AWARD NUMBER: W81XWH-17-1-0438

TITLE: Multimodality Imaging Platform for Neurovascular Bundle Sparing Prostate Radiotherapy to Preserve Sexual Function

PRINCIPAL INVESTIGATOR: Tian Liu, PhD

CONTRACTING ORGANIZATION: Emory University, Atlanta, GA

REPORT DATE: September 2020

TYPE OF REPORT: Annual Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
September 2020	Annual	01Sep2019-31Aug2020

4. TITLE AND SUBTITLE

Multimodality Imaging Platform for Neurovascular Bundle Sparing Prostate Radiotherapy to Preserve Sexual Function

5a. CONTRACT NUMBER
W81XWH-17-1-0438

5b. GRANT NUMBER
PC160820

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Tian Liu, PhD

E-Mail: tian.liu@emory.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Emory University
Holly Sommers
1599 Clifton Rd.
Atlanta, GA 30322-4250

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The overall project goal is to develop an advanced imaging technology to assess treatment-related erectile dysfunction and thereby improve sexual outcomes and quality of life of prostate-cancer survivors. This project has 3 major steps: (1) develop a new technique for combining MRI and ultrasound images in a way that can allow physicians to see the neurovascular bundle (NVB) clearly and reliably, (2) confirm that these new imaging findings match the patients' outcome with respect to their erectile function, and (3) use the new imaging techniques to reduce radiation dose to the NVB. Over the current period, we have made significant progress in the development of the technology to combine MRI and Ultrasound images (including Doppler images), and made significant progress on subject accrual for the calibration study and longitudinal clinical study that is the core of the project.

15. SUBJECT TERMS
Prostate cancer, Radiotherapy, Ultrasound, Magnetic Resonance Imaging, Erectile Dysfunction

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	71	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

1. Introduction……. 4

2. Keywords……. 4

3. Accomplishments……..……... 4

4. Impact……...…… 8

5. Changes/Problems...…..…… 8

6. Products……..……..…..……. 9

7. Participants & Other Collaborating Organizations…….. 12

8. Special Reporting Requirements……..… 13

9. Appendix…… 13

1. INTRODUCTION:

The ultimate goal of our project is to advance imaging technology to assess treatment-related erectile dysfunction and thereby improve sexual outcomes and quality of life of prostate-cancer survivors. Loss of sexual potency is the most common and debilitating long-term complication in radiotherapy for prostate cancer. Over 50% of men will suffer from post-radiation erectile dysfunction which may lead to diminished sexual activity and desire, decreased sexual satisfaction, changes in intimate relationships and reduced quality of life post-radiotherapy. This project has 3 major steps. First, we will develop a new technique for combining MRI and ultrasound images in a way that can allow physicians to see the neurovascular bundle (NVB) clearly and reliably. Second, we will confirm that these new imaging findings match the patients' outcome with respect to their erectile function. Third, we will use the new imaging techniques to reduce radiation dose to the NVB to improve preservation of sexual function.

2. KEYWORDS:

Prostate cancer, Radiotherapy, Ultrasound, Magnetic Resonance Imaging, Erectile Dysfunction

3. ACCOMPLISHMENTS:

What were the major goals of the project?

Below are the **Specific Aims** and **Major Tasks** (with embedded timelines) as listed in the original/approved Statement of Work (SOW):

Specific Aim 1: Develop a novel MRI and US technology to measure radiation-related NVB injury
- **Major Task 1.1:** Integration of MRI and Quantitative Doppler Ultrasound (months 1-6)
- **Major Task 1.2:** Conduct phantom study of 3D Doppler ultrasound imaging (months 1-6)
- **Major Task 1.3:** Verify reliability of the multimodality imaging technology (months 2-6)

Specific Aim 2: Conduct a clinical study to validate that the quantitative results of the novel MR-US technology correlate with standard clinical endpoints used to evaluate ED
- **Major Task 2.1:** Preparation for clinical study (months 1-6)
- **Major Task 2.2:** Interim Analysis of Clinical Study (months 7-18)
- **Major Task 2.3:** Final Analysis of Clinical Study (months 7-36)

Specific Aim 3: Conduct a feasibility study to demonstrate utility of the novel MR-US technology in improving patient sexual potency through NVB-sparing radiotherapy
- **Major Task 3.1:** Determine NVB dose received during the original treatment plans (months 7-26)
- **Major Task 3.2:** Correlate NVB dose with clinical outcomes data (months 26-36)
- **Major Task 3.3:** Re-plan with constraints on DVH's (months 30-36)

What was accomplished under these goals?

Overall, work is progressing well along the timeline indicated in the previous section:

1. Major activities:

Under Aim 1, we have calibrated the algorithms developed under Major Task 1.1 and 1.3 above on patient data. For Major Task 1.2, we did construct a unique prostate phantom consisting of a standard prostate phantom and custom-adding neurovascular bundle components.

Scientifically, we have three major achievements for Aim 1. One achievement is to develop in-house software to analyze Doppler ultrasound waveform of the neurovascular bundles. Our Doppler analysis software provides more accurate blood flow measurements than current commercial software (Figure 1).

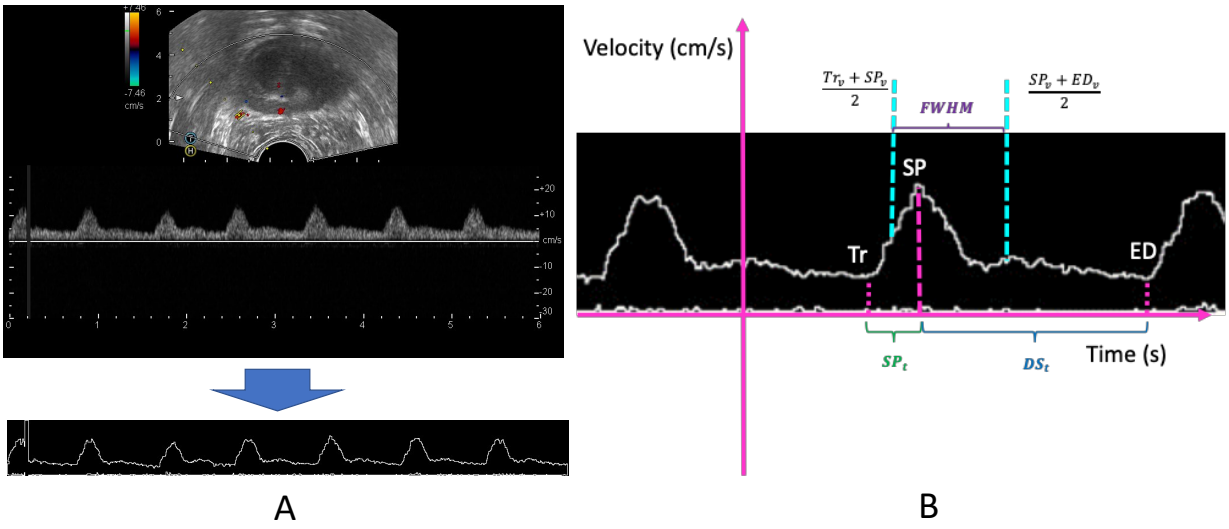


Figure 1: Quantitative analysis of artery flow at neurovascular bundle. (A) Pulse wave extraction from Doppler image using ImageJ. (B) Morphological features are quantified for pulse waveform.

The second major achievement is to develop an ultrasound prostate segmentation based on multidirectional deeply supervised V-Net (Figure 2 below). Prostate segmentation is an important step in building a robust multimodality imaging platform for prostate radiotherapy. Our segmentation technology based on deep-learning concept is accurate and reliable.

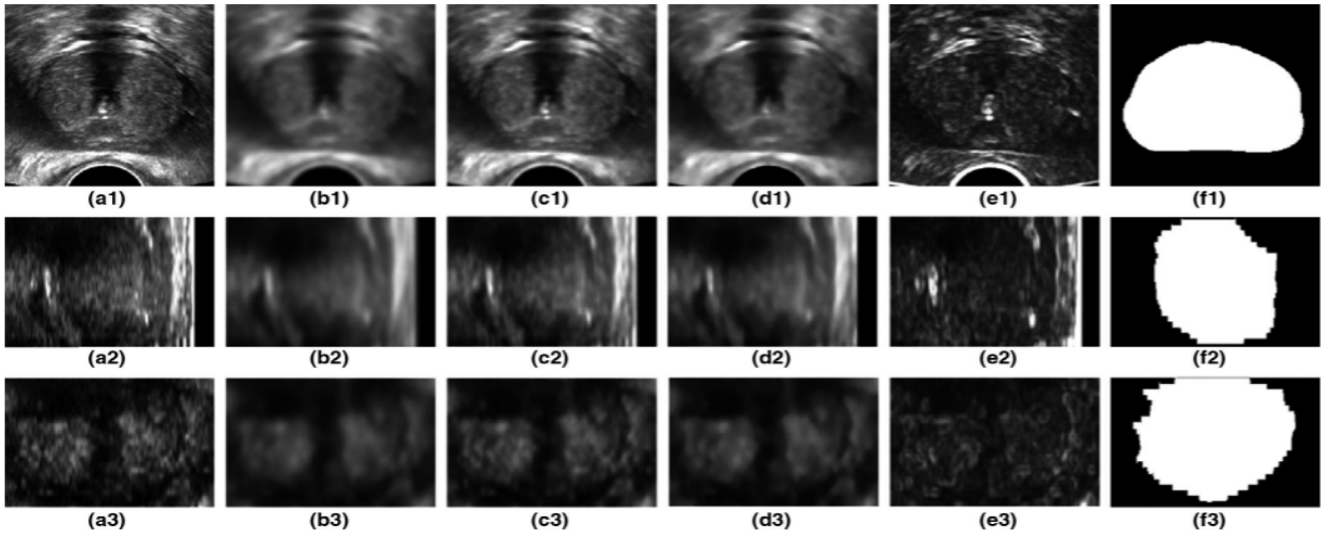


FIG. 2. The multiderivative images and the corresponding prostate. (a1–a3) TRUS images represented in transverse plane, sagittal plane, and coronal plane. (b1–b3), (c1–c3), (d1–d3) Corresponding images generated by 3D Gaussian filter, 3D mean filter, and 3D median filter. (e1–e3) Corresponding prostate masks based on physicians' manual contours. The display window for (a–d) is [0, 200].

The third major achievement in the image registration between MRI and transrectal ultrasound (TRUS) outlined in Aim 1. The major achievement is to develop a novel deep learning-based 3D point cloud matching method for MR-ultrasound prostate image registration. The registration framework consists of a convolutional neural networks (CNN) for MR prostate segmentation, a CNN for TRUS prostate segmentation and a point-

cloud based network for rapid 3D point cloud matching (Figure 3). The point cloud matching network was trained using deformation field that was generated by finite element analysis. Therefore, the network implicitly models the underlying biomechanical constraint when performing point cloud matching. Population-based FE models were conducted to simulate prostate deformations between MR and TRUS. Figure 4 is an example case showing the performance of our ProRegNet as compared with other registration methods including Rigid, SurfReg and CnnReg. Our results demonstrated that the proposed method was able to accurately and rapidly register MR and TRUS prostate images for image-guided prostate radiotherapy.

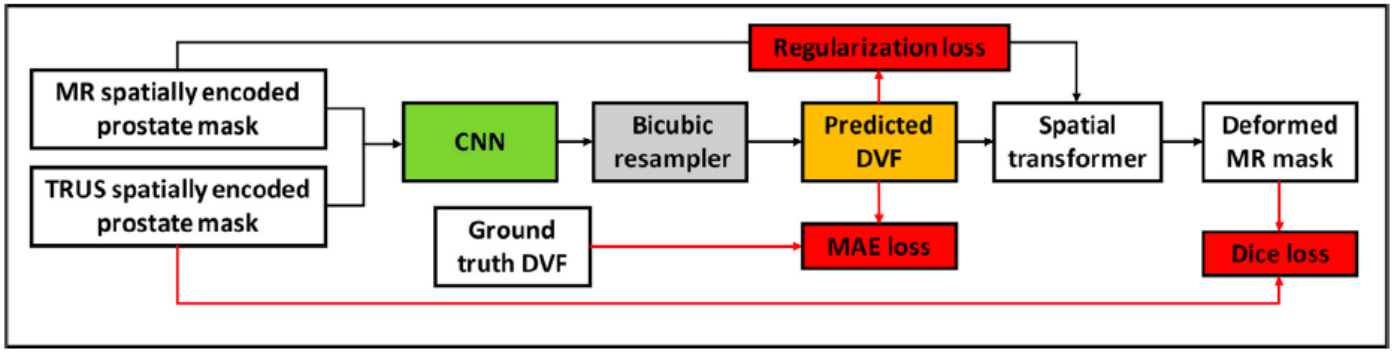


Figure 3. Network design of CnnReg. The CNN network consists of 12 convolutional layers and 3 max pooling layers. The loss function contains deformation vector field (DVF) regularization loss, DVF mean absolute error loss and Dice loss.

