values in cm^3 for AUS and MR images is given in top part, and it is symmetrical. The MAVD values between the proposed system, the baseline system, and the experts are given in the bottom part. Green cells show the smallest (best), and red cells show the highest MAVD values.

<i>MEAN</i>	<i>AUS^{mark1}_{exp1}</i>	<i>AUS^{mark2}_{exp1}</i>	<i>AUS^{mark1}_{exp2}</i>	<i>AUS^{mark2}_{exp2}</i>	<i>MR^{mark1}_{exp1}</i>	<i>MR^{mark2}_{exp1}</i>	<i>MR^{mark1}_{exp2}</i>	<i>MR^{mark2}_{exp2}</i>
<i>AUS^{mark1}_{exp1}</i>		3.46	5.19	5.44	7.70	7.20	6.36	7.12
<i>AUS^{mark2}_{exp1}</i>	3.46		4.76	4.89	8.67	7.96	6.43	7.06
<i>AUS^{mark1}_{exp2}</i>	5.19	4.76		4.01	9.53	9.08	7.15	8.05
<i>AUS^{mark2}_{exp2}</i>	5.44	4.89	4.01		9.68	9.26	6.93	7.38
<i>MR^{mark1}_{exp1}</i>	7.70	8.67	9.53	9.68		2.63	4.96	5.48
<i>MR^{mark2}_{exp1}</i>	7.20	7.96	9.08	9.26	2.63		4.79	4.92
<i>MR^{mark1}_{exp2}</i>	6.36	6.43	7.15	6.93	4.96	4.79		3.00
<i>MR^{mark2}_{exp2}</i>	7.12	7.06	8.05	7.38	5.48	4.92	3.00	
<i>AUS_{QN}</i>	4.86	4.52	5.08	5.37	7.10	6.55	5.37	5.86
<i>AUS_{BL}</i>	6.62	6.13	5.79	6.46	7.92	7.27	5.97	5.77

Table 5.2: SDAVD values for Table 5.1.

<i>MEAN</i>	<i>AUS</i> ^{mark1} _{exp1}	<i>AUS</i> ^{mark2} _{exp1}	<i>AUS</i> ^{mark1} _{exp2}	<i>AUS</i> ^{mark2} _{exp2}	<i>MR</i> ^{mark1} _{exp1}	<i>MR</i> ^{mark2} _{exp1}	<i>MR</i> ^{mark1} _{exp2}	<i>MR</i> ^{mark2} _{exp2}
<i>AUS</i> ^{mark1} _{exp1}		3.67	5.38	8.35	6.97	6.29	5.32	5.33
<i>AUS</i> ^{mark2} _{exp1}	3.67		5.03	7.70	8.05	7.21	5.88	5.75
<i>AUS</i> ^{mark1} _{exp2}	5.38	5.03		7.56	10.51	9.68	7.49	6.73
<i>AUS</i> ^{mark2} _{exp2}	8.35	7.70	7.56		10.68	9.71	6.92	6.50
<i>MR</i> ^{mark1} _{exp1}	6.97	8.05	10.51	10.68		2.63	6.01	5.92
<i>MR</i> ^{mark2} _{exp1}	6.29	7.21	9.68	9.71	2.63		4.83	5.16
<i>MR</i> ^{mark1} _{exp2}	5.32	5.88	7.49	6.92	6.01	4.83		2.60
<i>MR</i> ^{mark2} _{exp2}	5.33	5.75	6.73	6.50	5.92	5.16	2.60	
<i>AUS</i> _{QN}	5.70	5.05	5.46	8.12	9.15	7.88	6.51	5.25
<i>AUS</i> _{BL}	7.46	7.12	5.98	8.90	11.17	10.07	7.36	7.09

Table 5.3: Ratio values. The average sample-wise volume estimation ratio for inter-expert and intra-expert results from AUS and MR images is given in the top part, and it is symmetrical. The ratio values between the proposed system, the baseline system, and the experts are given in the bottom part. Green cells show the ratio values closest to 1 (best), and red cells show the ratio values farthest to 1.

<i>RATIO</i>	AUS_{exp1}^{mark1}	AUS_{exp1}^{mark2}	AUS_{exp2}^{mark1}	AUS_{exp2}^{mark2}	MR_{exp1}^{mark1}	MR_{exp1}^{mark2}	MR_{exp2}^{mark1}	MR_{exp2}^{mark2}
AUS_{exp1}^{mark1}		1.01	1.12	1.11	1.38	1.39	1.27	1.34
AUS_{exp1}^{mark2}	1.01		1.12	1.11	1.43	1.42	1.30	1.37
AUS_{exp2}^{mark1}	1.12	1.12		1.01	1.53	1.52	1.40	1.46
AUS_{exp2}^{mark2}	1.11	1.11	1.01		1.51	1.50	1.36	1.42
MR_{exp1}^{mark1}	1.38	1.43	1.53	1.51		1.01	1.13	1.08
MR_{exp1}^{mark2}	1.39	1.42	1.52	1.50	1.01		1.12	1.08
MR_{exp2}^{mark1}	1.27	1.30	1.40	1.36	1.13	1.12		0.97
MR_{exp2}^{mark2}	1.34	1.37	1.46	1.42	1.08	1.08	0.97	
AUS_{QN}	1.06	1.05	1.14	1.14	0.85	0.85	0.94	0.90
AUS_{BL}	1.09	1.08	1.16	1.16	0.89	0.89	0.98	0.94

The average sample-wise volume estimation ratio for inter-expert and intra-expert results from AUS and MR images is given in the top part of Table 5.3. The bottom part of the table shows the average sample-wise volume estimation ratio between the proposed system, the baseline system, and experts. Green cells show the ratio values closest to 1 (best), and red cells show the ratio values farthest to 1.

