

TECHNIQUES FOR DEVELOPMENT OF ULTRASOUND GUIDED FAST, MULTI-MODAL LIVER RESECTION METHOD

DEEPAK MISHRA



DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

AUGUST 2019

© Indian Institute of Technology Delhi (IITD), New Delhi, 2019

TECHNIQUES FOR DEVELOPMENT OF ULTRASOUND GUIDED FAST, MULTI-MODAL LIVER RESECTION METHOD

by

DEEPAK MISHRA

DEPARTMENT OF ELECTRICAL ENGINEERING

Submitted

in the fulfillment of the requirement of the Degree of Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

AUGUST 2019

To my loving parents ...

Certificate

This is to certify that the thesis entitled “**Techniques for Development of Ultrasound Guided Fast, Multi-modal Liver Resection Method**”, being submitted by **Deepak Mishra** for the award of the degree of **Doctor of Philosophy** to the Department of Electrical Engineering, Indian Institute of Technology Delhi, is a record of bonafide work carried out by him under my guidance and supervision. The matter embodied in this thesis has not been submitted in part or full, to any other University or Institute for the award of any degree or diploma.

Mukul Sarkar

Associate Professor,
Department of Electrical Engineering,
Indian Institute of Technology Delhi,
Hauz Khas, New Delhi - 110016,
INDIA.

Santanu Chaudhury

Director,
Indian Institute of Technology Jodhpur,
Professor,
Department of Electrical Engineering,
Indian Institute of Technology Delhi,
Hauz Khas, New Delhi - 110016,
INDIA.

Arvinder Singh Soin

Chairman,
Institute of Liver Transplantation and Regenerative Medicine,
Medanta - The Medicity,
Gurgaon - 122006,
INDIA.

Acknowledgments

I would like to express my sincere gratitude to my supervisors Dr. Mukul Sarkar, Prof. Santanu Chaudhury, and Dr. Arvinder Singh Soin for their support and guidance, and also like to thank for providing me an opportunity to work with them. My special thanks to Prof. Brejesh Lall, Prof. Indra Narayan Kar, and Prof. Sudipto Mukherjee, members of my research committee and Dr. Pratosh A.P. for all their support and motivation.

My heartfelt gratitude to my parents for having faith in me and patience to support me on an extended journey. My special thanks to Ms. Amandeep Kaur for her consistent support and help. This journey would have not been possible without these people.

I express my gratitude to medical personnel of Medanta Hospital, Gurgaon, India, especially Dr. Vivek Sharma, Dr. Prashant Bhangui, and Mr. Dilip, for the help they provided in form of discussions, data acquisition, and clinical understanding of the problem.

I thank my colleagues from Multimedia and Impact laboratories, Department of Electrical Engineering, IIT Delhi, in particular Dr. Gajendranath Chaudhury, Bhuvan B., Swati Bhugra, and Supriya Chakraborty, for their continuous support during the work. I acknowledge the help provided by undergraduate and graduate students who worked with me on different occasions. Special thanks to Sarthak Tyagi, Siddhant Jain, Sidharth Manohar, Shivam Chandra, Abhay Chandra, Gavit Gogna, Anubhav Barsaiyan, Shubham Rao, Shubham Singla, Rajan Goyal, Mayuna Gupta, Tanuj Garg, Nikhil Shenoy, Kaustubh Prakash, Harsh Pare, Shashank Vasisht,

Udit Singla, Tarun Gupta, and Vivek Arora. I am very thankful to my friends Upendra Pathrikar and Rajeev Ranjan for all the technical and non-technical discussions, which kept me motivated.

I would also like to thank AKAB Healthcare Pvt. Ltd. and FITT, IIT Delhi for the funding. I am grateful to the Department of Electrical engineering and the High Performance Computing (HPC) facility of IIT Delhi. A special thanks to Mr. Satish, Mr. Yatindra, and Mr. Rakesh for their help.

Deepak Mishra

Abstract

Liver resection requires extreme precision, as the surgeon aims to remove damaged tissue or a predefined portion of liver leaving the remnant intact with its blood vessel supply. Thus, selection of the best resection technique becomes a crucial step. Blood loss and operating time are some of the key factors which help the surgeon in choosing the best resection technique, however, a single technique cannot satisfy all the requirements simultaneously. Random trials of combination of two techniques, for multi-modal resection, are reported as the solution. However, such combination is not efficient due to the manual switching of cutting tools during the mode change as well as tool blindness. An ultrasound (US) guided multi-modal resection can address these issues where US images are used to identify vessels present near the transection point and provide a real-time feedback for automatic switching of the cutting tools. US image analysis, thus, becomes the most critical part of the US guided multi-modal liver resection.