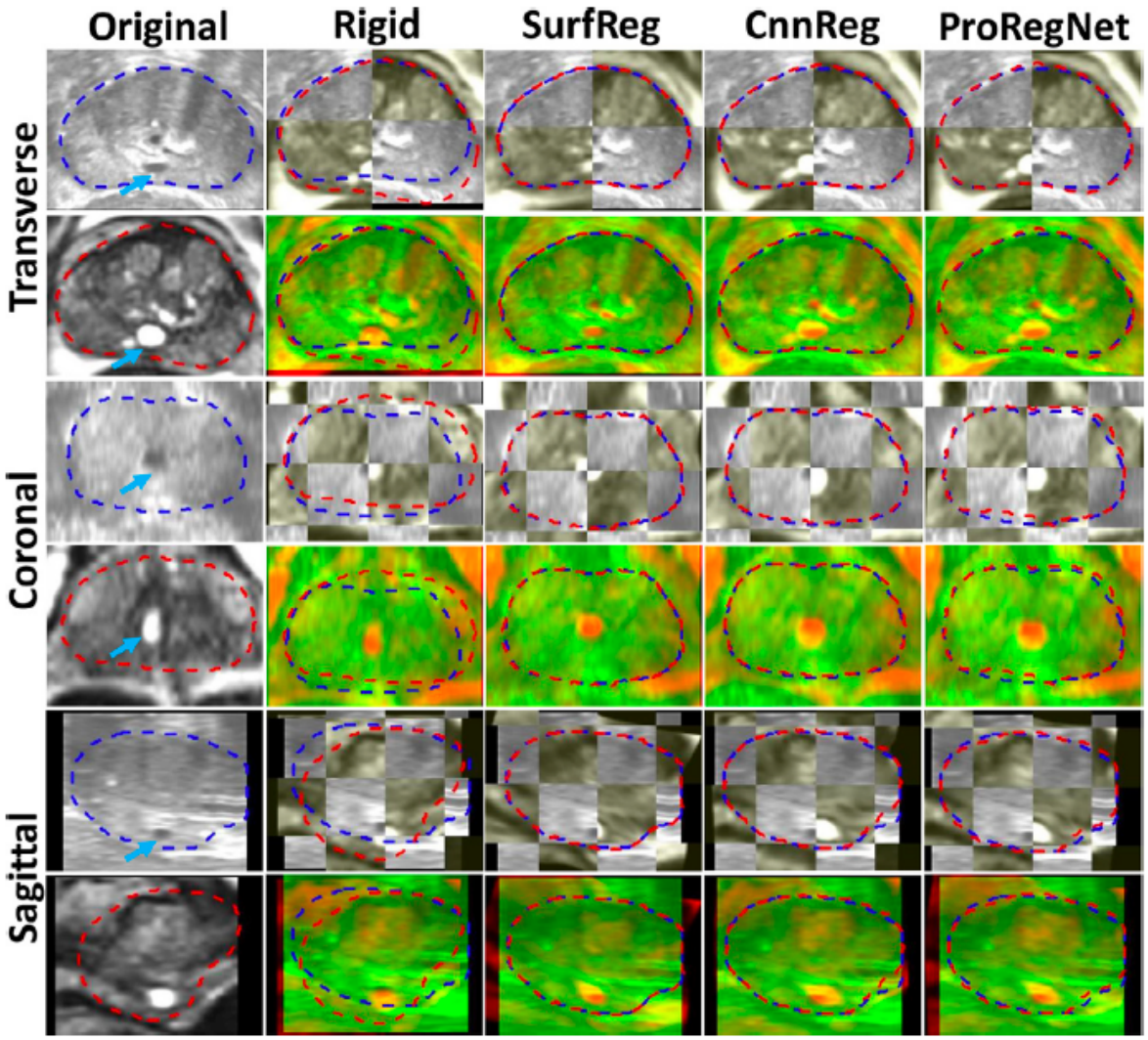


Figure 4. Example image slices from one case. Results are shown in the transverse, coronal and sagittal planes. First column shows the original TRUS and MR images. Second column shows the images after rigid registration in checkboard and red-green fusion. Third, fourth and fifth columns show

the registration results for SurfReg, CnnReg and ProRegNet, respectively. Blue dashed contour indicates the TRUS prostate shape while red dashed contour indicates MR prostate shape. The cyst within prostate were highlighted by arrows on the original images.

Under Aim 2, we have made significant progress on the clinical study and have now consented 36 subjects in 2 cohorts: 31 for the technology development and 4 for the longitudinal study (Table 1). One subject was an overall screen failure, and most have been screen failures for the longitudinal study (due to receiving ADT or having an IIEF score too low) but their imaging data has been captured for further refinement and calibration of the registration algorithms in Major Tasks 1.1-1.3.

Table 1 Patient Characteristics – Total Consented: 36 patients

Characteristics	Value
Age (y)	
Median	66
Range	52-84
Race	
African-American	22
Asian	2
White	12
Marital Status	
Divorced	1
Married	22
Separated	1
Single	10
Unknown	2
Study Cohort	
Longitudinal	4
Reliability	31
Screen Failure	1
Exclusion Criteria for Longitudinal Cohort	
Pre Hormonal Therapy	18
Incomplete Pre-RT MRI	7
Documented Moderate-Severe ED	7
Radiation Therapy	
EBRT alone	16
EBRT plus brachytherapy	13
Brachytherapy alone	7
Imaging Data Collected	57
Ultrasound (B-mode & Color Doppler)	
Pre RT	35
Pre Brachytherapy	19
6 months F/U	3
MRI	
Pre RT	4
6 months F/U	3
1 year F/U	1

Abbreviation: EBRT = external beam radiation therapy, ED = erectile dysfunction, RT = radiation therapy, F/U = follow up.

Under Aim 3, we have updated and maintained the constructed imaging database and an outcomes database with the above subjects to later correlate imaging, dosimetric data, and clinical data.

2. Specific Objectives:
All specific objectives, as listed in the Specific Aims above, are on target. The only items that are still underway are Major Tasks 2.3 and Major Tasks 3.3, which will be completed under no-cost extension.

3. Significant Results or Key Outcomes:
We were able to successfully develop and calibrate the quantitative ultrasound/Doppler technology (including prostate segmentation and registration with MRI data) using patient data (for subjects on both the reliability and longitudinal categories of subjects on the clinical study), as reported in numerous abstract and publications listed in section 6 (PRODUCTS) below.

4. Other Achievements:
Not applicable

What opportunities for training and professional development has the project provided?
Nothing to Report
How were the results disseminated to communities of interest?
Nothing to Report
What do you plan to do during the next reporting period to accomplish the goals?
We plan to continue consenting/screening subjects to our longitudinal clinical study to continue to collect imaging and clinical data for correlation.

4. **IMPACT:**
What was the impact on the development of the principal discipline(s) of the project?
The impact on the principal discipline is to date that of technology development, with respect to (a) Ultrasound Doppler technology to detect and measure blood flow to the neurovascular bundles, and (b) MRI/Ultrasound registration technology.
What was the impact on other disciplines?
Nothing to Report
What was the impact on technology transfer?
Nothing to Report
What was the impact on society beyond science and technology?
Nothing to Report

5. **CHANGES/PROBLEMS:**
Changes in approach and reasons for change
Nothing to Report

Actual or anticipated problems or delays and actions or plans to resolve them
One difficulty encountered, which was described in the previous annual reports, is that the majority of our consented patients [who have data that can be used for the initial ultrasound/MR imaging technology development] are screen failures for the longitudinal study, as a result of having started ADT or having an IIEF score that is too low. Though we believed we would have an increase in favorable risk prostate cancer patients due to our opening NRG GU 005 open at our institution this was not the case and thus we still had very few patients meeting eligibility for the longitudinal study. An additional difficulty is that a few subjects who were otherwise eligible for the longitudinal study could not participate in that portion of the study due to not having a pre-treatment MRI scan. Though we considered expand the eligibility criteria - to allow for a wider range of IIEF scores and/or allow ADT use and/or allow more aggressive pathology and/or relax the pre-

treatment MRI requirement - in order to facilitate the longitudinal study enrollment – we felt that this would not be a sound approach.

The other major difficulty encountered was the impact on accrual by COVID-19. While as shown in Table 1 above, we did have overall robust enrollment (36 consented subjects [including 8 subjects in Year 3 of the grant, which is the year in which COVID-19 pandemic started], which is relatively close to the overall initial accrual goal of 39 total subjects), COVID-19 did impact our ability to enroll patients on the longitudinal study, on which we have relatively few subjects.
Hence we have concentrated our efforts more on using the subject data to calibrate the algorithms on the subject data, which was the main overall objective of our grant.

Changes that had a significant impact on expenditures
Nothing to Report
Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents
Nothing to Report
Significant changes in use or care of human subjects
Nothing to Report
Significant changes in use or care of vertebrate animals.
Not Applicable
Significant changes in use of biohazards and/or select agents
Nothing to Report

6. **PRODUCTS** (note that all products that are related in any way to the current project are listed – note also that ***as indicated below, the two most significant items listed below are included in their entirety in the Appendix*****):**
Publications, conference papers, and presentations
Journal publications.

1. Wang B, Lei Y, Tian S, Wang T, Liu Y, Pretesh P, Jani A, Mao H, Curran W, Liu T and Yang X. "Deeply Supervised 3D FCN with Group Dilated Convolution for Automatic MRI Prostate Segmentation" *Medical Physics*, 46(4):1707-1718, 2019.
2. Lei Y, Tian S, He X, Wang T, Wang B, Pretesh P, Jani A, Mao H, Curran W, Liu T and Yang X. "Ultrasound Prostate Segmentation Based on Multi-Directional Deeply Supervised V-Net," *Medical Physics*, 46(6):3194-3206, 2019. ***[PLEASE SEE APPENDIX]***
3. Wang T, Zhou J, Tian S, Wang Y, Patel P, Jani A, Langen K, Curran W, Liu T and Yang X. A Feasibility Study of Focal Dose Escalations to Multiparametric MRI-defined Dominant Intraprostatic Lesions in Prostate Proton Radiation Therapy. *Br J Radiol*. 96: 20190845, 2019.
4. Lei Y, Dong X, Tian Z, Liu Y, Wang T, Tian S, Patel P, Jani A, Curran W, Mao H, Liu T and Yang X. CT Prostate Segmentation Based on Synthetic MRI-aided Deep Attention Fully Convolution Network. *Med Phys*. 47(1), 530-540, 2020.
5. Zhang Y, Lei Y, Qiu R, Wang T, Wang H, Jani A, Curran W, Patel P, Liu T and Yang X. Multi-needle Localization with Total Variation Regularized Deep Supervised Attention U-Net in Ultrasound-guided HDR Prostate Brachytherapy. *Med Phys*. 2020, PMID: 32155666, https://doi.org/10.1002/mp.14128.
6. Fu Y, Lei Y, Wang T, Patel P, Jani A, Mao H, Curran W, Liu T and Yang X. Biomechanically Constrained Non-rigid MR-TRUS Prostate Registration using Deep Learning based 3D Point Cloud Matching. *Med Image Anal*. 2020. ***[PLEASE SEE APPENDIX]***
7. Dai X, Lei Y, Zhang Y, Qiu R, Wang T, Dresser S, Curran W, Patel P, Liu T and Yang X. Automatic Multi-Needle Detection using Deeply Supervised Convolutional Neural Network in MRI-guided HDR Prostate Brachytherapy. *Med Phys*. 2020, https://doi.org/10.1002/mp.14307. (online)