5.3.3. Ablation Study

Three different models and a subset of the dataset were used for the ablation study. Patches with pixel sizes (64×64) and (128×128) was used, and a ResNet-18 was trained for the transverse and another one for the sagittal plane for each of these patch sizes. A twin deep convolutional neural network (TDCNN) was also trained for the transverse and another one for the sagittal plane. A TDCNN consists of two DCNNs and gets two patches with sizes (64×64) and (128×128) pixels similarly to the QDCNN, which gets four patches. The model outputs the *cd* and the *co* values for each endpoint.

Patch sizes that are smaller than the prostate sizes were used in the data set for the ablation study. Thus, one can observe the effect of the quadruplet patch structure, which consists of smaller and larger patches than the prostate.

When the volume estimations of these individual ResNet-18 networks and TDCNN were compared with the volume estimations of experts on AUS images, an average of 13.5 cm³ MAVD was obtained between experts and these models for the 64×64 ResNet-18 network, 11.4 cm³ for the 128×128 ResNet-18 network, and 8.92 cm³ for the TDCNN. MAVD values between two different markings of two experts and five different models are shown in Figure 5.8 as a bar chart. The models of the ablation study are the first three models. The fourth model is the baseline model, and the last model is the model of the proposed system. The proposed system, QDCNN, had the best MAVD values, which showed the effect of the quadruplet patches and the quadruplet model. TDCNN had better MAVD values than single networks, but it cannot achieve the MAVD values of the baseline system. These MAVD values of the TDCNN showed that using patches only smaller than the prostate is not enough to obtain results ranging among experts.

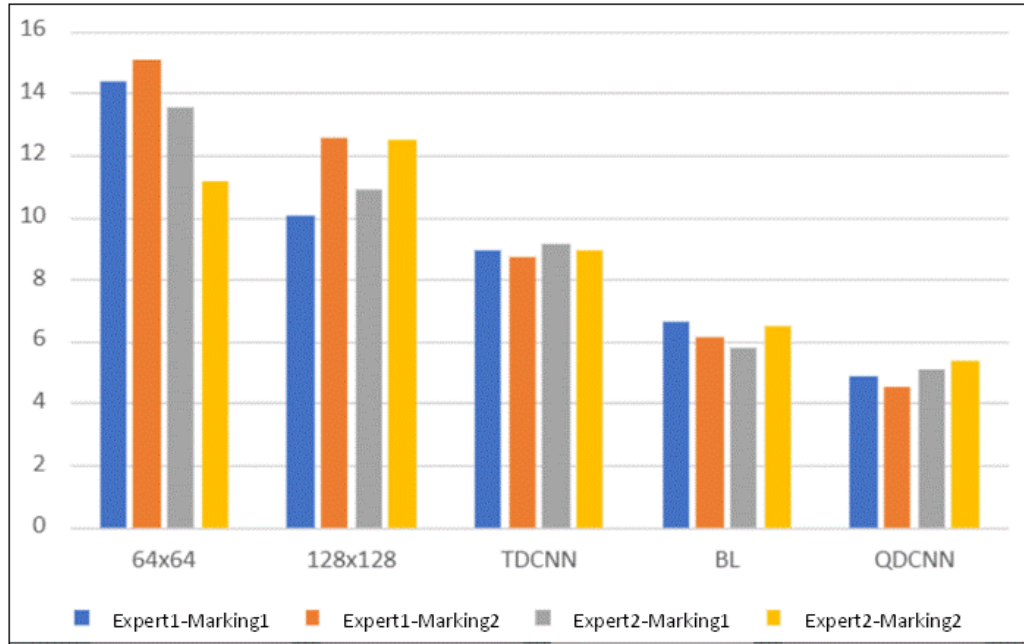


Figure 5.8: MAVD values between experts and models. MAVD values between two expert markings at two different marking sessions and five models are represented. Each group of bars shows the comparison of a model and expert markings at different sessions.

Convolutional neural networks can be used as feature extractors [Jogin et al., 2018]. Good classification results can only come from good features. Thus, deep classifiers can also be used as good feature extractors. To examine the feature extraction ability of the proposed QDCNN, the outputs of the last layer before the classification layer of the network was visualized as a feature vector. t-SNE graphs [Van der Maaten et al., 2008] were used for 2D visualization of the feature vectors of the distance task for the QDCNN, the individual 64×64 ResNet-18, and the individual 128×128 ResNet-18. Two examples of t-SNE graphs for each network are shown in Figure 5.9. Each point represents a feature vector, and its colors represent different distance classes. The color charts show the association between colors and class numbers. Classes with smaller numbers represent short distances, and larger ones represent long distances. As shown in the figure, class values are nicely separated from others and grouped mainly for the small distances for the QDCNN. Single-scale networks cannot achieve the same grouping. These visual results indicate the quality of the features extracted by the proposed network.