US images are adversely affected by artifacts such as, low contrast, speckle, and acoustic shadows. In addition, variations in size and shapes of blood vessels make the vessel detection and segmentation challenging. In this work, novel techniques are developed for US image enhancement, segmentation, and registration with an end goal of accurate measurements of blood vessels and other anatomical structures such as lesions. Among the four enhancement approaches developed in this work, edge probability based anisotropic diffusion and structure oriented generative adversarial network (SO-GAN) are the most important. In particular SO-GAN,

which is fast and free from input-dependent parameter tuning. In contrast to the existing speckle elimination approaches, it tries only to reduce speckle extent while maintaining structural and qualitative attributes of the images. A quantitative measurement of output speckle characteristics shows that the proposed GAN results in symmetric KL-divergence value of 0.404. It is better than the value 0.547, which is the best achieved by the existing approaches.

Similarly two convolutional neural network (CNN) based segmentation approaches are developed. Among these, the CNN trained with sub-problem specific deep supervision provides the best performance. In this, the fine resolution layers are trained to give more emphasis to the anatomical boundaries whereas coarse resolution layers are focused to discriminate object regions. The approach is tested for blood vessel segmentation in liver images and results in the dice index value of 0.79. The best value observed among the existing approaches is produced by U-net as 0.75. The proposed network also results in dice index value of 0.91 in lumen segmentation experiments on MICCAI 2011 IVUS challenge dataset, which is near to the provided reference value of 0.93.

Segmentation outputs are further used as landmark information to develop a CNN based coupled hybrid convolutional method for deformable medical image registration. The existing deep learning registration methods are either geometric (landmark based), iconic (intensity based) or decoupled hybrid in nature. A decoupled hybrid method sequentially uses landmark and intensity information to perform registration. In contrast, a coupled hybrid method uses the two information simultaneously to provide better results. To best of our knowledge, the proposed method is the first to perform a coupled hybrid registration in deep learning framework. The experiments performed using cardiac magnetic resonance images, introduced in the MICCAI ACDC challenge 2017, show that the proposed method is able to achieve a matching of 86.2%, 64.3%, and 66.3% for left ventricle cavity, myocardium, and right ventricle cavity

regions, respectively. This is considerably better than the matching figures of 69.9%, 42.8%, and 53.8% which are achieved by the state-of-the-art deep learning registration method. The proposed method is also extended for multi-modal registration between CT and US images.

All the developed analysis techniques are generic in nature and can be extended to other applications in future. Furthermore, a clinical device or an augmented reality system can be developed using these techniques to obtain the desired image guided feedback for an autonomous multi-modal resection.

सार

यकृत उच्छेदन हेतु अत्यधिक सटीकता की आवश्यकता होती है, क्योंकि सर्जन का उद्देश्य क्षतिग्रस्त ऊतक या यकृत के पूर्वनिर्धारित हिस्से को इस प्रकार हटाने का होता है कि रक्त वाहिका आपूर्ति के साथ अवशेष बरकरार रहे। इस प्रकार, सर्वश्रेष्ठ उच्छेदन तकनीक का चयन एक महत्वपूर्ण कदम है। रक्त की हानि और सर्जरी का समय कुछ प्रमुख कारक हैं जो सर्जन को सर्वश्रेष्ठ उच्छेदन तकनीक चुनने में मदद करते हैं, हालांकि, एकमात्र तकनीक सभी आवश्यकताओं को एक साथ संतुष्ट नहीं कर सकती है। बहुविध उच्छेदन हेतु दो तकनीकों के संयोजन के यादृच्छिक परीक्षणों को समाधान के रूप में प्रकशित किया जाता है। हालांकि, इस तरह के संयोजन सर्जन द्वारा मोड़ परिवर्तन के साथ-साथ उपकरण अंधेपन के कारण कुशल नहीं है। एक अल्ट्रासाउंड निर्देशित बहुविध उच्छेदन इन मुद्दों को संबोधित कर सकता है जहां अल्ट्रासाउंड छवियों का उपयोग उच्छेदन बिंदु के पास मौजूद धमनियों की पहचान करने के लिए तथा काटने के उपकरण के स्वचालित परिवर्तन हेतु वास्तविक समय प्रतिक्रिया प्रदान करने के लिए करते हैं। इस प्रकार, अल्ट्रासाउंड छवि विश्लेषण, अल्ट्रासाउंड निर्देशित बहुविध यकृत उच्छेदन का सबसे