8. Zeng Q, Fu Y, Tian Z, Lei Y, Zhang Y, Wang T, Mao H, Liu T, Curran W, Jani A, Patel P and Yang X. Label-driven Magnetic Resonance Imaging (MRI)-transrectal Ultrasound (TRUS) Registration Using Weakly Supervised Learning for MRI-guided Prostate Radiotherapy. *Phys Med Biol*. 65(13), 135002, 2020, https://doi.org/10.1002/mp.14307 ***[PLEASE SEE APPENDIX]***
9. Fu Y, Lei Y, Wang T, Tian S, Patel P, Jani A, Curran W, Liu T and Yang X. Pelvic Multi-organ Segmentation on Cone Beam CT for Prostate Adaptive Radiotherapy. *Med Phys*. 2020. (In press; online ahead of print)
10. Zhang Y, Zhen T, Yang L, Wang T, Patel P, Jani A, Curran W, Liu T and Yang X. Automatic Multi-needle Localization in Ultrasound Images using Large Margin Mask RCNN for Ultrasound-guided Prostate Brachytherapy. *Physics in Medicine and Biology*. 2020. (Major revision)
11. Zhang Y, Lei Y, Qiu R, Wang T, Wang H, Jani A, Curran W, Patel P, Liu T and Yang X. Multi-needle Localization with Attention U-Net in US-guided HDR Prostate Brachytherapy. *Med Phys*. 2020 (online ahead of print), PMID: 32155666, https://doi.org/ 10.1002/mp.14128

Books or other non-periodical, one-time publications.
Nothing to report

Other publications, conference papers, and presentations.
1. Wang B, Lei Y, Jeong J, Wang T, Tian S, Jiang X, Jani A, Mao H, Curran W, Patel P, Liu T and Yang X. "Automatic MR Prostate Segmentation Using 3D Deeply Supervised FCN With Concatenated Atrous Convolution," *SPIE Medical Imaging*, 2018. (Poster)
2. Lei Y, Tian S, Wang B, Wang T, Jani A, Curran W, Patel P, Liu T and Yang X. " Ultrasound Prostate Segmentation Based on 3D V-Net with Deep Supervision," *SPIE Medical Imaging*, 2018. (poster presentation)
3. Wang B, Lei Y, Wang T, Dong X, Tian S, Jiang X, Jani A, Mao H, Curran W, Patel P, Liu T and Yang X. "Automated Prostate Segmentation of Volumetric CT images Using 3D Deeply Supervised Dilated FCN," *SPIE Medical Imaging*, 2018. (Poster)
4. Giles M, Press R, Wang T, Schreibmann E, Jani A, Liu T, Rossi P, Curran W, Patel P and Yang X. "Multiparametric MRI-Guided High-Dose-Rate Brachytherapy Boost to the Dominant Intraprostatic Lesion Based On Deformable Registration: A Prospective Study," *Medical Physics,* 45 (6), E443-E443 (2018).
5. Wang T, Lei Y, Jani A, Patel P, Dhabaan A, Liu T, Dong X, Jiang X, Tang X, Curran W and Yang X. "Learning-Based Cone-Beans CT Correction: Dosimetric Calculation Accuracy in Prostate Cancer Radiation Therapy," *Medical Physics,* 45 (6), E169-E170 (2018).
6. Yang X, Lei Y, Tian S, Wang T, Jani A, Curran W, Patel P, and Liu T. "3D Ultrasound Prostate Segmentation Using 3D Deeply Supervised V-Net," *Medical Physics,* 45 (6), E473-E473 (2018).
7. Yang X, Wang B, Lei Y, Wang T, Jiang X, Jani A, Mao H, Curran W, Patel P and Liu T. "3D Prostate Segmentation in MR Image Using 3D Deeply Supervised Convolutional Neural Networks," *Medical Physics,* 45 (6), E582-E583 (2018).
8. Wang B, Lei Y, Wang T, Dong X, Tian S, Jiang X, Jani A, Mao H, Curran W, Patel P, Liu T and Yang X. "Automated Prostate Segmentation of Volumetric CT images Using 3D Deeply Supervised Dilated FCN," Presented at the Annual Meeting of International Society for Optics and Photonics (SPIE) Medical Imaging, San Diego, CA, USA 2019. (Poster)
9. Wang B, Lei Y, Jeong J, Wang T, Tian S, Jiang X, Jani A, Mao H, Curran W, Patel P, Liu T and Yang X. "Automatic MR Prostate Segmentation Using 3D Deeply Supervised FCN With Concatenated Atrous Convolution," Presented at the Annual Meeting of International Society for Optics and Photonics (SPIE) Medical Imaging, San Diego, CA, USA 2019. (Poster)
10. Lei Y, Tian S, Wang B, Wang T, Jani A, Curran W, Patel P, Liu T and Yang X. " Ultrasound Prostate Segmentation Based on 3D V-Net with Deep Supervision," Presented at the Annual Meeting of International Society for Optics and Photonics (SPIE) Medical Imaging, San Diego, CA, USA 2019. (Oral)
11. He X, Yang X, Jani A, Sohn J, Patel P, Curran W and Liu T. "Reliability of Doppler Blood Flow Evaluation of Neurovascular Bundle Vessels in Patients Receiving Prostate Radiotherapy," Presented at 61st annual meeting

of the American Association of Physicists in Medicine (AAPM), San Antonio, TX, USA, 2019. (Oral) *[**PLEASE SEE APPENDIX**]*

12. Zeng Q, Jeong J, Lei Y, Tian Z, Wang T, Dong X, Jani A, Patel P, Mao H, Curran W, Liu T and Yang X. "Surface-Driven MRI-US Registration Using Weakly-Supervised Learning in Prostate Brachytherapy," Presented at 61st annual meeting of the American Association of Physicists in Medicine (AAPM), San Antonio, TX, USA, 2019. (Oral)

13. Lei Y, Tian S, Wang T, Liu Y, Dong X, Jiang X, Patel P, Jani A, Curran W, Liu T and Yang X. "Synthetic MRI-Aided Prostate Segmentation in CT Image," Presented at 61st annual meeting of the American Association of Physicists in Medicine (AAPM), San Antonio, TX, USA, 2019. (Oral)

14. Yang X, Zeng Q, Lei Y, Tian S, Wang T, Dong X, Jani A, Mao H, Curran W, Patel P and Liu T. "MRI-US Registration Using Label-driven Weakly-supervised Learning for Multiparametric MRI-guided HDR Prostate Brachytherapy," Presented at 60th Annual Meeting of American Society for Radiation Oncology (ASTRO), Chicago, IL, USA, 2019. (E-poster)

15. Zeng Q, Jeong J, Zhen T, Wang T, Lei Y, Mao H, Jani A, Patel P, Curran W, Liu T, and Yang X, "Surface-driven Deformable MRI-US Prostate Registration Using Weakly-supervised Deep Learning," Presented at the Annual Meeting of International Society for Optics and Photonics (SPIE) Medical Imaging, Houston, TX, USA, 2020. (Poster Presentation)

16. Zeng Q, Jeong J, Wang T, Lei Y, Mao H, Jani A, Patel P, Curran W, Liu T, and Yang X, "Non-rigid MRI-TRUS Prostate Registration Using Fully Convolutional and Recurrent Neural Networks," Presented at the Annual Meeting of International Society for Optics and Photonics (SPIE) Medical Imaging, Houston, TX, USA, 2020. (Poster Presentation)

17. Wang T, Giles M, Press R, Dai X, Jani A, Rossi P, Lei Y, Curran W, Patel P, Liu T, and Yang X, "Multiparametric MRI-guided High-dose-rate Prostate Brachytherapy with Focal Dose Boost to Dominant Intraprostatic Lesions," Presented at the Annual Meeting of International Society for Optics and Photonics (SPIE) Medical Imaging, Houston, TX, USA, 2020. (Poster Presentation)

Website(s) or other Internet site(s).

Work on this grant (both initiating and partnering PI contributions) was highlighted by the CDMRP website during Prostate Cancer Awareness month at the below link:
https://cdmrp.army.mil/pcrp/research_highlights/20liu_highlight.aspx

Note that this content is included in its entirety in PDF form. [***PLEASE SEE APPENDIX***]

Below is the CDRMP liaison who was involved in this highlight:

Jodi Belz, Ph.D.
PCRP Scientific Manager
Biomedical Life Scientist | Leidos
Supporting the Congressionally Directed Medical Research Programs, USAMRDC
5202 Presidents Court, Suite 110
Frederick, MD 21703
Phone: 206.276.9227
jodi.e.belz.ctr@mail.mil

The corresponding program book is still in process of being published.

Technologies or techniques.

1. An ultrasound Doppler technology that can detect blood flow in right and left neurovascular bundles in patients receiving radiotherapy for prostate cancer.

2. An image registration technology that can accurate register MRI and ultrasound images of the prostate.

Inventions, patent applications, and/or licenses.

n/a

[Note in particular that the item listed in the 2018 annual report: Liu T, Yang X and Jani A. "MRI-guided NVB Sparing in Prostate Cancer Radiation Therapy." (Emory Ref: 15229) was an application for a patent that was ultimately not approved, and the DD Form 882 associated with the closeout of this grant accurately lists no inventions or patents were generated from this work].

Other Products.

We have constructed an imaging database and an outcomes database to be able to correlate imaging and clinical data, in order to reach Aims 2 and 3 of the project over the next year.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name	**Tian Liu, PhD**
Research Identifier:	NIH ERA COMMONS ID: liutian
Nearest person month worked:	1.92
Contribution to project:	Principal Investigator. Dr. Liu leads the technical aspects of the proposed study. In this funding period, she and her team made progress in prostate segmentation and registration of ultrasound and MRI images. She works closely with Drs. Jani, Patel and Mehta to refine the ultrasound Doppler technology to assess neurovascular function and correlate the NVB doses and imaging findings with patient-reported sexual outcomes.

Name	**Hui Mao, PhD**
Research Identifier:	N/A
Nearest person month worked:	0.6
Contribution to project:	Co-I. Mr. Mao provides guidance in MRI protocol design, as well as provide resources to support the MRI part of the proposed project. In this funding period, he refined the MRI protocol for patients' 6-month and 1-year follow-up scans

Name	**Xiaofeng Yang, PhD**
Research Identifier:	N/A
Nearest person month worked:	0.6
Contribution to project:	Co-I. Dr. Yang has made significant progress in building the multimodality NVB imaging platform for the proposed study. He introduced AI-based technology to prostate registration and improved its accuracy.

Name	**Xiuxiu He, PhD**
Research Identifier:	N/A
Nearest person month worked:	10
Contribution to project:	Dr. He assisted physicians (Dr. Jani and Patel) in patient ultrasound scanning. Through the proposed study, she has analyzed ultrasound images from all clinical sites, and ensure inter- and intra-observer reliability of the imaging technology.

Name	**Simone Henry, PhD**
Research Identifier:	N/A

Nearest person month worked:	1.2
Contribution to project:	Ms Henry is assisting with regulatory matters and patient consent/accrual for the clinical study and data management portion of the project. She is assisting Drs. Jani in the recruitment and screening of study participants and coordinating the day-to-day activities of the study including the collection of all patient data. Ms. Henry is in charge of the storage of ultrasound and MRI data. She prepares and submits IRB documents and reports to the Data Safety and Management Board

Name	**Jooyoung Jung, PhD**
Research Identifier:	N/A
Nearest person month worked:	4
Contribution to project:	Dr Jung has designed and built a Doppler phantom which is critical in testing our ultrasound technology. He will analyze MR and ultrasound images from all clinical sites and ensure inter- and intra-observer reliability of the imaging technology.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to Report

What other organizations were involved as partners?

Nothing to Report

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS:

An independent report will be submitted by the partnering PI (Dr. Ashesh Jani, Award number W81XWH-17-1-0439 / PC160820P1).