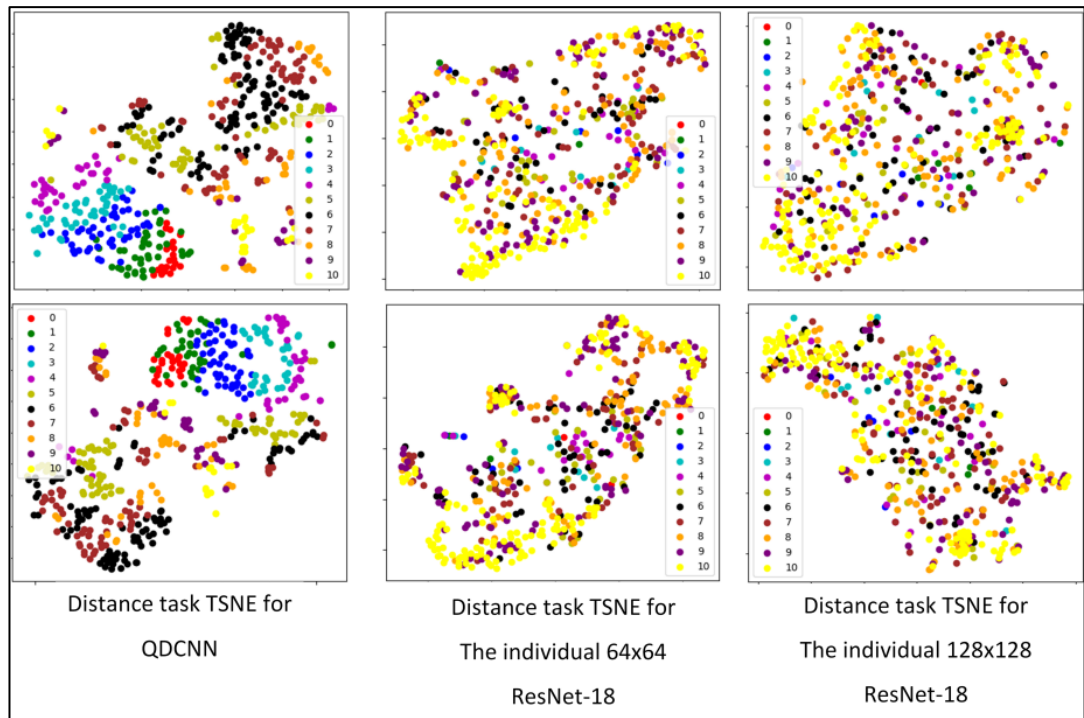


Figure 5.9: t-SNE graph samples. Two examples of t-SNE graphs are represented for each of the QDCNN, the 64×64 ResNet-18, and the 128×128 ResNet-18 networks for the distance task. Each color represents a distance class label and each dot represents a feature vector for the distance task.

6. CONCLUSION AND FUTURE WORK

6.1. Conclusion

According to data from both the world and Turkey, prostate cancer is one of the most common types of cancer in men. As with other types of cancer, early diagnosis of prostate cancer is very important for treating the disease. Ultrasound technology is one of the most frequently used imaging techniques in both the diagnosis and treatment process of prostate cancer. US technology offers two alternatives for prostate imaging: TRUS and AUS. Although the TRUS method produces very reliable results, it is not suitable for early diagnosis or frequent use during treatment because it is uncomfortable for patients. Although the AUS method is a more preferred technique for patients, it has problems in manual and automatic image analysis since the noise ratio of the images produced is higher than the TRUS technique, and the prostate area is small on the images.

This thesis has aimed to propose techniques for prostate volume measurement on AUS images that can provide efficient data acquisition opportunities for patients and specialists in the early diagnosis and treatment of prostate-related diseases like PCa or BPH. Due to the high noise and varying structure of AUS images, the application of known medical image processing techniques to this problem by improving them with original techniques is the creative side of the study.

This thesis study proposed three main methods for the localization of the prostate and prostate volume measurement. The first one is a unique technique that uses DPM to locate the prostate center on the AUS image. The proposed method for finding the prostate center from AUS images is designed to solve the problem that the noise in AUS images is too high and the prostate center and shape vary significantly in the image. Thanks to the customization of the DPM technique for this area, it has been ensured that all features on the image contribute to the prostate region's detection by using the local features of not only the prostate region but also other prominent parts of the image. The experiments performed on the created dataset have shown numerically that quite good improvements are achieved compared to the classical methods that only use local features in feature extraction and then classification.

The main contribution of this method to the literature was the DPM-based design that turns the disadvantages that arise because of the character of images like AUS

images that include many other anatomical structures. Another contribution was the use of a machine learning technique together with the DPM-based design, thus providing a solution to the problems of AUS-like images like low SNR and artifacts, by combining local and object-level information.

The second method developed a deep learning-based contour finding technique to be made around the prostate center obtained in the first stage. Since the mentioned techniques are applied for both transverse and sagittal AUS images of the same patient, the prostate volume is estimated in 3D, and the result obtained can be compared with MR data, which is known as the gold standard. A deep learning-based system, which has been preferred recently, has been used for segmentation and size measurement for prostate size measurement from AUS images. Although deep learning systems have given successful results in many areas, they generally cannot produce good results in the case of direct application in studies involving medical images. The main reason for this is the scarcity of training data, the most basic need for deep learning techniques. Although the dataset created for this study is the largest known so far, it was impossible to use it directly for deep learning because the number of patients was less than thousands. Therefore, a unique hybrid solution has been developed in which deep learning and other techniques are used together. This technique aims to estimate the distance of local particles on the image from the contour instead of directly finding a contour in a given AUS image. Since only image fragments are used instead of the whole image, tens of thousands of training image fragments can be created from a dataset of a few hundred samples. The distance estimations obtained are used for prostate size estimation thanks to a more classical ellipse fitting technique. When the results obtained by using cross-validation techniques are examined thanks to this unique method developed, it is seen that it is close to obtaining a size error difference as much as the marking differences between experts.

The second method contributed to the literature with its hybrid design that gathered a deep learning technique with a patch-based map creation strategy. The method also proposed the extraction of a partite contour from the map as another contribution.

Most radiology measurement systems are designed in an end-to-end fashion. Expert in the loop solutions that are more compatible with experts' normal workflow is complicated with these kinds of designs. To know how the results were produced is important for the experts to trust an automatic system. Examining the results of an end-

to-end system is not easy to understand and explain for many. The operations of these systems should be transparent to experts to be explainable.