महत्वपूर्ण हिस्सा बन जाता है।

अल्ट्रासाउंड छवियां विषमताओं जैसे कि कम दृश्यता, स्पेकल और ध्वनिक छाया से प्रतिकूल रूप से प्रभावित होती हैं। इसके अलावा, रक्त वाहिकाओं के आकार और आकार में भिन्नता उनके पता लगाने और विभाजन को चुनौतीपूर्ण बनती हैं। इस काम में, रक्त वाहिकाओं एवं अन्य संरचनात्मक संरचनाओं के सटीक माप हेतु अल्ट्रासाउंड छवि वृद्धि, विभाजन और पंजीकरण की नई तकनीक विकसित की गई हैं। इस कार्य में विकसित चार संवर्द्धन पद्धतियों में, एज प्रोबेबिलिटी आधारित एनिसोट्रोपिक डिफ्यूजन और स्ट्रक्चर ओरिएंटेड जेनरेटिव एडवरसैरियल नेटवर्क (एसओ-जीएन) सबसे महत्वपूर्ण हैं। विशेष रूप से एसओ-जीएन, जो तेज और इनपुट-निर्भर पैरामीटर ट्यूनिंग से मुक्त है। मौजूदा स्पेकल उन्मूलन तरीकों के विपरीत, यह केवल छवियों की संरचनात्मक और गुणात्मक विशेषताओं को बनाए रखते हुए स्पेकल के प्रभाव को कम करने की कोशिश करता है। आउटपुट स्पेकल विशेषताओं की एक मात्रात्मक माप से पता चलता है कि प्रस्तावित जीएन का सममित केएल-विचलन मूल्य 0.404 है। यह मूल्य 0.547 से बेहतर है, जो कि मौजूदा पद्धतियों में सबसे अच्छा है।

इसी तरह दो कन्वोलुशनल न्यूरल नेटवर्क (सीएनएन) आधारित विभाजन पद्धतियां विकसित की गई हैं। इनमें, उप-समस्या विशिष्ट गहन पर्यवेक्षण के साथ प्रशिक्षित सीएनएन सर्वश्रेष्ठ प्रदर्शन प्रदान करता है। इसमें बारीक रिज़ॉल्यूशन लेयर को संरचनात्मक सीमाओं

पर अधिक जोर देने के लिए प्रशिक्षित किया जाता है जबकि मोटे रेजोल्यूशन लेयर्स को ऑब्जेक्ट क्षेत्रों में भेदभाव करने के लिए केंद्रित किया जाता है। इस पद्धति का यकृत छवियों में रक्त वाहिका विभाजन के लिए परीक्षण किया गया है और परिणाम 0.79 के डाइस सूचकांक मूल्य में प्राप्त हुआ है। मौजूदा तरीकों के बीच सबसे अच्छा मूल्य यू-नेट का 0.75 है। प्रस्तावित नेटवर्क MICCAI 2011 IVUS चुनौती के डेटासेट पर भी लुमेन विभाजन प्रयोगों में 0.91 के डाइस सूचकांक मूल्य का परिणाम देता है, जो प्रदान की गई संदर्भ डाइस मूल्य, 0.93, के पास है।

विखंडन चिकित्सा छवि पंजीकरण के लिए सी-एनएन आधारित युग्मित संकर कन्वोल्यूशनल विधि विकसित करने के लिए विभाजन आउटपुट को लैंडमार्क सूचना के रूप में आगे उपयोग किया गया है। मौजूदा डीप लर्निंग पंजीकरण विधियाँ या तो ज्यामितीय (लैंडमार्क आधारित), प्रतिष्ठित (छवि आधारित) या प्रकृति में संकरित हैं। क्रमिक रूप से डिकपल्ड हाइब्रिड विधि पंजीकरण करने के लिए लैंडमार्क और छवि की जानकारी का उपयोग करती है। इसके विपरीत, एक युग्मित संकर विधि बेहतर परिणाम प्रदान करने के लिए एक साथ दो सूचनाओं का उपयोग करती है। हमारे ज्ञान अनुसार, प्रस्तावित विधि डीप लर्निंग ढांचे में प्रथम युग्मित संकर पंजीकरण विधि है। MICCAI ACDC चुनौती, 2017 में पेश किए गए हृदय चुंबकीय अनुनाद छवियों का उपयोग करते हुए किए गए प्रयोगों से पता चलता है कि प्रस्तावित विधि बाएं वेंट्रिकल गुहा, मायोकार्डियम, और दाएं