QUAD CHARTS:

Nothing to Report

9. APPENDIX – the below 2 items (also listed in section 6 above) are included in their entirety:

1. Lei Y, Tian S, He X, Wang T, Wang B, Pretesh P, Jani A, Mao H, Curran W, Liu T and Yang X. "Ultrasound Prostate Segmentation Based on Multi-Directional Deeply Supervised V-Net," *Medical Physics*, 46(6):3194-3206, 2019.
2. Zeng Q, Fu Y, Tian Z, Lei Y, Zhang Y, Wang T, Mao H, Liu T, Curran W, Jani A, Patel P and Yang X. Label-driven Magnetic Resonance Imaging (MRI)-transrectal Ultrasound (TRUS) Registration Using Weakly Supervised Learning for MRI-guided Prostate Radiotherapy. *Phys Med Biol*. 65(13), 135002, 2020, https://doi.org/10.1002/mp.14307
3. Fu Y, Lei Y, Wang T, Patel P, Jani A, Mao H, Curran W, Liu T and Yang X. Biomechanically Constrained Non-rigid MR-TRUS Prostate Registration using Deep Learning based 3D Point Cloud Matching. *Med Image Anal*. 2020.
4. He X, Yang X, Jani A, Sohn J, Patel P, Curran W and Liu T. "Reliability of Doppler Blood Flow Evaluation of Neurovascular Bundle Vessels in Patients Receiving Prostate Radiotherapy," Presented at 61st annual meeting of the American Association of Physicists in Medicine (AAPM), San Antonio, TX, USA, 2019. (Oral)
5. PDF of the following CDRMB highlight link: https://cdmrp.army.mil/pcrp/research_highlights/20liu_highlight.aspx

Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net

Yang Lei, Sibo Tian, Xiuxiu He, Tonghe Wang, Bo Wang, Pretesh Patel, and Ashesh B. Jani
Department of Radiation Oncology and Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA

Hui Mao
Department of Radiology and Imaging Sciences and Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA

Walter J. Curran, Tian Liu, and Xiaofeng Yang[a)]
Department of Radiation Oncology and Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA,

(Received 18 September 2018; revised 14 April 2019; accepted for publication 1 May 2019; published xx xxxx xxxx)

Purpose: Transrectal ultrasound (TRUS) is a versatile and real-time imaging modality that is commonly used in image-guided prostate cancer interventions (e.g., biopsy and brachytherapy). Accurate segmentation of the prostate is key to biopsy needle placement, brachytherapy treatment planning, and motion management. Manual segmentation during these interventions is time-consuming and subject to inter- and intraobserver variation. To address these drawbacks, we aimed to develop a deep learning-based method which integrates deep supervision into a three-dimensional (3D) patch-based V-Net for prostate segmentation.

Methods and materials: We developed a multidirectional deep-learning-based method to automatically segment the prostate for ultrasound-guided radiation therapy. A 3D supervision mechanism is integrated into the V-Net stages to deal with the optimization difficulties when training a deep network with limited training data. We combine a binary cross-entropy (BCE) loss and a batch-based Dice loss into the stage-wise hybrid loss function for a deep supervision training. During the segmentation stage, the patches are extracted from the newly acquired ultrasound image as the input of the well-trained network and the well-trained network adaptively labels the prostate tissue. The final segmented prostate volume is reconstructed using patch fusion and further refined through a contour refinement processing.

Results: Forty-four patients' TRUS images were used to test our segmentation method. Our segmentation results were compared with the manually segmented contours (ground truth). The mean prostate volume Dice similarity coefficient (DSC), Hausdorff distance (HD), mean surface distance (MSD), and residual mean surface distance (RMSD) were 0.92 ± 0.03, 3.94 ± 1.55, 0.60 ± 0.23, and 0.90 ± 0.38 mm, respectively.

Conclusion: We developed a novel deeply supervised deep learning-based approach with reliable contour refinement to automatically segment the TRUS prostate, demonstrated its clinical feasibility, and validated its accuracy compared to manual segmentation. The proposed technique could be a useful tool for diagnostic and therapeutic applications in prostate cancer. © *2019 American Association of Physicists in Medicine* [https://doi.org/10.1002/mp.13577]

Key words: deeply supervised network, deep learning, prostate segmentation, transrectal ultrasound (TRUS)

1. INTRODUCTION

Prostate cancer is the second leading cause of cancer-related death in men in the United States.[1] Transrectal ultrasound (TRUS) is a standard imaging modality for image-guided prostate-cancer procedures (e.g., biopsy and brachytherapy).[2] Accurate segmentation of the prostate plays a key role in biopsy needle placement, radiotherapy treatment planning, and motion monitoring.[3] Manual segmentation during biopsy or radiation therapy planning can be time-consuming and subject to inter- and intraobserver variation.[4] As ultrasound images have a relatively low signal-to-noise ratio (SNR), automated segmentation of the prostate is challenging.

Recently, a number of techniques have been developed to segment the prostate from TRUS images. The current TRUS segmentation techniques can be briefly summarized as the following:

1. Non-machine-learning-based methods include contour and shape-based methods and region-based methods.[5] Contour and shape-based methods segment the prostate based on boundary information,[6] which can be affected by ambiguous boundaries of the prostate apex and base in TRUS images. Prior shape information has been applied to address this issue.[7,8] Region-based methods use predominant intensity distributions of the prostate region to

segment the TRUS contour, which are affected by the speckle noise in TRUS images. In addition, there are many non-machine-learning-based methods including atlas-based,[9,10] graph-based,[11,12] or level sets-based methods.[13,14] Ghose et al. reviewed these methods in detail.[5]

2. Machine learning-based methods cluster the TRUS voxels into prostate and non-prostate tissues based on different leaning-based models. These methods can be classified into two types, either using unsupervised or supervised models. The unsupervised methods perform tissue classification based on TRUS contour information,[15,16] and shape priority.[17] The supervised methods train a classifier using a set of training data with their associated labels (prostate or non-prostate) and then the well-trained classifier performs the segmentation for a newly acquired ultrasound image.[18,19] The supervised methods can be grouped into support vector machines (SVM)-based, random forest-based, and the deep learning-based methods. The SVM-based and random forest-based methods use TRUS contour boundary information, such as texture features or shape statistic information, to train a SVM or random forest classifier for future segmentation.[20–25] To address the problem that traditional machine-learning-based methods are challenging to handcrafted, high-dimensional, and ill-posed mapping from TRUS image to binary segmentation, a deep learning method has been introduced into medical image segmentation.[26–30] Yang et al. incorporated an auto-context model into recurrent neural networks to deal with the severe boundary incompleteness and enhanced the performance in prostate boundary delineation.[31] However, this method was based on two-dimensional (2D) patch inputs, which lacks spatial information and thus results in ambiguous boundary segmentation in low-contrast regions such as the prostate apex and base. Ghavami et al. proposed a method based on an improved convolutional neural networks (CNN) for 2D and three-dimensional (3D) TRUS images' prostate segmentation.[32] However, due to the lack of a stage-wise deep supervision, training such a network with limited patient data is difficult when all the convolutional kernels of each stage are optimized only based on a loss function at the final stage. Zhu et al. involved a deep supervision strategy into CNN for prostate segmentation.[33] Zeng et al. utilized magnetic resonance imaging priors for TRUS prostate segmentation.[34]

In this work, we proposed a multidirectional and multiderivative deep learning-based method to automatically segment the prostate. The contributions of the paper are as follows:

1. To cope with the optimization difficulties of training the deep learning-based model with limited training data, a deep supervision strategy with a hybrid loss function (logistic and Dice loss) was introduced to the different stages. These mechanisms could make the residual information semantically meaningful for the early stages and final stage in the network, and thus reduce convergence time and improve the segmentation performance of the network when training with limited patient data.
2. To reduce possible segmentation errors at the prostate apex and base in TRUS images, we introduced a multidirectional-based contour refinement model to fuse transverse, sagittal, and coronal plane-based segmentation.
3. The multiderivative images including the 3D original TRUS and multiple filtered images were used as multichannel samples to train the proposed network, which could perform a cross-modality feature learning to enhance the networks' capacity and performance.[35]

The paper is organized as follows: We first provide an overview of the proposed TRUS prostate segmentation framework in Materials and Methods, followed by detailed description of the multiderivative preprocessing, deeply supervised V-Net, and then multidirectional-based contour refinement post-processing. We evaluated the proposed method through a comparison with state-of-art segmentation methods U-Net[36] and V-Net[26] and verified their performance using clinical data. Finally, along with an extended discussion, we conclude the presentation of our novel TRUS prostate segmentation framework.

2. MATERIALS AND METHODS

2.A. Overview

The proposed prostate segmentation method consists of a training stage and a segmentation stage. For a given pair of TRUS image, the corresponding manual contour was used as the learning-based target. Both TRUS images from the training and testing sets were preprocessed to remove noise by bias correction and a despeckling method.[37] Multidirectional-based deep learning networks were trained using images in the transverse, sagittal, and coronal planes. For each plane, TRUS images were filtered by 3D Gaussian, mean, and median filters. The filtered images, together with original images were used to constitute a 4-channel image data, or multiderivative-based data. A 3D patch-based V-Net[26] architecture was introduced to enable end-to-end learning. A deep supervision strategy[38] combined with hybrid loss was used to deeply supervise the network. During the segmentation stage, 3D patches were extracted from multiderivative images of newly acquired images as the input of the well-trained networks, which performed a patch-based segmentation. The segmented prostate volume was obtained by patch fusion and was refined by multidirectional-based contour refinement. Figure 1 outlines the workflow schematic of our segmentation method.

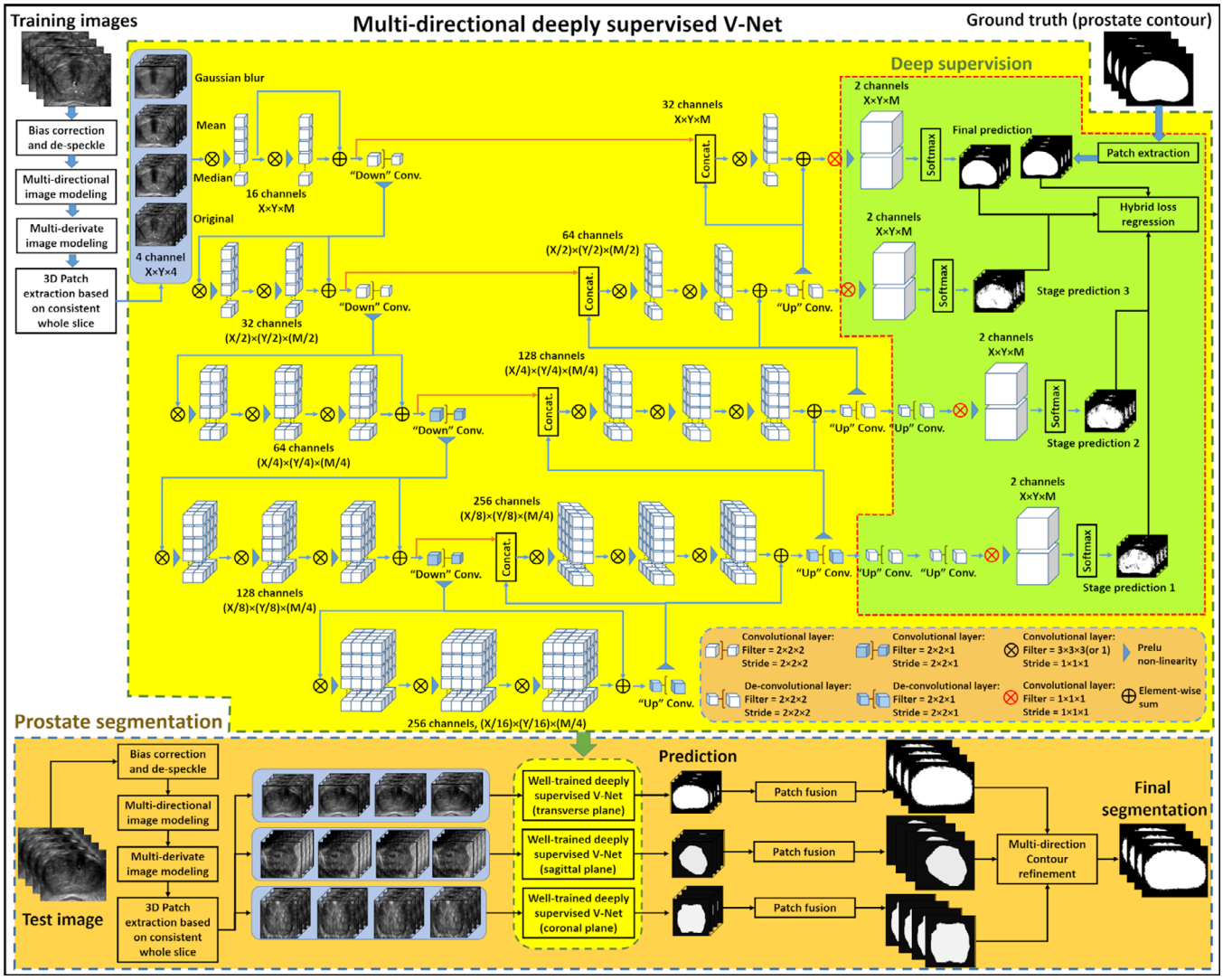


FIG. 1. The schematic flow diagram of the proposed method.