The third method proposed an image-patch voting system to automate the commonly used ellipsoid formula-based prostate-volume-estimation method by addressing the explainability issues mentioned above. The method follows the classical prostate volume estimation process outputting intermediate results for manual intervention and final results for automatic use. Experts can see the detected endpoints of the ellipsoid diameters and change the endpoint positions if necessary, providing explainability and confidence in the final measurement results.

The effectiveness of the image-patch voting method was verified against a common baseline model which gets image windows containing the whole prostate and outputs the diameter endpoint location estimations. An ablation study was also done to show the effect of the proposed data creation method and the proposed model.

Since some of the sample patients had both AUS and MR images, the proposed system's AUS volume estimations were compared with the gold standard. By comparing both the proposed system and experts to the gold standard, it was shown that the proposed system's volume estimations fall within the expert estimations. The markings made by two experts in two different marking sessions showed unignorable intra-expert and inter-expert MAVD values in the estimations made on both the same and different modalities. On the other hand, the MAVD values of the proposed system, which were less than inter-expert MAVD values on AUS images and intra- and inter-expert MAVD values on different modalities, indicate the good level of guidance ability of the proposed method. The system can help to enhance expert volume estimations' stability and consistency.

The main contribution of the third method was an expert-in-the-loop technique that works with raw AUS images. Another contribution was the novel use of a QDCNN with a unique voting mechanism.

6.2. Future Work

The methods proposed in this work as solutions to the prostate localization and volume estimation problems on AUS images may be applied to other modalities. Many tasks in medical image analysis need localization as a prerequisite. The variation in anatomical structures and the existence of more than one anatomical structure in a

medical image make the localization process a challenge for automatic methods. The DPM-based method proposed in this work provides a localization technique that also turns the disadvantage of multiple organ existence into an advantage. The image-patch voting-based mechanism with a novel deep model proposed in this work removes the necessity of the localization step as a prerequisite for the detection of any part of any anatomical structure. In future work, these proposed methods might be applied to other modalities.

The proposed methods may be strengthened by implementing their mechanisms to work with more than one AUS slice. Implementing the DPM-based method with more than one slice from different planes would provide 3-dimensional information that would yield a better localization. Using more than one slice for the patch-based deep learning methods proposed in this work would provide more information for the learning models and enhance the results. Using more than one slice would be a good solution to the problems of differentiating different anatomical structures and the existence of artifacts by giving the system the shape and scale information of the anatomic structures from different imaging layers.

Different learning methods might provide solutions to the localization and volume estimation problems on AUS images. The following paragraphs will explain graph convolutional network (GCN) [Kipf and Welling, 2016], deformable convolutional network [Dai et. al., 2017] (DCN), and generative adversarial network [Goodfellow et. al., 2014] (GAN) structures as three sample methods that would provide solutions to the mentioned problems.

GCN is one of the graph neural networks that were proposed to implement neural networks with data of graph-based structures. GCN generalizes the convolution operation from grid data to graph data. As shown in Figure 6.1, in 2D convolution the neighbors of a node are ordered and have a fixed size whereas in graph convolution the neighbors of a node are unordered and variable in size [Wu et. al., 2020].

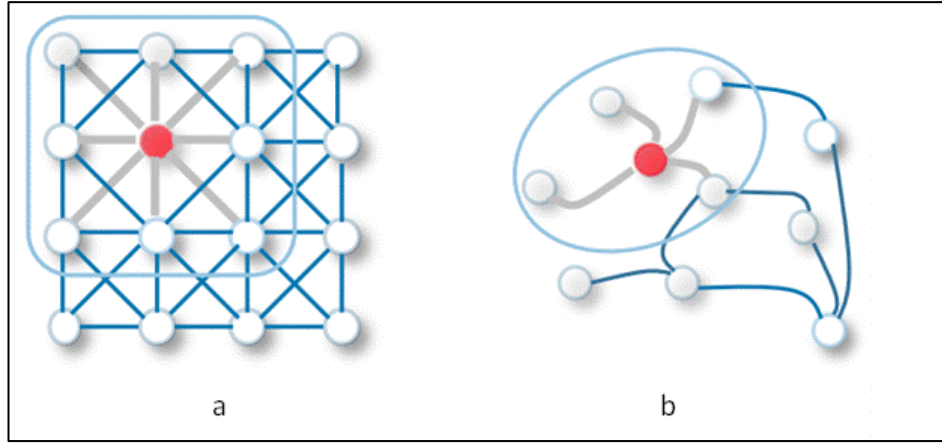


Figure 6.1: Difference between 2D convolution and graph convolution. (a) In 2D convolution, the neighbors of a node are ordered and have a fixed size. (b) The neighbors of a node are unordered and variable in size in graph convolution.

GCN would be a good solution to the localization problem on AUS images and other modalities where more than one anatomical structure exists. GCN provides solutions to many graph-based problems that can be formulated as node classification, graph classification, node clustering, link prediction, and influence maximization. Different anatomical structures, parts of anatomical structures, or both of them on different slices might be represented as a graph. The localization problem might be formulated as one of the mentioned problem settings according to the graph definition chosen. The formulated localization problem then might be solved by using a GCN that utilizes both the representation power of graphs and the learning ability of convolutional networks.

The shape and size variation in anatomical structures on medical images challenge medical image analysis studies. DCN implements two modules that enhance the geometric transformation modeling capability of CNN, namely deformable convolution and deformable RoI pooling. Deformable convolution allows free form deformations of the sampling grid, by adding offsets to the regular grid sampling locations of the standard convolution. RoI pooling similarly adds an offset to each bin position in the regular bin partition of the previous RoI pooling, enabling adaptive part localization for objects with different shapes. Figure 6.2 represents the deformable convolution [Dai et. al., 2017]. DCN would be a good model to learn the anatomical structures on AUS images and other modalities where variations exist that challenges regular CNNs.