वैट्रिकल गुहा क्षेत्रों के लिए क्रमशः 86.2%, 64.3% और 66.3% का मिलान प्राप्त करने में सक्षम है। यह 69.9%, 42.8% और 53.8% के मिलान आंकड़ों की तुलना में काफी बेहतर है, जो कि अत्याधुनिक डीप लर्निंग पंजीकरण पद्धति द्वारा प्राप्त किया जाता है। प्रस्तावित विधि सीटी और अल्ट्रासाउंड छवियों के बीच मल्टी-मोडल पंजीकरण के लिए भी विस्तारित है।

सभी विकसित विश्लेषण तकनीक प्रकृति में व्यापक हैं और भविष्य में अन्य अनुप्रयोगों के लिए विस्तारित की जा सकती हैं। इसके अलावा, एक स्वायत्त बहु-विध उच्छेदन के लिए वांछित छवि निर्देशित प्रतिक्रिया प्राप्त करने के लिए इन तकनीकों का उपयोग करके एक नैदानिक उपकरण या संवर्धित वास्तविकता प्रणाली विकसित की जा सकती है।

Contents

List of Figures	xvii
List of Tables	xxvii
List of Abbreviation	xxix
1 Introduction	1
1.1 Liver resection methods	2
1.2 Image guided multi-modal resection	4
1.3 Ultrasound imaging and analysis	6
1.3.1 Speckle reduction for US image enhancement	7
1.3.2 Vessel segmentation	7
1.3.3 Image registration	9
1.4 Conceptual design of multi-modal resection	9
1.5 Objectives	11
1.6 Major contributions	12
1.7 Thesis structure	14
2 Ultrasound Image Enhancement using Speckle Reduction	17
2.1 Introduction	21

2.2	Existing approaches	22
2.2.1	Uniform speckle reduction	23
2.2.2	Selective speckle reduction	29
2.3	Edge probability and pixel relativity based anisotropic diffusion	31
2.3.1	Implementation	33
2.3.2	Stability analysis	38
2.3.3	Selective filtering	40
2.3.4	Complexity control	44
2.3.5	Experimental results	45
2.3.6	Discussion	59
2.4	Structure oriented adversarial network	61
2.4.1	Network architecture	63
2.4.2	Training	64
2.4.3	Dataset	67
2.4.4	Experimental results	68
2.5	Despeckling CNN with ensembles of classical outputs	74
2.5.1	Network architecture	75
2.5.2	Data generation	76
2.5.3	Training	78
2.5.4	Experimental results	78
2.5.5	Extended study	81
2.6	Summary	85

3 Vascular Region Segmentation in Ultrasound Images 87

3.1	Introduction	90
-----	------------------------	----

3.2	Background	90
3.3	Pipeline of CNN and unsupervised clustering	93
3.3.1	Localization of vascular regions	93
3.3.2	Pixel-wise segmentation	94
3.3.3	Experimental Results	95
3.4	Subproblem specific deep supervision	100
3.4.1	Network architecture	102
3.4.2	Fusion	105
3.4.3	Dataset	106
3.4.4	Training	107
3.4.5	Experimental results	109
3.5	Extension to segmentation of other anatomical structures	120
3.5.1	Lumen segmentation	120
3.5.2	Lesion segmentation	123
3.5.3	Discussion	127
3.6	Summary	127
4	Deformable Image Registration	129
4.1	Introduction	131
4.2	Background	133
4.3	Methodology	134
4.3.1	Segmentation network	136
4.3.2	Registration network	137
4.4	Experiments and results	139
4.4.1	Dataset	139

4.4.2	Baseline	141
4.4.3	2D Segmentation experiment	142
4.4.4	2D Registration experiment	142
4.4.5	3D segmentation and registration	151
4.5	Multi-modal registration	152
4.5.1	Dataset	153
4.5.2	Results	154
4.5.3	Localization: A missing link	154
4.6	Summary	158
5	Conclusions and Future Scope	159
5.1	Future work	161
5.1.1	Prototype	161
5.1.2	Disentangled representation learning using GANs	163
	Bibliography	165
	List of Publications	191