2.B. Multiderivative image

In the training of a segmentation network, extracted features from the entire image can enhance the segmentation accuracy by providing structural, global, and spatial information. However, it is typically not feasible to use whole 3D TRUS images due to the computational cost. In order to address the challenge of large-scale TRUS data, serial 3D image patches were extracted for the training. Moreover, for a successful segmentation task, it is important for a network to extract more and deep features of prostate in TRUS images. Thus, handcrafted filters, for example, 3D mean, median, and Gaussian filters were introduced to enhance the feature extraction of the prostate boundary. This was done through smoothing and denoising the original TRUS images to reduce the uninformative artifact texture in the original TRUS images, as shown in Fig. 2. Then, 3D patches of the original TRUS image and its multiderivative counterparts were used as the multichannel input samples to train our network.

2.C. Deeply supervised V-Net

Our proposed network architecture (yellow part of Fig. 1) was inspired by a well-known end-to-end V-Net.[26] The principle advantage of a V-Net is its ability to perform voxel-wise error back-propagation during the training stage, and generate a segmented patch with the same size as the input patch during testing.[26]

As shown in Fig. 1, the network consists of compression and decompression paths, and a bridge path that connects these two. The compression path is constructed by 1, 2, or 3 convolutional layers, which are called convolutional components, followed by a "down" convolutional layer to reduce the resolution. The decompression path is constructed by 1, 2, or 3 convolutional layers and followed by an "up" convolutional layer to enhance the resolution. The bridge path concatenates the feature maps from equal-sized compression and decompression paths.

From up to down, the network is grouped as five stages with different resolutions. Each stage consists of a

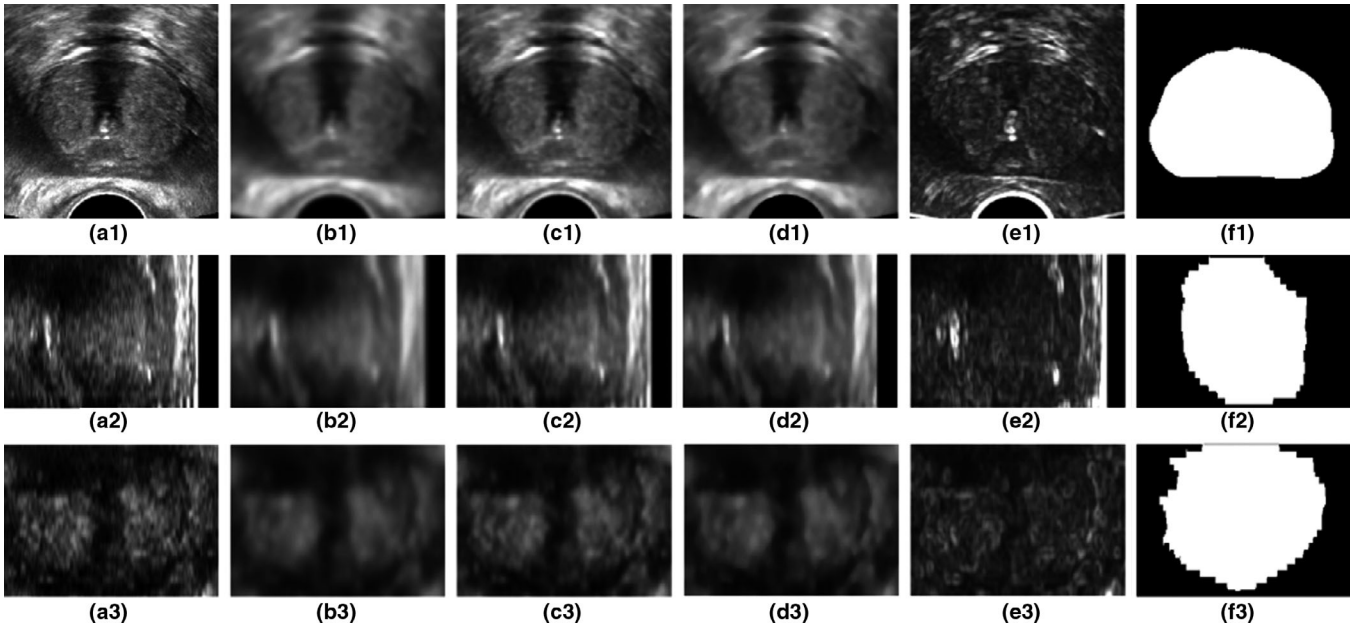


FIG. 2. The multiderivative images and the corresponding prostate. (a1–a3) TRUS images represented in transverse plane, sagittal plane, and coronal plane. (b1–b3), (c1–c3), (d1–d3) Corresponding images generated by 3D Gaussian filter, 3D mean filter, and 3D median filter. (e1–e3) Corresponding prostate masks based on physicians' manual contours. The display window for (a–d) is [0, 200].

compression path, a bridge path, a decompression path, a soft-max operator, and a threshold to binarize the output (with 0 and 1 denoting prostate and non-prostate regions, respectively). Assume input patch size of our network is $X \times Y \times M$, where X and Y denote the length and width of slice in our work, M denotes the number of slices. From up to down, the first stage generates feature maps with a size of $X \times Y \times M$, and is the final stage. The second stage generates feature maps with a size of $\frac{X}{2} \times \frac{Y}{2} \times \frac{M}{2}$, and is the high-resolution stage. The third stage generates feature maps with a size of $\frac{X}{4} \times \frac{Y}{4} \times \frac{M}{4}$, and is the modest-resolution stage. The fourth stage generates feature maps with a size of $\frac{X}{8} \times \frac{Y}{8} \times \frac{M}{4}$, and is the low-resolution stage. The last stage's output size is $\frac{X}{16} \times \frac{Y}{16} \times \frac{M}{4}$, which we denote the last stage as the bridge stage. The kernel size and stride size of the "down" convolutional layer and the corresponding "up" convolutional layer in the final, and the high-resolution stages are both set as $2 \times 2 \times 2$. The kernel size and stride size of the convolutional component in the final and the high-resolution stages are set as $3 \times 3 \times 3$ and $1 \times 1 \times 1$, respectively. Since the depth of patch is reduced to one in the last three stages, the kernel size and stride size of the "down" convolutional layer and the corresponding "up" convolutional layer in the modest-resolution, low-resolution, and bridge stages are set as $2 \times 2 \times 1$ and $2 \times 2 \times 1$, respectively. The kernel size and stride size of the convolutional component in the modest-resolution, low-resolution, and bridge stages are set as $3 \times 3 \times 1$ and $1 \times 1 \times 1$, respectively.

2.C.1. Deep supervision

To cope with the optimization difficulties when training a deep network with limited data, we incorporated deep supervision[33,38] into a V-Net, as shown in green part of Fig. 1. Since the output sizes of these stages are equal to the original input size, for the final and high-resolution stages, an "up" convolution operator was not needed to retrieve the image size. Since the patch is downsampled by factors of two and four, in order to obtain an equal-size output in the modest-resolution and low-resolution stages, these two stages are followed by one or two more "up" convolution operators, and then followed by a soft-max and threshold operator to obtain the equal-size segmentation.

2.C.2. Hybrid loss function

Recent work has used either logistic or Dice loss as loss functions in their networks.[26,36] We proposed to combine logistic loss which is used to measure dissimilarity, with Dice loss which is used to measure similarity, into a hybrid loss function to supervise our network at four stages. Voxel-wise binary cross-entropy (BCE) loss is commonly used as a logistic loss. Since the segmentation task can be regarded as a binary regression, we used the voxel-wise BCE loss as the logistic loss function. The BCE loss is defined as follows:

$$L_{\mathrm{BCE}}(I_B, \hat{I}_B) = -\sum_j I_{Bj} \log(\hat{I}_{B_j}) + (1 - I_{B_j}) \log(1 - \hat{I}_{B_j}) \tag{1}$$

where I_B denotes the voxel value in manual segmentation, $\hat{I}_B$ denotes the voxel value of the automatic segmentation generated in final, high, modest, and low-resolution stages. j denotes the $j-$th voxel. I_B or $\hat{I}_B$ equal to 1 indicates the voxel belongs to the segmented prostate, vice versa.

A Dice similarity coefficient (DSC) loss[26] is also introduced, which is defined as:

$$L_{\mathrm{DSC}}(I_S, \hat{I}_S) = 1 - \frac{2|I_S \cap \hat{I}_S|}{|I_S| + |\hat{I}_S|} \tag{2}$$

where I_S denotes the prostate mask from manual contour, and $\hat{I}_S$ denotes the prostate mask from automatic segmentation in final, high-, modest- and low-resolution stages.

Combining the above two loss functions, the hybrid loss function for deep supervision at the different stages is defined as follows:

$$L_{\mathrm{final}}(I, \hat{I}) = \sum_{l=1}^{4} \lambda_l \left(L_{\mathrm{BCE}}(I_B, \hat{I}_B^l) + \mu L_{\mathrm{DSC}}(I_S, \hat{I}_S^l) \right) \tag{3}$$

where l denotes the stage of our network. λ_l denotes the regularization weights of each stage's loss, and is set by $\lambda_l = \rho^{l-1}$, empirically. Since the resolution difference of each stage, we set $\rho \in (0, 1]$. μ is a balancing parameter, which balances the BCE and DSC losses.

2.D. Multidirectional-based contour refinement

As stated in Section 2.B, if we used only input slices from the transverse plane, the prostate apex and base cannot be accurately identified on TRUS. To address this issue, we propose a post-processing method called multidirectional-based model to automatically refine the segmentation results.

Suppose $\hat{I}_{\mathrm{transverse}}$, $\hat{I}_{\mathrm{sagittal}}$, and $\hat{I}_{\mathrm{coronal}}$ are the three segmentations from the transverse, sagittal, and coronal planes. Since the refinement's range-of-view occurs in the apex and base of the prostate, that is, around the poles of the contour surface in the z-axis. Thus, we cast N rays from prostate contour's central position O. These N rays are generated by restricting the variation range of the polar angle along the z-axis, as shown in Fig. 3. The orange lines of Fig. 3(c) shows the distribution range of N rays, where θ denotes the polar ray's polar angle, the range of θ is $[-30°, 30°]$, φ denotes the polar ray's azimuthal angle, the range of φ is $[0°, 360°)$. The orange lines of Figs. 3(a) and 3(b) show the distribution of N rays in the sagittal and coronal planes, respectively.