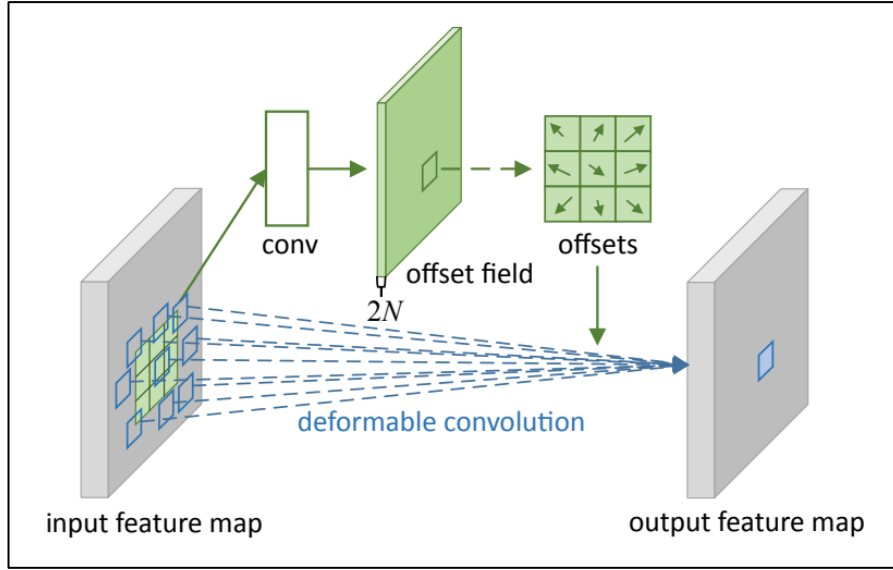


Figure 6.2: Representation of deformable convolution.

GANs are a kind of generative model. A generative model can generate new data instances by modeling the distribution of the data. A GAN model consists of two parts, the generator, and the discriminator. The generator learns to generate fake data that the discriminator can not distinguish from real data. The discriminator learns to detect fake data and penalizes the generator for generating implausible data. The generator and the discriminator incorporate learning to generate and detect fake data, respectively. The representation of the GAN model is given in Figure 6.3. By using GANs the image-patch voting map creation mechanisms in the proposed methods might be enhanced by generating a voting map for each image patch and gathering the patch-based voting maps together by adding vote values.

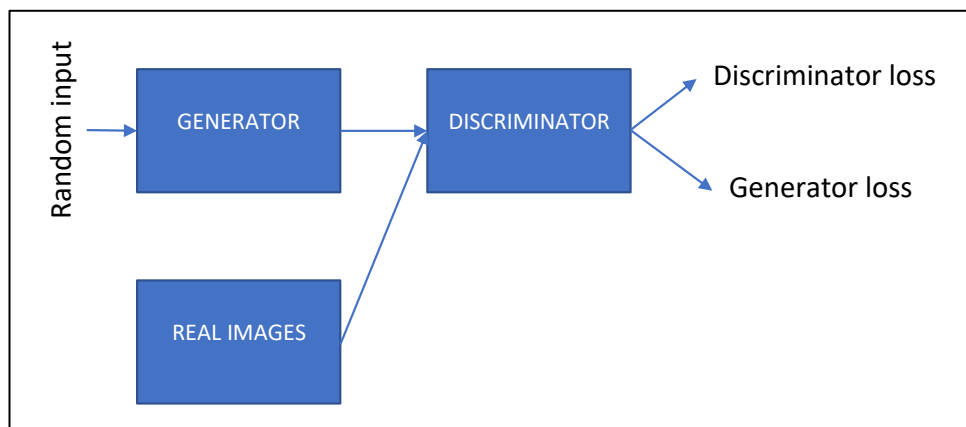


Figure 6.3: Representation of the GAN model.

REFERENCES

- Aarnink R. G., Giesen R. J. B., Huynen A. L., De La Rosette J. J. M. C. H., Debruyne F. M. J., Wijkstra, H., (1994), "A practical clinical method for contour determination in ultrasonographic prostate images.", *Ultrasound in Medicine & Biology*, 20(8), 705-717.
- Aarnink R., Pathak S. D., De La Rosette J. J., Debruyne F. M., Kim Y., Wijkstra H., (1998), "Edge detection in prostatic ultrasound images using integrated edge maps.", *Ultrasonics*, 36(1-5), 635-642.
- Akbari H., Fei B., (2012), "3D ultrasound image segmentation using wavelet support vector machines.", *Medical Physics*, 39(6Part1), 2972-2984.
- Albayrak N. B., Akgul Y. S., (2022), "Estimation of the Prostate Volume from Abdominal Ultrasound Images by Image-Patch Voting.", *Applied Sciences*, 12(3), 1390.
- Albayrak N. B., Oktay A. B., Akgul Y. S., (2015), "Prostate detection from abdominal ultrasound images: A part based approach.", *IEEE International Conference on Image Processing*, 1955-1959, Quebec City, Canada, 27-30 September.
- Albayrak N. B., Oktay A. B., Akgül Y. S., (2017), "Prostate localization from abdominal ultrasound images by using a two-level approach.", *25th Signal Processing and Communications Applications Conference*, Antalya, Türkiye, 1-4, 15-18 May.
- Albayrak N. B., Yildirim E., Akgul Y. S., (2017), "Prostate Size Inference from Abdominal Ultrasound Images with Patch Based Prior Information.", *International Conference on Advanced Concepts for Intelligent Vision Systems*, 249-259, Antwerp, Belgium, 18-21 September.
- Anas E. M. A., Mousavi P., Abolmaesumi P., (2018), "A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy.", *Medical Image Analysis*, 48, 107-116.
- Barry M. J., Roehrborn C. G., (2001), "Benign prostatic hyperplasia." *BMJ*, 323(7320), 1042-1046.
- Betrouni N., Vermandel M., Pasquier D., Maouche S., Rousseau J., (2005), "Segmentation of abdominal ultrasound images of the prostate using a priori information and an adapted noise filter.", *Computerized Medical Imaging and Graphics*, 29(1), 43-51.
- Carneiro G., Nascimento J. C., Freitas A., (2011), "The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods.", *IEEE Transactions on Image Processing*, 21(3), 968-982.