List of Figures

1.1	Liver US images with different speckle extent.	8
1.2	Degradation of vessel visibility due to acoustic shadow. Vessel boundaries are highlighted using green contours.	8
1.3	Conceptual design of the multi-modal resection method. It consists of ultrasound unit, processing unit, switching circuit, liver transection unit, and power supply.	10
2.1	Speckle reduction with uniform and selective filtering approaches.	22
2.2	Overview of the complete PDF calculation process. For each pixel, orientation binning of the gradients inside $M \times M$ surrounding window is performed to calculate weights of HOG bins. The bin weights are normalized to obtain the desired PDF.	34
2.3	Test image used to understand the effect of HOG window size variation on edge probabilities. An edge and a non-edge pixel, marked by red and yellow squares, are considered for the experiment.	36

2.4	Edge probability values observed for different window sizes $M \times M$ around the edge and non-edge pixel marked in the test image shown in Fig. 2.3. Probability values for the edge pixel are high in the vertical direction due to the vertical edge of the object. For the noisy non-edge pixel, the probability values are random, in general.	36
2.5	(a) Test image used to understand the behaviour of HOG at the extremity of a thin object, (b) the edge probability values observed for different window sizes $M \times M$ around the marked pixel.	37
2.6	Illustration of pixel relativity. The two edge pixels have higher relativity between them as compared to their common neighbour.	41
2.7	Synthetic phantom image with multiplicative Rayleigh noise. Regions of low echogenicity (yellow) and high echogenicity (red) are marked in the image for experiments.	43
2.8	Variation of relativity (a) r and (b) r in two regions marked in the synthetic phantom image shown in Fig. 2.7.	43
2.9	Filtered phantom images obtained using (a)-(g) SRAD, DPAD, OBNLM, RTAD, ADMSS and EPPRSRAD without and with GPR. EPPRSRAD shows better performance as compared to other filters. Speckle in the desired regions is best preserved by EPPRSRAD in presence of GPR.	48
2.10	Performance of different filters with noise variance increasing from 0.2 to 1.0. .	49
2.11	(a) US image simulated using Field-II simulator, (b)-(h) filtered images obtained using SRAD, DPAD, OBNLM, RTAD, ADMSS and EPPRSRAD without and with GPR.	50

2.12 (a) Real US image of a liver showing multiple vessels and an echolucent cyst, (b)-(h) Filtered liver images obtained using SRAD, DPAD, OBNLM, RTAD, ADMSS and EPPRSRAD without and with GPR. Speckle in parenchymal re- gion (a small portion marked using red circle in the original image) is best preserved by EPPRSRAD with GPR.	52
2.13 Endocardial border detection using BFAC segmentation in (a) original image, (b)-(h) outputs of SRAD, DPAD, OBNLM, RTAD, ADMSS and EPPRSRAD without and with GPR. The contour segmented in the output of EPPRSRAD with GPR is closest to the ground truth.	58
2.14 Despeckling or generator network architecture. The first convolutional (Conv) layer contains 7×7 kernels and has 64 channels (ch). It is combined with batch normalization (Batch Norm) and rectified linear units (ReLU). The first layer is followed by a chain of six ResNet blocks and a final Conv layer.	63
2.15 Discriminator network architecture. The strided convolutional layers are used to avoid the maxpooling layers along with the Leaky ReLU ($\alpha_r = 0.2$). The final output is a single value from the sigmoid activation.	64
2.16 Examples of (a) low speckle extent image (V) acquired using RAB6-D probe, (b) high speckle extent image (Z) acquired using X7-2t probe. Marked regions are used for performance evaluation of different approaches.	68
2.17 Despeckled outputs obtained using different approaches. The proposed ap- proach shows the best output.	70
2.18 Despeckled output obtained from DRNN and the observed PMF.	72
2.19 Despeckled outputs obtained using different approaches. PMFs observed from the reference (V), input (Z) and output (U) are also shown.	73