As shown in Figs. 3(a) and 3(b), these polar rays have N intersection points with the surface of prostate contour. We denote the intersection coordinates of segment contours $\hat{I}_{\mathrm{transverse}}$, $\hat{I}_{\mathrm{sagittal}}$ and $\hat{I}_{\mathrm{coronal}}$ with ray n as $\hat{I}_{\mathrm{transverse}}(n)$, $\hat{I}_{\mathrm{sagittal}}(n)$ and $\hat{I}_{\mathrm{coronal}}(n)$, $n = 1, 2, ...N$. The multidirectional-based contour refinement is performed by calculating the distance between these three contours. The distance is defined as follows:

$$\begin{aligned} d(n) = &\left\| \hat{I}_{\mathrm{transverse}}(n), \hat{I}_{\mathrm{sagittal}}(n) \right\| \\ &+ \left\| \hat{I}_{\mathrm{transverse}}(n), \hat{I}_{\mathrm{coronal}}(n) \right\| \\ &+ \left\| \hat{I}_{\mathrm{sagittal}}(n), \hat{I}_{\mathrm{coronal}}(n) \right\| \end{aligned} \tag{4}$$

where $||I_{\mathrm{A}}(n), I_{\mathrm{B}}(n)||$ denotes the Euclidean distance between the coordinates $I_{\mathrm{A}}(n)$ and $I_{\mathrm{B}}(n)$. Then, inspired by Ding's work,[39] the multidirectional-based contour refinement is implemented differently based on this distance. The refinement is given as follows:

$$\hat{I}_{\mathrm{final}}(n) = \begin{cases} \frac{\hat{I}_{\mathrm{transverse}}(n) + \hat{I}_{\mathrm{sagittal}}(n) + \hat{I}_{\mathrm{coronal}}(n)}{3}, & if\ d(n) < d_0 \\ \frac{\hat{I}_{\mathrm{sagittal}}(n) + \hat{I}_{\mathrm{coronal}}(n)}{2}, & otherwise \end{cases} \tag{5}$$

where d_0 is a fixed threshold, as previously recommended by Ding et al.[39] If $d(n) < d_0$, we assume that the three contours at this point are well-matched, the final segmentation is refined by the average of the three contours. Otherwise, we assume that the three contours at this point are not well-matched, and the final segmentation is refined by the average of contours generated from sagittal and coronal planes. Finally, the refined segmentation around apex and base is obtained by mesh smoothing on these points.

2.E. Dataset and quantitative measurements

We tested the proposed method using 44 patients' TRUS data. All were acquired using a Hitachi ultrasound scanner with a 7.5-MHz biplane probe. Each 3D TRUS image is composed of $1024 \times 768 \times 216$ voxels. The voxel size is $0.12 \times 0.12 \times 1.0$ mm^3. We used leave-one-out cross-validation methods to evaluate the proposed segmentation algorithm. Specifically, we excluded one patient from dataset for training our deep learning-based segmentation model. During training stage, one fivefold cross-validation was used to train the model, that is, the random selected 80% patch samples were used to train the model, and the rest 20% patch samples

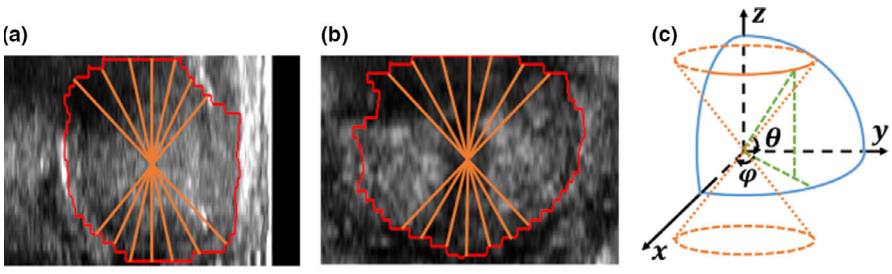


FIG. 3. The visual result of polar rays' distribution. (a) and (b) Polar rays' distribution by orange line in sagittal plane and coronal plane, where the red lines show the manual contour of prostate. (c) Distribution of polar rays in both polar coordinate and Cartesian coordinate. The display window for (a) and (b) is [0, 200].

were used for validation. After training, this excluded patient's TRUS image was used for segmentation test. Our segmentation results were compared with the manually created contours. All manual prostate glands contours were created on TRUS images by an experienced physician. We calculated the DSC, precision score, recall score, Hausdorff distance (HD), mean surface distance (MSD), and the residual mean square distance (RMSD) between the two contours to evaluate the accuracy of our segmentation method. The DSC, precision, and recall scores are used to quantify volume similarity between two contours. The HD, MSD, and RMSD metrics are used to quantify boundary similarity between two surfaces. Generally, more accurate segmentation results are associated with lower HD, MSD, and RMSD scores and higher DSC, precision, and recall scores. Due to low contrast, the most challenging regions in TRUS prostate segmentation are the base and apex. We also conducted a regional analysis to measure errors in the base and apex sections of the prostate.

In order to illustrate the significant improvement of our proposed step-by-step enhancement, a paired two-tailed t-test were used for comparison of the outcomes between two numerical results groups calculated from all patients' data. To furfure evaluate the performance of matrices (DSC, precision, recall, HD, MSD, and RMSD), we computed the corrected p-values through the Holm-Bonferroni method, for which we set $\alpha < 0.05$.[40]

2.F. Parameter performance

In general, segmentation performance can be improved by one of several parameters: larger patch size, larger batch size, and more epochs. There is a tradeoff between segmentation performance and the computation complexity and memory requirements raised by these parameters. In order balance competing demands, we empirically set the initial patch size as $256 \times 256 \times 4$, the number of epochs to 180, and the batch size to 40.

To test the influence of weighting parameter ρ and balancing parameter μ in our proposed hybrid loss function, we fixed the parameters setting, detailed in Appendix Table S1, and tested these two parameters. Fourfold cross-validation was used for this evaluation. Appendix Fig. S1 plots the averaged DSC as a function of these two parameters, which illustrates that $\rho = 0.8$ and $\mu = 2$ are adequate for our TRUS prostate segmentation. We set the number of rays as 360 to make the refined contours smooth and more detailed at boundaries.

3. RESULTS

3.A. Comparison between multiderivative-based and single image input

To evaluate the influence of multiderivative-based image input, we compared the convergence of our proposed algorithm using the original image and the multiderivative-based images as an input. Appendix Fig. S2 shows the average DSC and loss convergence curve of these two methods. The mean DSC based on a batch converges faster with the multiderivative-based input, especially if the epoch value is <40. Figure S2 also shows the best DSC and loss, and their epoch number of these two algorithms. The epoch number with the best DSC's and loss of a multiderivative-based input is much smaller than with a single input.

Compares the segmentation produced with the proposed method using a single (original image) or multiderivative-based image input. Because more informative and structural features can be more easily captured from a multiderivative-based input than an original image input, the final segmented prostate [Appendix Fig. S3(b2)] and the prostate probability maps at high-, modest-, and low-resolution stages of multiderivative-based image input are closer to the manual contour (ground truth) compared with those of an original image input alone. Table I quantitatively compares 44 patients' data based on leave-one-out cross-validation, showing the results of multiderivative input are better than the single input in precision, recall, HD, and RMSD.

3.B. Contribution of deep supervision

To demonstrate the efficiency of deep supervision, we compared the results of a V-Net and our proposed algorithm without contour refinement, that is, deeply supervised V-Net (DS-V-Net). The aim was to show the contribution in three different aspects: (a) the segmented prostate probability maps of four stages, (b) the 3D scatter plot of the first three principle components of randomly selected patches from the probability maps at each stage, and (c) the batch-based mean DSC convergence.

As shown in Appendix Fig. S4, the probability maps of each stage for the apex, middle, and base of the prostate generated by DS-V-Net can identify the prostate boundary with deep supervision. In contrast, the probability maps generated by the traditional V-Net cannot do the same, especially for the high-resolution stage.

Figure 4 shows 3D scatter plots of the first three principle components of patch samples in the probability map of

TABLE I. Quantitative comparison of the proposed deep supervised V-Net with multiderivative-based and single input.

Metric	DSC	Precision	Recall	HD (mm)	MSD (mm)	RMSD (mm)
Original	0.908 ± 0.030	0.893 ± 0.058	0.927 ± 0.046	3.911 ± 1.558	0.605 ± 0.228	0.904 ± 0.377
Multiderivative-based	0.912 ± 0.026	0.897 ± 0.056	0.930 ± 0.043	3.996 ± 1.560	0.607 ± 0.228	0.907 ± 0.377
P-value	0.078	0.012	0.041	<0.001	0.076	0.038

DSC, Dice similarity coefficient; HD, Hausdorff distance; MSD, mean surface distance; RMSD, residual mean surface distance.

each stage. We randomly selected 4000 samples from the prostate region as well as 4000 samples from a non-prostate region around the prostate boundary, as shown in Fig. 4(a2). From the scatter plots of a V-Net at final, high-, modest-, and low-resolution stages, as shown in Figs. 4(b1)–4(e1), there can be seen a large overlap between the samples from prostate and non-prostate regions. It is difficult to directly separate the prostate and non-prostate samples using a traditional V-Net. Whereas in Figs. 4(b2)–4(e2), the prostate and non-prostate samples can be approximately separated by a plane, demonstrating the benefit of using a deep supervision strategy.

The batch-based mean DSC was used as a metric to compare the convergence of a V-Net and our DS-V-Net (shown in Appendix Fig. S5). The epoch values that our DS-V-Net reaches its best DSC are much smaller than the corresponding values that the V-Net reaches its best DSC for both the training and validation folds. This demonstrates that the deep supervision strategy can accelerate the training convergence of our deep-learning-based segmentation. We did not use loss as a metric, because the loss function of a V-Net is computed only in the final stage, whereas the loss function of our DS-V-Net is calculated by the summation of all four stages' loss. Table II quantitatively compares our DS-V-Net vs the V-Net based on the leave-one-out cross-validation of 44 patients' data. As shown in Table II, our DS-V-Net significantly improves the DSC, precision, HD, and RMSD over that of the V-Net.

3.C. Comparison of the loss function

In order to compare the influence of different loss functions, we compared the proposed DS-V-Net with the Dice, the BCE, and the hybrid loss functions. Appendix Fig. S6 shows a segmentation comparison from our DS-V-Net based on three different loss functions in transverse, sagittal and coronal planes. The segmented prostate using the hybrid loss function most closely resembles the manual segmentation. Moreover, since the manual segmentation is delineated on transverse slices, it can have some gradient and sharp shape regions, as seen in the manual contours in sagittal and coronal planes [Appendix Figs. S6(a4) and S6(a6)]. The segmentation results of our DS-V-Net based on the hybrid loss function [Appendix Figs. S6(b6) and S6(b9)] can also smooth the edge of prostate contour, whereas DS-V-Net based on a BCE or Dice loss function cannot, especially for a BCE loss function alone. Appendix Table S2 quantitatively compares the performance of the proposed DS-V-Net with that of three different loss functions, showing that our DS-V-Net with the hybrid loss

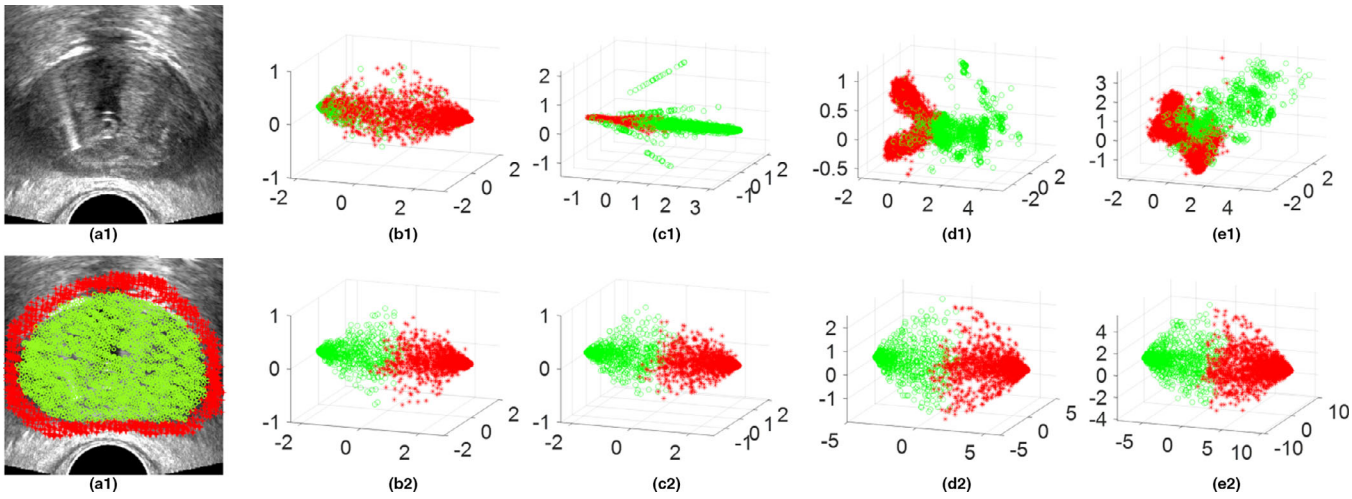


FIG. 4. An illustrative example of the benefit of our deeply supervised V-Net compared with V-Net without deeply supervision, (a1) Transrectal ultrasound (TRUS) image in transverse plane. (a2) Sample patches' central positions drawn from test TRUS images, where the samples belonging to the prostate are highlighted by green circles and the samples belonging to the non-prostate are highlighted by red asterisks. (b1–e1) Scatter plots of the first three principle components of corresponding patch samples in probability map at the final, high-resolution, modest-resolution, and low-resolution stages by using a V-Net, respectively. (b2–e2) Scatter plots of first three principle components of corresponding patch samples in the probability maps at the final, high-resolution, modest-resolution, and low-resolution stages by using our deeply supervised V-Net (DS-V-Net), respectively. The position of the viewer in (b1–e1) and (b2–e2) is azimuth = 20°and elevation = 20°.