Chan T. F., Vese L. A., (2001), "Active contours without edges.", IEEE Transactions on Image Processing, 10(2), 266-277.

Chen W., Chen X., Zhang J., Huang K., (2017), "Beyond triplet loss: a deep quadruplet network for person re-identification.", IEEE Conference on Computer Vision and Pattern Recognition, 403-412, Honolulu, Hawaii, 21-26 July.

Choi Y. J., Kim J. K., Kim H. J., Cho K. S., (2009), "Interobserver variability of transrectal ultrasound for prostate volume measurement according to volume and observer experience.", American Journal of Roentgenology, 192(2), 444-449.

Chughtai B., Forde J. C., Thomas D. D. M., Laor L., Hossack T., Woo H. H., Te A. E., Kaplan, S. A., (2016), "Benign prostatic hyperplasia.", Nature Reviews Disease Primers, 2(1), 1-15.

Cootes T. F., Hill A., Taylor C. J., Haslam J., (1994), "Use of active shape models for locating structures in medical images.", Image and Vision Computing, 12(6), 355-365.

Cosío, F. A., (2008), "Automatic initialization of an active shape model of the prostate.", Medical Image Analysis, 12(4), 469-483.

Dai J., Qi H., Xiong Y., Li Y., Zhang G., Hu H., Wei Y., (2017), "Deformable convolutional networks.", IEEE International Conference on Computer Vision 764-773, Venice, Italy, 22-29 October.

Dokur Z., Ölmez T., (2002), "Segmentation of ultrasound images by using a hybrid neural network.", Pattern Recognition Letters, 23(14), 1825-1836.

Dunn M. W., (2017), "Prostate cancer screening.", Seminars in Oncology Nursing, 33(2), 156-164.

Fan S., Voon L. K., Sing N. W., (2002), "3D prostate surface detection from ultrasound images based on level set method.", International Conference on Medical Image Computing and Computer-Assisted Intervention, 389-396, Tokyo, Japan, 25-28 September.

Felzenszwalb P. F., Girshick R. B., McAllester D., (2010), "Cascade object detection with deformable part models.", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2241-2248, 13-18 June.

Fischler M. A., Elschlager R. A., (1973), "The representation and matching of pictorial structures.", IEEE Transactions on Computers, 100(1), 67-92.

Foster H. E., Barry M. J., Dahm P., Gandhi M. C., Kaplan S. A., Kohler T. S., Lerner L. B., Lightner D. J., Parsons J. K., Roehrborn C. G., Welliver C., Wilt T. J., McVary K. T., (2018), "Surgical management of lower urinary tract symptoms attributed to benign prostatic hyperplasia: AUA guideline.", The Journal of Urology, 200(3), 612-619.

Garvey B., Türkbey B., Truong H., Bernardo M., Periaswamy S., Choyke P. L., (2014), "Clinical value of prostate segmentation and volume determination on MRI in benign prostatic hyperplasia.", *Diagnostic and Interventional Radiology*, 20(3), 229.

Ghanei A., Soltanian-Zadeh H., Ratkewicz A., Yin F. F., (2001), "A three-dimensional deformable model for segmentation of human prostate from ultrasound images.", *Medical Physics*, 28(10), 2147-2153.

Ghavami N., Hu Y., Bonmati E., Rodell R., Gibson E., Moore C., Barratt D., (2018), "Automatic slice segmentation of intraoperative transrectal ultrasound images using convolutional neural networks.", *Medical Imaging*, 10576, 1057603.

Ghose S., Oliver A., Martí R., Lladó X., Vilanova J. C., Freixenet J., Mitra J., Sidibe D., Meriaudeau, F., (2012), "A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images.", *Computer Methods and Programs in Biomedicine*, 108(1), 262-287.

Ghose S., Oliver A., Mitra J., Martí R., Lladó X., Freixenet J., Sidibe D., Vilanova J. C., Comet J., Meriaudeau F., (2013), "A supervised learning framework of statistical shape and probability priors for automatic prostate segmentation in ultrasound images.", *Medical Image Analysis*, 17(6), 587-600.

GLOBOCAN 2020, (2020), <https://gco.iarc.fr/>, (Date of access: 28.12.2021)

Gong L., Pathak S. D., Haynor D. R., Cho P. S., Kim Y., (2004), "Parametric shape modeling using deformable superellipses for prostate segmentation." *IEEE Transactions on Medical Imaging*, 23(3), 340-349.

Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., (2014), "Generative adversarial nets.", *Advances in Neural Information Processing Systems*, 27.

HSGM, (2017), https://hsgm.saglik.gov.tr/depo/birimler/kanser-db/istatistik/Turkiye_Kanser_Istatistikleri_2017.pdf, (Date of access: 28.12.2021)

Jia Y., Shelhamer E., Donahue J., Karayev S., Long J., Girshick R., Guadarrama S., Darrell T., (2014), "Caffe: Convolutional architecture for fast feature embedding.", *22nd ACM International Conference on Multimedia* 675-678, Orlando, Florida, USA, 3-7 November.

Jogin M., Madhulika M. S., Divya G. D., Meghana R. K., Apoorva S., (2018), "Feature extraction using convolution neural networks (CNN) and deep learning.", *3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology*, 2319-2323, Bengaluru, Karnataka, India, 18-19 May.