2.20	Proposed CNN architecture. The first convolutional (Conv) layer contains 7×7 kernels with 128 channels (ch). It is followed by a chain of three ResNet blocks and a final Conv layer to produce the desired outputs.	76
2.21	Sample ensembles generated using the combination of DPAD, OBNLM and EAGF with ADMSS.	77
2.22	Sample of size 96×96 pixels showing different anatomical structures.	79
2.23	Sample outputs obtained using the proposed DsCNN. First row: test images, second row: $DsCNN_i$ output, third row: $DsCNN_e$ output, fourth row: ground truth.	80
2.24	Sample outputs obtained using different approaches. $DsCNN_e$ results in the best visual quality.	84
3.1	Examples of blood vessel segmentation. Variations in shapes and sizes of different vessels along with signal loss and speckle make the task challenging. Left column: real US liver images, right column: manually segmented vessel regions.	91
3.2	First row: real US liver images, second row: outputs from CNN based patch-wise classification. Positively classified patches are marked by white squares. False positives are shown with red squares.	96
3.3	First row: a positive sample and corresponding feature maps, second row: a negative sample and corresponding feature maps.	97
3.4	Segmented outputs of the images shown in Fig. 3.2. Outputs from K -means clustering, region growing and thresholding are in first, second and third row respectively.	97

3.5	Performance of the considered filters on test images. Segmented regions are shown in magenta and ground truth region boundaries are shown by green contours. First column: output using the method of Nam et al., second column: output using the method of Blackall et al., third column: output of Frangi filter, fourth column: proposed CNN + K -means output.	99
3.6	Complete architecture of the DSN-OB network. The convolutional layers from $C_{1.1}$ to $C_{5.3}$ form the core network and deeply supervised by the auxiliary layers. The auxiliary layers represented by E_i 's are used to learn the boundary definition and the layers represented by the O_i 's are focused on the object region discrimination from the background. This subproblem specific deep supervision results in the desired segmentation.	103
3.7	Training behaviour of DSN-O and DSN-Os. The plot clearly shows that DSN-O converges faster as compared to DSN-Os.	110
3.8	Fusion layer parameters ($\mathbf{h}$) observed from the trained DSN-O network. A large non-zero value of h_0 shows that network gains considerable information from this layer.	111
3.9	Comparison of segmentation performances of the networks. The outputs obtained using DSN-Os, DSN-O, and DSN-OB are shown in the first, second, and third row, respectively. Segmented regions are shown in magenta and ground truth region boundaries are shown by green contours. The performance of DSN-O is better than DSN-Os, however, the best results are shown by DSN-OB. . . .	112
3.10	Observations of the experiments performed with DSN-OB for different values of T_s	113

3.11	Dice index variation for validation set in the experiments performed with DSN-OB for different values of T_s . Dotted lines show dice index values measured at every epoch whereas solid lines show the best value of dice index observed over all previous epochs.	113
3.12	Comparison of the output probability maps observed from DSN-O and DSN-OB networks. Clearly the boundaries in the output of DSN-OB are sharper than DSN-O.	115
3.13	The outputs obtained using Frangi filter, U-net, and DSN-OB are shown in the first, second, and third row, respectively. Segmented regions are shown in magenta and ground truth region boundaries are shown by green contours. Clearly the proposed DSN-OB network shows the best performance with least false negatives and false positives.	117
3.14	(a) Input image, (b) DSN-OB output, (c) intermediate outputs from the auxiliary layers, first row: up-sampled output of the core network layer, second row: application of learned 3×3 convolutional filter on first row images, third row: application of sigmoid on second row images.	119
3.15	Examples of lumen segmentation. First row: IVUS liver images, second row: manually segmented lumen regions.	121
3.16	Boundaries of the lumen regions in the test images from IVUS dataset, green curve shows the ground truth and magenta and yellow curves represent the DSN-OB and U-net outputs, respectively.	123