TABLE II. Quantitative comparison of the proposed deep supervised V-Net vs a V-Net without deep supervision.

Metric	DSC	Precision	Recall	HD (mm)	MSD (mm)	RMSD (mm)
V-Net	0.905 ± 0.030	0.881 ± 0.060	0.935 ± 0.043	4.643 ± 1.926	0.657 ± 0.270	0.977 ± 0.377
DS-V-Net	0.912 ± 0.026	0.897 ± 0.056	0.930 ± 0.043	3.996 ± 1.560	0.607 ± 0.228	0.907 ± 0.377
P-value	<0.001	<0.001	0.561	0.001	0.007	0.021

DSC, Dice similarity coefficient; DS-V-Net, deeply supervised V-Net; HD, Hausdorff distance; MSD, mean surface distance; RMSD, residual mean surface distance.

function slightly better than DS-V-Net with BCE or DSC loss functions. Thus, we adopted the hybrid loss as the loss function for our method.

3.D. Contribution of multidirectional-based contour refinement

The main challenge in TRUS prostate segmentation is to accurately delineate the prostate at the apex and base. We compared segmentation results from our proposed DS-V-Net with and without contour refinement to evaluate the influence of multidirectional-based refinement, especially at the prostate apex and base. Figure 5 shows our segmentation results with and without a multidirectional-based refinement. The segmentation generated from only the transverse plane has discontinuous boundaries around the apex and base regions [Fig. 5(c1)–5(c2)], whereas the segmentation generated by sagittal and coronal planes can keep the continuity of the contour boundary [Figs. 5(d1)–5(d2), 5(e1)–5(e2)]. In addition, as shown in Figs. 5(a3)–5(f3) and 5(a4)–5(f4), the boundaries of manual contour are not well-matched to the original TRUS images, thus the segmentation results based on any one of three planes introduce ambiguous boundaries. Even in these situations, the proposed DS-V-Net with a contour refinement (DS-CR-V-Net) can still maintain a reasonable and smooth boundary. Table III quantitatively compares segmentation results with and without contour refinement based on leave-one-out cross-validation of 44 patients' data.

3.E. Comparison with state-of-art methods

In order to evaluate and verify the performance of our proposed method, we compared its performance against state-of-the-art prostate segmentation algorithms based on U-Net,[36] V-Net,[26] deeply supervised U-Net (DS-U-Net),[33] and CNNs.[32] Figure 6 compares the segmentation results between these methods and our proposed DS-CR-V-Net. In Figs. 6(a1)–6(g1), five algorithms have similarly segmented prostate in a high-contrast TRUS image. However, in a low-contrast TRUS image, the segmented prostate from our DS-CR-V-Net much more closely represents the real prostate than comparison methods as shown in Figs. 6(a2)–6(g2). Furthermore, as shown in Figs. 6(a3)–6(g3), 6(a4)–6(g4), all the comparison methods could not accurately identify the prostate boundary at the apex and base with a low contrast in prostate TRUS images, while our proposed DS-CR-V-Net with a contour refinement model, demonstrating that our method outperforms these four state-of-art deep learning segmentation algorithms.

Table IV shows a quantitative metrics comparison of prior methods as well as our DS-CR-V-Net based on the leave-one-out cross-validation. All the comparing algorithms were performed using their best parameter setting. As shown in Table V, there is a significant improvement on all metrics between our proposed DS-CR-V-Net method over the V-Net and DS-U-Net methods. Table V also shows the corrected *P*-value calculated through the Holm-Bonferroni method[40] with p-values obtained by comparison between our method and four other methods. Any corrected *P*-value less than alpha (0.05) is significant. A binary vector h has the same

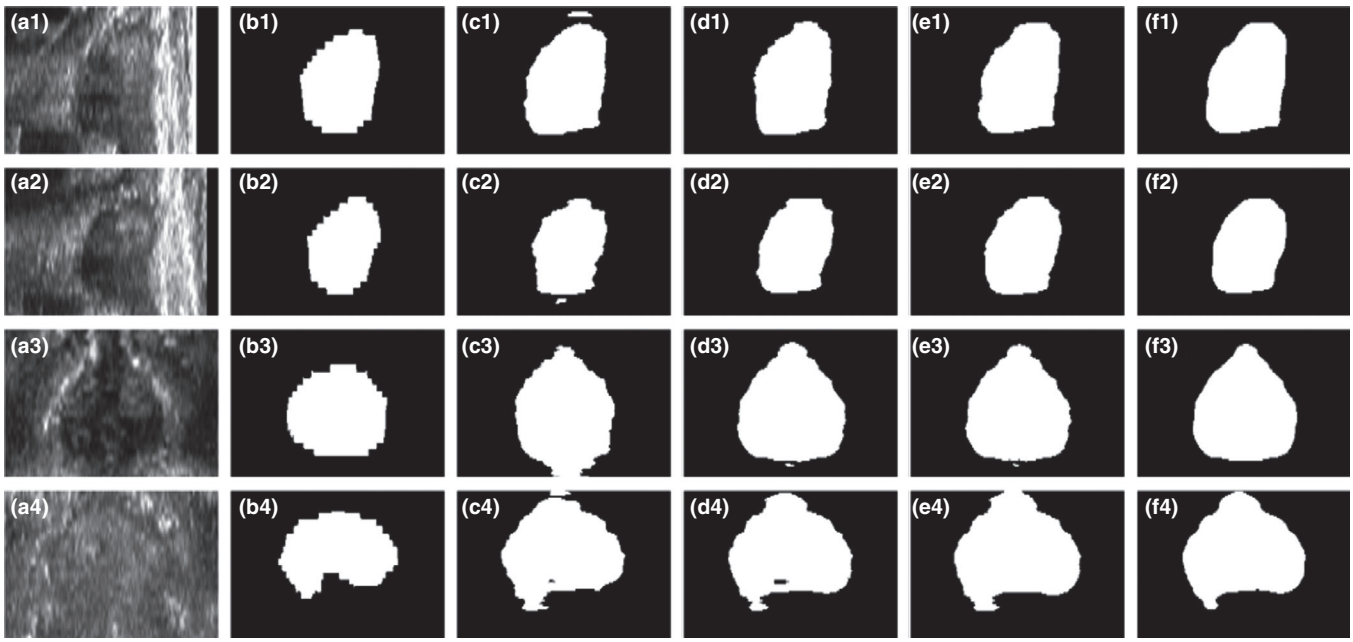


FIG. 5. Comparison of segmented prostates with and without multidirectional-based refinement. (a1–a2) Transrectal ultrasound (TRUS) images shown in sagittal plane, (b1–f1) and (b2–f2) corresponding manually contoured prostates, the segmented prostate from transverse, coronal, and sagittal planes, and the segmented prostate with the multidirectional-based refinement, respectively. (a3–a4) TRUS images shown in coronal plane, (b3–f3) and (b4–f4) Corresponding manual contour, the segmented prostate from only transverse, coronal, and sagittal planes, and the segmented prostate with the refinement, respectively. The display window size of (a1–a4) is [0, 200].

TABLE III. Quantitative metrics comparison with and without contour refinement.

Metric	DSC	Precision	Recall	HD (mm)	MSD (mm)	RMSD (mm)
Transverse	0.912 ± 0.026	0.897 ± 0.056	0.930 ± 0.043	3.996 ± 1.560	0.607 ± 0.228	0.907 ± 0.377
Sagittal	0.914 ± 0.025	0.897 ± 0.057	0.931 ± 0.043	3.981 ± 1.573	0.607 ± 0.228	0.907 ± 0.378
Coronal	0.916 ± 0.028	0.899 ± 0.059	0.930 ± 0.044	3.999 ± 1.546	0.606 ± 0.228	0.907 ± 0.376
Our DS-CR-V-Net	0.919 ± 0.028	0.906 ± 0.055	0.938 ± 0.043	3.938 ± 1.550	0.599 ± 0.225	0.900 ± 0.377
P-value (Our *vs* Trans)	<0.001	<0.001	<0.001	<0.001	0.002	<0.001
P-value (Our *vs* Sag)	0.004	<0.001	<0.001	0.009	0.002	<0.001
P-value (Our *vs* Cor)	0.064	<0.001	<0.001	<0.001	<0.001	<0.001

DSC, Dice similarity coefficient; DS-CR-V-Net, deeply supervised contour refinement V-Net; HD, Hausdorff distance; MSD, mean surface distance; RMSD, residual mean surface distance.