Kachouie N. N., Fieguth P., Rahnamayan S., (2006), "An elliptical level set method for automatic TRUS prostate image segmentation.", *IEEE International Symposium on Signal Processing and Information Technology*, 191-196, Vancouver, BC, Canada, 27-30 August.

Karimi D., Zeng Q., Mathur P., Avinash A., Mahdavi S., Spadinger I., Abolmaesumi P., Salcudean S. E., (2019), “Accurate and robust deep learning-based segmentation of the prostate clinical target volume in ultrasound images.”, *Medical Image Analysis*, 57, 186-196.

Kass M., Witkin A., Terzopoulos D., (1988), “Snakes: Active contour models.”, *International Journal of Computer Vision*, 1(4), 321-331.

Kipf T. N., Welling M., (2016), “Semi-supervised classification with graph convolutional networks.”, *arXiv preprint arXiv:1609.02907*.

Knoll C., Alcañiz M., Grau V., Monserrat C., Juan M. C., (1999), “Outlining of the prostate using snakes with shape restrictions based on the wavelet transform”, *Pattern Recognition*, 32(10), 1767-1781.

Krizhevsky A., Sutskever I., Hinton G. E., (2012), “Imagenet lassification with deep convolutional neural networks.”, *Advances in Neural Information Processing Systems*, 25, 1097-1105.

Kwoh C. K., Teo M. Y., Ng W. S., Tan S. N., Jones L. M., (1998), “Outlining the prostate boundary using the harmonics method.”, *Medical and Biological Engineering and Computing*, 36(6), 768-771.

Ladak H. M., Mao F., Wang Y., Downey D. B., Steinman D. A., Fenster A., (2000), “Prostate boundary segmentation from 2D ultrasound images.”, *Medical Physics*, 27(8), 1777-1788.

Lei Y., Tian S., He X., Wang T., Wang B., Patel P., Jani A. B., Mao H., Curran W. J., Liu T., Yang, X., (2019), “Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net.”, *Medical Physics*, 46(7), 3194-3206.

Ladak H. M., Mao F., Wang Y., Downey D. B., Steinman D. A., Fenster A., (2000), “Prostate segmentation from 2D ultrasound images.”, *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4, 3188-3191.

Liu D., Shoag J. E., Poliak D., Goueli R. S., Ravikumar V., Redmond D., Vosoughi A., Fontugne J., Pan H., Lee D., Thomas D., Salari K., Wang Z., Romanel A., Te A., Lee R., Chughtai B., Olumi A. F., Mosquera J. M., Demichelis F., Elemento O., Rubin M. A., Sboner A., Barbieri C. E., (2020), “Integrative multiplatform molecular profiling of benign prostatic hyperplasia identifies distinct subtypes.”, *Nature Communications*, 11(1), 1-9.

Liu Y. J., Ng W. S., Teo M. Y., Lim H. C., (1997), “Computerized prostate boundary estimation of ultrasound images using radial bas-relief method.”, *Medical and Biological Engineering and Computing*, 35(5), 445-454.

Marquardt D. W., (1963), “An algorithm for least-squares estimation of nonlinear parameters.”, *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431-441.

Murphy K., Ruggiero E. D., Upshur R., Willison D. J., Malhotra N., Cai J. C., Malhotra N., Lui V., Gibson J., (2021), "Artificial intelligence for good health: a scoping review of the ethics literature.", *BMC Medical Ethics*, 22(1), 1-17.

Nanayakkara N. D., Samarabandu J., Fenster A., (2006), "Prostate segmentation by feature enhancement using domain knowledge and adaptive region based operations.", *Physics in Medicine & Biology*, 51(7), 1831.

Osher S., Sethian J. A., (1988), "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations.", *Journal of Computational Physics*, 79(1), 12-49.

Pathak S. D., Haynor D. R., Kim Y., (2000), "Edge-guided boundary delineation in prostate ultrasound images.", *IEEE Transactions on Medical Imaging*, 19(12), 1211-1219.

Ronneberger O., Fischer P., Brox T., (2015), "U-net: Convolutional networks for biomedical image segmentation.", *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234-241, Munich, Germany, 5-9 October.

Rundo L., Pirrone R., Vitabile S., Sala E., Gambino O., (2020), "Recent advances of HCI in decision-making tasks for optimized clinical workflows and precision medicine.", *Journal of Biomedical Informatics*, 108, 103479.

Shen D., Zhan Y., Davatzikos C., (2003), "Segmentation of prostate boundaries from ultrasound images using statistical shape model.", *IEEE Transactions on Medical Imaging*, 22(4), 539-551.

Sunarti S., Rahman F. F., Naufal M., Risky M., Febriyanto K., Masnina R., (2021), "Artificial intelligence in healthcare: opportunities and risk for future.", *Gaceta Sanitaria*, 35, 67-70.

Thorpe A., Neal D., (2003), "Benign prostatic hyperplasia.", *The Lancet*, 361(9366), 1359-1367.

Turkbey B., Pinto P. A., Choyke P. L., (2009), "Imaging techniques for prostate cancer: implications for focal therapy.", *Nature Reviews Urology*, 6(4), 191-203.

Tyloch J. F., Wieczorek A. P., (2016), "The standards of an ultrasound examination of the prostate gland. Part 1.", *Journal of Ultrasonography*, 16(67), 378.

Van der Maaten L., Hinton G., (2008), "Visualizing data using t-SNE.", *Journal of Machine Learning Research*, 9(11), 2579-2605.

Viola P., Jones M., (2001), "Robust real-time object detection.", *International Journal of Computer Cision*, 4(34-47), 4.