3.17	Segmented outputs from the approach of Wang et al., U-net and DSN-OB are shown in the first, second, and third row, respectively. Segmented regions are shown in magenta and ground truth region boundaries are shown by green contours. The better performance of the proposed network can be appreciated from the outputs in second and fourth column.	126
4.1	Complete overview of the proposed registration method. Initially the masks corresponding to S and T are generated, denoted by $\hat{S}$ and $\hat{T}$, respectively. These masks are used by a registration network, which maps the input $(S, \hat{S}, T, \hat{T})$ to ϕ and subsequently produces the warped images $S(\phi)$ and $\hat{S}(\phi)$	135
4.2	Detailed deformation field plots, in (a) absence and (b) presence of segmentation masks. At each location of the plots, arrows point out the direction of the field and their lengths show the magnitude.	135
4.3	Architecture of the model used for segmentation. Except first and last rectangle, all the other represent convolutional layers, where number of filters/channels are mentioned inside the rectangles. X and $\hat{X}$ are input and output of the network. $\hat{X}$ has a same resolution as X with C channels, where each channel corresponds to one part of the object. Reduction in size of the rectangle represents reduction of resolution by a factor of 2. Auxiliary branches contain a combination of convolutional and upsampling layers and produce outputs of same resolution as X with C number of channels.	137
4.4	Architecture of the model used for registration is similar to the segmentation network. The concatenation of $S, \hat{S}, T$, and $\hat{T}$ is given as input to the network to generate the warped image $S(\phi)$. ϕ is a 2(3) channel output of the fusion layer which is used by an ST layer to produce the warped images.	138

4.5	Auxiliary deformation fields observed at the output of auxiliary branches. The outputs are obtained from finest (O_0 and O_1) to coarsest (O_5) resolution. Effect of segmentation masks can be clearly observed at different scales.	140
4.6	Test image outputs of the segmentation networks. Boundaries of the segmented right ventricle cavity, myocardium, and left ventricle cavity regions are shown in the top, middle, and bottom rows, respectively. Green curves show the ground truth and magenta curves show the outcome of the segmentation network. Left ventricle cavity regions are segmented with relatively higher accuracy due to a good contrast and regularity of the surrounding myocardium.	143
4.7	Sample pairs from test set with overlaid boundaries of right (green) and left (red) ventricle cavities, row 1 and 2. Registered images from log-demons, the method by Balakrishnan et al., non-hybrid DSN-R, and hybrid DSN-R, are respectively shown in row 3, 4, 5, and 6. The maximum similarity is observed between the structures of the images in row 2 and 6, which reflects the best performance of the proposed method.	146
4.8	Qualitative evaluation of inverse consistency of DSN-R. The order of source and target image pairs shown in Fig. 4.7 is reversed (left and middle column) and given as input to DSN-R for registration. The resultant warped images are shown in the right column. DSN-R shows a consistent performance in most cases. The performance can be attributed to the ordered pairing of the images used to create the training dataset.	148

4.9	Improvement in registration performance (dice index) with improving landmark information provided by the segmentation network at different training epochs. At epoch 0, the dice index values produced by the hybrid DSN-R are similar to the non-hybrid DSN-R. The values gradually improve with improving segmentation masks. RV: right ventricle cavity, Myo: myocardium, LV: left ventricle cavity.	149
4.10	(a) Source and target input slices, (b) registered outputs obtained using non-hybrid and hybrid DSN-R, (c) deformation field magnitude ($ \phi $) observed from non-hybrid and hybrid DSN-R.	150
4.11	Slices from 3D test volumes, first row: source, second row: target. Boundaries of right (green) and left (red) ventricle cavities are overlaid for comparison. Registered outputs from log-demons, VoxelMorph-2 by Balakrishnan et al., and hybrid DSN-R, are respectively shown in third, fourth, and fifth rows.	153
4.12	DSN-R performance on multi-modal registration with CT and US images. Even rows show the lesion regions corresponding to the images shown in preceding odd rows. Poor performance can be observed from the registered image and corresponding lesion masks shown in the right column.	155
4.13	Overview of the proposed registration framework with localization network. SSD, DSN-OB, and DSN-R form a pipeline to perform localization, segmentation, and registration respectively. Registration is done only between the mask to handle the difference in modalities of CT and US.	156
4.14	DSN-R performance on multi-modal registration with SSD. Considerably better outputs as compared to the outputs shown in Fig. 4.12 are observed.	157

5.1 Diagram of a conceptual system for US guided multi-modal liver resection. It consists of US acquisition system, a server/workstation for US image analysis, a device containing transection probes attached to actuators for both swift and selective cutting, and AR glasses for surgeon’s assistance. 162

List of Tables

1.1	Comparison of liver resection methods	3
2.1	Optimized parameters values used for synthesized phantom image along with the observed $SSNR_d$	47
2.2	Optimized parameters values used for cyst phantom image experiment	50
2.3	SSNR values observed from the filtered outputs of cyst phantom image	51
2.4	Implementation parameter values used for real liver US images	53
2.5	SSIM values for the filtered outputs of liver US images	54
2.6	$SSNR_d$ values observed from the filtered liver images	55
2.7	Implementation parameter values used for real cardiac US image set	56
2.8	SSIM values for the filtered outputs of cardiac US images	56
2.9	Hausdorff distances between the segmented BFAC contour and the ground truth for the filtered outputs obtained using different filters	59
2.10	Diagnostic preference scores given by expert radiologists for the filtered outputs of real images obtained using different filters	60
2.11	Evaluation of the approaches on simulated images	71
2.12	Evaluation of despeckling approaches on real image	72
2.13	Comparison of DsCNN performance on test data	81

2.14	DsCNN performance evaluation for different training sample sizes	82
2.15	Cross-validation and comparison of DsCNN with state-of-the-art approaches .	83
3.1	Architecture of the proposed deep CNN. The names conv, pool and FC represent the convolutional, pooling and fully connected layers, respectively	95
3.2	Quantitative evaluation of the CNN + K -means clustering pipeline	99
3.3	Evaluation metric values observed for DSN-Os and DSN-O for blood vessel segmentation	111
3.4	Evaluation metric values observed from the outputs of different networks for blood vessel segmentation	115
3.5	Dice index values observed from the outputs of the DSN-OB for the lumen segmentation in MICAAI 2011 IVUS challenge dataset	123
3.6	Evaluation metric values observed from the outputs of different approaches for lesion segmentation dataset	125
4.1	Evaluation of the segmentation network on 2D slices	143
4.2	Quantitative evaluation of the proposed registration method on 2D slices	144
4.3	Evaluation of the segmentation network on 3D CMR volumes	151
4.4	Quantitative evaluation of the proposed registration method on 3D volumes . .	152
4.5	Quantitative evaluation of the proposed multi-modal registration method on the pairs of CT and simulated US images in presence of despeckling	158

List of Abbreviation

ADMSS	Anisotropic diffusion with memory based on speckle statistic
AR	Augmented reality
BFAC	Balloon force active contour
CEUS	Contrast-Enhanced ultrasound
CMR	Cardiac magnetic resonance
CNN	Convolutional neural network
CNN-ff	Feedforward convolutional neural network
CT	Computed Tomography
CUSA	Cavitron Ultrasonic Surgical Aspirator
CV	Coefficient of variation
DH	Donor Hepatectomy
DPAD	Detail preserving anisotropic diffusion
DRNN	Despeckling residual neural network
DSN	Deeply supervised net
DSN-O	Deeply supervised network - object
DSN-OB	Deeply supervised network - object and boundaries
DSN-Os	Deeply supervised network - object small
DSN-R	Deeply supervised network - registration

DsCNN	Despeckling convolutional neural network
EAGF	Edge aware geometric filter
EPI	Edge preservation index
EPPRSRAD	Edge probability and pixel relativity based speckle reducing anisotropic diffusion
FC	Fully connected
FCNN	Fully convolutional neural network
FLL	Focal liver lesion
FNH	Focal Nodular Hyperplasia
FOM	Pratt's figure of merit
GF	Geometric filter
GOSF	General order statistics filter
GPR	Global pixel relativity
HOG	Histogram of oriented gradients
HS	Harmonic Scalpel
IVUS	Intravascular ultrasound
IoU	Intersection over union
JM	Jaccard measure
LDDMM	deformation diffeomorphic metric mapping
MRI	Magnetic resonance imaging
MS-SSIM	Multi-scale structural similarity index measure
NLM	Non-local mean
OBNLM	Optimized Bayesian Non Local Mean
OSRAD	Oriented speckle reducing anisotropic diffusion
PDF	Probability density function

PMF	Probability mass function
POSRAD	Probability driven oriented speckle reducing anisotropic diffusion
PR	Pixel relativity
PSNR	Peak signal to noise ratio
RF	Radio-frequency
RFA	Radio-frequency ablation
RTAD	Rayleigh trimmed anisotropic diffusion
ReLU	Rectified linear units
SBF	Squeeze box filter
SEM	Scanning electron microscopic
SI	Speckle index
SLIC	simple linear iterative clustering
SO-GAN	Structure oriented generative adversarial network
SRAD	Speckle reducing anisotropic diffusion
SSD	Single shot multi-box detector
SSIM	Structural similarity index measure
SSNR	Speckle region's signal to noise ratio
US	Ultrasound
WJS	Water-jet dissection