Wang D., Wang L., Zhang Z., Wang D., Zhu H., Gao Y., Fan X., Tian F., (2021), "Brilliant AI Doctor in Rural Clinics: Challenges in AI-Powered Clinical Decision

Support System Deployment.”, CHI Conference on Human Factors in Computing Systems, 1-18, Yokohama, Japan, 8-13 May.

Wang Y., Cardinal H. N., Downey D. B., Fenster A., (2003), “Semiautomatic three-dimensional segmentation of the prostate using two-dimensional ultrasound images.”, *Medical Physics*, 30(5), 887-897.

Wang Y., Dou H., Hu X., Zhu L., Yang X., Xu M., Qin J., Heng P., Wang T., Ni, D., (2019), “Deep attentive features for prostate segmentation in 3D transrectal ultrasound.”, *IEEE Transactions on Medical Imaging*, 38(12), 2768-2778.

Wasserman N. F., Niendorf E., Spilseth B., (2020), “Measurement of prostate volume with MRI (a guide for the perplexed): bproximate method with analysis of precision and accuracy.”, *Scientific Reports*, 10(1), 1-12.

Wasserman N. F., Spilseth B., Golzarian J., Metzger G. J., (2015), “Use of MRI for lobar classification of benign prostatic hyperplasia: potential phenotypic biomarkers for research on treatment strategies.”, *American Journal of Roentgenology*, 205(3), 564.

Wu Z., Pan S., Chen F., Long G., Zhang C., Philip S. Y., (2020), “A comprehensive survey on graph neural networks.” *IEEE Transactions on Neural Networks and Learning Systems*, 32(1), 4-24.

Yamashiro J. R., de Riese W. T., (2021), “Any correlation between prostate volume and incidence of prostate cancer: a review of reported data for the last thirty years.”, *Research and Reports in Urology*, 13, 749.

Yan P., Xu S., Turkbey B., Kruecker J., (2010), “Discrete deformable model guided by partial active shape model for TRUS image segmentation.”, *IEEE Transactions on Biomedical Engineering*, 57(5), 1158-1166.

Yang F., Suri J., Fenster A., (2006), “Segmentation of prostate from 3-D ultrasound volumes using shape and intensity priors in level set framework.”, *International Conference of the IEEE Engineering in Medicine and Biology Society*, 2341-2344, Orlando, FL, USA, 16-20 August.

Yang X., Rossi P. J., Jani A. B., Mao H., Curran W. J., Liu T., (2016), “3D transrectal ultrasound (TRUS) prostate segmentation based on optimal feature learning framework.”, *Medical Imaging 2016: Image Processing*, 654-660, San Diego, California, USA, 27 Feb-3 March.

Yang X., Yu L., Wu L., Wang Y., Ni D., Qin J., Heng P. A., (2017), “Fine-grained recurrent neural networks for automatic prostate segmentation in ultrasound images.”, *AAAI Conference on Artificial Intelligence*, 31(1),

Yu H. J., Lai M. K., (1998), “The usefulness of prostate-specific antigen (PSA) density in patients with intermediate serum PSA level in a country with low incidence of prostate cancer.”, *Urology*, 51(5), 125-130.

Zaim A., (2005), “Automatic segmentation of the prostate from ultrasound data using feature-based self organizing map.” Scandinavian Conference on Image Analysis, 1259-1265, Joensuu, Finland, 19-22 June.

Zaim A., Jankun J., (2007), “An energy-based segmentation of prostate from ultrasound images using dot-pattern select cells.”, IEEE International Conference on Acoustics, Speech and Signal Processing, 297-300, Honolulu Hawaii, 15-20 April.

Zaim A., Yi T., Keck R., (2007), “Feature-based classification of prostate ultrasound images using multiwavelet and kernel support vector machines.”, International Joint Conference on Neural Networks, 278-281, Orlando, Florida, USA, 12-17 August.

Zhan Y., Shen D., (2003), “Automated segmentation of 3D US prostate images using statistical texture-based matching method.”, International Conference on Medical Image Computing and Computer-Assisted Intervention, 688-696, Montreal, Canada, 15-18 November.

Zhan Y., Shen D., (2006), “Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method.”, IEEE Transactions on Medical Imaging, 25(3), 256-272.

Zouqi M., Samarabandu J., (2008), “Prostate segmentation from 2-D ultrasound images using graph cuts and domain knowledge.”, Canadian Conference on Computer and Robot Vision, 359-362, Windsor, ON, 28-30 May.

BIOGRAPHY

Nur Banu Albayrak graduated from the Department of Computer Engineering of Kocaeli University in 2008. She received her M.S. degree from Fatih University in 2011. She started her Ph.D. at Gebze Technical University in 2012. During her Ph.D. study, she worked on medical image analysis and developed image analysis methods on abdominal ultrasound images. Her research interests include medical image analysis, computer vision, and machine learning.

APPENDIX

Appendix A: Publications Based on this Thesis

Albayrak N. B., Akgul Y. S. (2022), “Estimation of the Prostate Volume from Abdominal Ultrasound Images by Image-Patch Voting.”, *Applied Sciences*, 12(3), 1390.

Albayrak N. B., Oktay A. B., Akgul Y. S., (2015), “Prostate detection from abdominal ultrasound images: A part based approach.”, *IEEE International Conference on Image Processing*, 1955-1959, Quebec City, Canada, 27-30 September.

Albayrak N. B., Oktay A. B., Akgül Y. S., (2017), “Prostate localization from abdominal ultrasound images by using a two-level approach.”, *25th Signal Processing and Communications Applications Conference*, Antalya, Türkiye, 1-4, 15-18 May.

Albayrak N. B., Yildirim E., Akgul Y. S., (2017), “Prostate Size Inference from Abdominal Ultrasound Images with Patch Based Prior Information.”, *International Conference on Advanced Concepts for Intelligent Vision Systems*, 249-259, Antwerp, Belgium, 18-21 September.