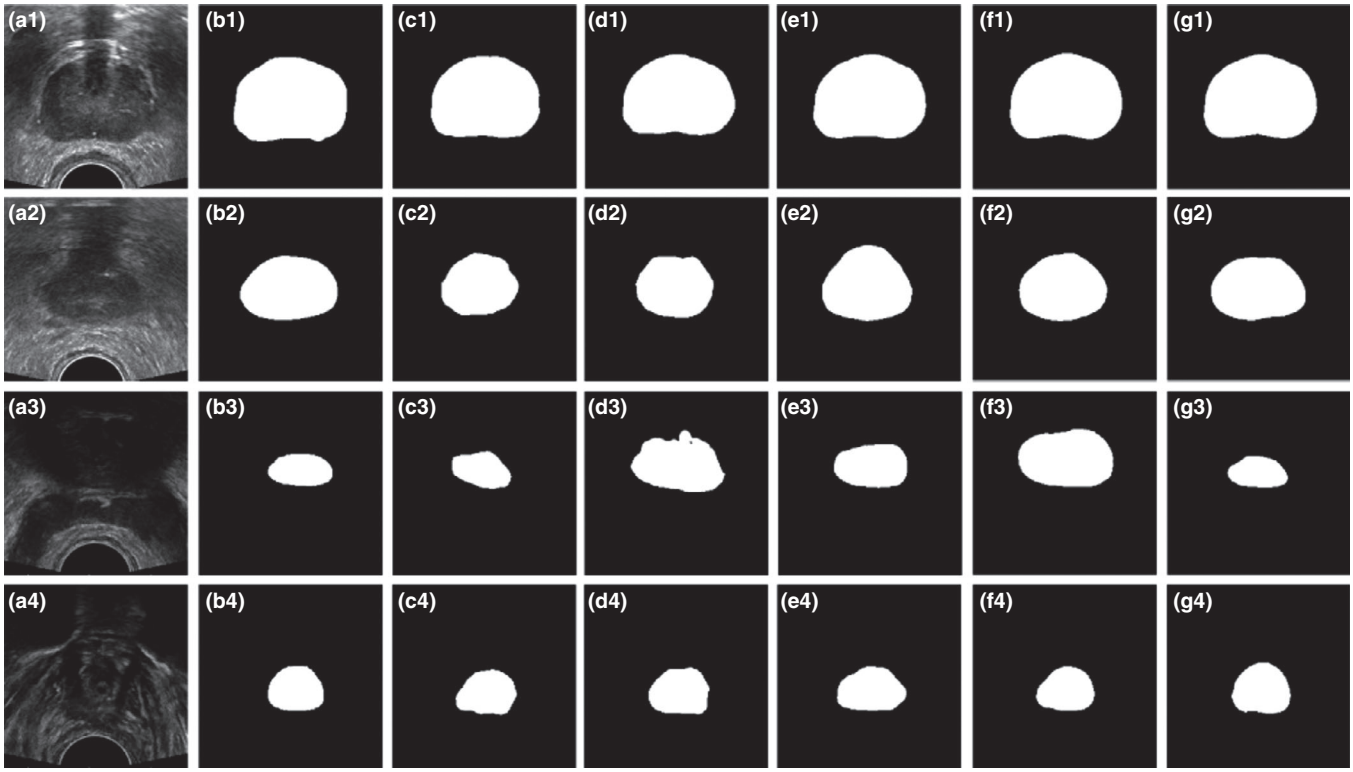


FIG. 6. Segmentation comparison between the proposed deeply supervised contour refinement V-Net (DS-CR-V-Net) and comparing methods. (a1–g1) High-contrast transrectal ultrasound (TRUS) image of the mid-prostate in the transverse plane, the corresponding manual contour, segmented prostates using U-Net, V-Net, convolutional neural networks (CNNs), deeply supervised U-Net (DS-U-Net), and our DS-CR-V-Net algorithms, respectively. (a2–g2) Low-contrast TRUS image in mid-prostate on the transverse plane, the corresponding manual contour, segmented prostates using U-Net, V-Net, CNNs, DS-U-Net, and our DS-CR-V-Net algorithms, respectively. (a3–g3) TRUS image at the prostate apex, the corresponding manual contour, segmented prostates using U-Net, V-Net, CNNs, DS-U-Net, and our DS-CR-V-Net algorithms, respectively. (a4–g4) TRUS image at the prostate base, the corresponding manual contour, segmented prostate generated by U-Net, V-Net, CNNs, DS-U-Net, and our DS-CR-V-Net algorithms, respectively. The display windows for (a1–a4) are [0, 200].

dimensionality as corrected *P*-values. If the ith element of h is 1, then the ith *P*-value is significant. As can be shown in Table V, all *P*-values are significant in DSC.

We also conducted a regional analysis to further demonstrate the improvement comparing with other state-of-the-art methods, which includes U-Net,[36] V-Net,[26] CNNs,[32] and DS-U-Net.[33] We varied the number of cropped base (superior) and apex (inferior) slices from 1 to 10 and measured the segmented contour accuracy within these regions. Dice similarity coefficient and HD were used to compare our proposed method to the comparator methods, as shown in Appendix Tables S3 and S4. It is shown that with respect to mean DSC of the base and apex regions, our method outperforms other methods.

Table VI show the p-values of the comparison between our method and other methods in base and apex regions, where the number of slices was four. Table VI also shows the corrected *P*-values calculated through the Holm-Bonferroni method[40] with *P*-values obtained by comparison between our method and four other methods. As can be shown in Table VI, all p-values are significant in DSC of both base and apex, and in HD of base region.

TABLE IV. Quantitative metrics comparison of our proposed algorithm vs state-of-the-art methods.

Metric	DSC	Precision	Recall	HD (mm)	MSD (mm)	RMSD (mm)
U-Net	0.906 ± 0.028	0.905 ± 0.062	0.912 ± 0.049	4.437 ± 2.010	0.619 ± 0.220	0.915 ± 0.346
V-Net	0.905 ± 0.030	0.881 ± 0.060	0.935 ± 0.035	4.643 ± 1.926	0.657 ± 0.270	0.977 ± 0.410
CNNs	0.901 ± 0.032	0.891 ± 0.069	0.910 ± 0.072	4.391 ± 1.788	0.711 ± 0.315	1.043 ± 0.479
DS-U-Net	0.911 ± 0.03	0.901 ± 0.05	0.926 ± 0.05	3.963 ± 1.51	0.599 ± 0.21	0.892 ± 0.33
DS-CR-V-Net	0.919 ± 0.028	0.906 ± 0.055	0.938 ± 0.043	3.938 ± 1.550	0.599 ± 0.225	0.900 ± 0.377

CNNs, convolutional neural networks; DSC, Dice similarity coefficient; DS-CR-V-Net, deeply supervised contour refinement V-Net; HD, Hausdorff distance; MSD, mean surface distance; RMSD, residual mean surface distance.

TABLE V. *P*-values, corrected *P*-values, and h obtained by comparing our proposed algorithm with state-of-the-art methods in whole prostate region.

	DSC	Precision	Recall	HD	MSD	RMSD
P-value						
U-Net	<0.001	0.854	<0.001	0.063	0.423	0.689
V-Net	<0.001	<0.001	0.561	0.002	0.007	0.021
CNNs	<0.001	0.171	0.004	0.854	0.991	0.771
DS-U-Net	<0.001	0.009	0.005	0.013	<0.001	0.001
Corrected *P*-value(h)						
U-Net	*<0.001 (**1**)*	*0.855 (**0**)*	*<0.001 (**1**)*	*0.126 (**0**)*	*0.846 (**0**)*	*1.378 (**0**)*
V-Net	*<0.001 (**1**)*	*<0.001 (**1**)*	*0.561 (**0**)*	*0.008 (**1**)*	*0.022 (**1**)*	*0.063 (**0**)*
CNNs	*<0.001 (**1**)*	*0.341 (**0**)*	*0.011 (**1**)*	*0.854 (**0**)*	*0.991 (**0**)*	*1.378 (**0**)*
DS-U-Net	*<0.001 (**1**)*	*0.028 (**1**)*	*0.011 (**1**)*	*0.039 (**1**)*	*0.002 (**1**)*	*0.005 (**1**)*

CNNs, convolutional neural networks; DSC, Dice similarity coefficient; DS-CR-V-Net, deeply supervised contour refinement V-Net; HD, Hausdorff distance; MSD, mean surface distance; RMSD, residual mean surface distance. The italics format is used to show the P-value through paired two-tailed t-test and corrected P-value through Holm-Boneferroni method, the italics and bold format is used to show the associated hypotheses for it corresponding corrected P-value, where 1 denoting significant, 0 denoting not significant.

TABLE VI. *P*-values, corrected *P*-values, and h obtained by comparing our proposed algorithm with state-of-the-art methods in low-contrast prostate region.

	DSC base	DSC apex	HD base	HD apex
P-value				
U-Net	0.014	0.001	<0.001	0.025
V-Net	0.009	0.042	0.010	0.031
DS-U-Net	0.018	0.002	0.003	0.031
CNNs	0.013	0.002	<0.001	0.149
Corrected *P*-value (*h*)				
U-Net	*0.040 (**1**)*	*0.004 (**1**)*	*<0.001 (**1**)*	*0.099 (**0**)*
V-Net	*0.037 (**1**)*	*0.042 (**1**)*	*0.010 (**1**)*	*0.092 (**0**)*
DS-U-Net	*0.029 (**1**)*	*0.005 (**1**)*	*0.006 (**1**)*	*0.099 (**0**)*
CNNs	*0.040 (**1**)*	*0.005 (**1**)*	*<0.001 (**1**)*	*0.149 (**0**)*

CNNs, convolutional neural networks; DSC, Dice similarity coefficient; DS-U-Net, deeply supervised U-Net; HD, Hausdorff distance; MSD, mean surface distance; RMSD, residual mean surface distance. The italics format is used to show the P-value through paired two-tailed t-test and corrected P-value through Holm-Boneferroni method, the italics and bold format is used to show the associated hypotheses for it corresponding corrected P-value, where 1 denoting significant, 0 denoting not significant.

3.F. Inter- and intraobserver reliability

In order to test the reliability of manual contours, we conducted an inter- and intraobserver reliability study with nine patients. The interobserver study was performed by three physicians contouring on same patients separately. The intraobserver study was performed by one physician contouring on the same patients separated by an interval of 3 weeks. The volume percentage difference among observer 1 (O1-1), 2 (O2), and 3 (O3) and our segmented contour were measured for interobserver reliability. The volume percentage difference between two-time contours of the observer 1 (O1-1, O1-2), and our segmented contour were measured for intraobserver reliability. Whole region, base, and apex were used to evaluate. The metrics of these comparisons are shown in Appendix Fig. S7.

Comparing the manual segmentations by three interobservers and our segmentation, the volume percentage difference variance of our vs. three observers' manual contour (whole region 3.12% ± 3.82%, base 16.95% ± 19.67%, apex 19.51% ± 19.15%) is smaller than that among three interobservers (whole region 4.46% ± 6.40%, base 27.18% ± 27.72%, apex 31.81% ± 25.31%), which demonstrated our method partially reduces interobserver variation. Comparing the two manual segmentations by same observer and our segmentation, the volume percentage difference variance of our vs intraobservers' manual contour (whole region 3.46% ± 4.01%, base 11.75% ± 15.76%, apex 15.82% ± 15.35%) is smaller than that among two intraobservers (whole region 5.14% ± 4.59%, base 20.83% ± 23.03%, apex 27.84% ± 23.26%), which demonstrated our method partially reduces intraobserver variation. In

prostate whole region, the inter- and intraobserver reliability study showed the consistency in the manual segmentations.

4. DISCUSSION

We proposed a new prostate segmentation method which incorporates a deep supervision strategy and multidirectional-based contour refinement into a V-Net architecture to automatically segment the prostate on TRUS images. Our proposed method was evaluated against state-of-the-art deep learning networks. As shown in Fig. 6 and Table IV, our proposed method outperformed these two both qualitatively and quantitatively. Ghavami et al. have reported their CNNs-based TRUS prostate segmentation results with 10-fold patient-level cross-validation.[32] The DSC and MSD of 4055 2D TRUS images were 0.91 ± 0.12 and 1.23 ± 1.46 mm, respectively. The DSC of 110 3D TRUS images was 0.91 ± 0.04. Our final results show the DSC and MSD of 44 3D TRUS images were 0.92 ± 0.03 and 0.60 ± 0.23 mm, respectively. Considering the different TRUS database used in our and Ghavami's paper, we applied the Ghavami's CNNs into our TRUS database with the best performing parameter settings, as shown in Fig. 6 and Table IV. Our proposed method outperformed this method, using the same TRUS database. In addition, the superior performance of our proposed method was further demonstrated at the apex and base regions from our region analysis study as shown in Table VI.

There are several limitations to our current methods. First, our ground truth of prostate volume is from the manual contours by physicians. These manual contours may have systematic errors and random errors. Our proposed method may mitigates random errors. However, the systematic errors (e.g., physician's contouring style) will affect our final segmented results, but this would be expected to be a limitation to all learning-based methods. Second, the computation complexity is higher than the state-of-the-art algorithms due to three more stages of deep supervision and corresponding upsampling convolutional kernels. In our leave-one-out experiments, the training time for a U-Net and V-Net are 1.35 and 1.70 h, whereas our proposed algorithm requires 1.85 h. However, after training, our proposed algorithm has similar segmentation times compared to these two algorithms. A single prostate segmentation can be completed in 1–2 s. All the algorithms were implemented in Tensorflow with Adam optimizer and were trained and tested on a NVIDIA TITAN XP GPU with 12 GB of memory. Third, it may require the introduction of an adaptive and nonlinear contour refinement model when the prostate surface of the three methods from three different directions are not well matched. Improving contour refinement, incorporating the conditional random field, is a direction of for our future segmentation work. We also plan to test the robustness and reliability of our TRUS prostate segmentation using more patients' data before we apply our method to our ultrasound-guided prostate-cancer radiotherapy procedure. In addition, the size of the patient population is relative small. Evaluating the clinical utility of the proposed method on more patients' data will be our future work. With limited patient datasets available as training samples, data augmentation is essential to train the network with the desired invariance and robustness properties. By artificially increasing the number of training examples, it helps reducing overfitting and improves generalization. Especially, for the computed tomography prostate segmentation, the prostates' size, shape, and position vary significantly among individuals. Data augmentation can help introducing more such training data diversity. In our implementation, data augmentation was applied during training by flipping the images left and right in transverse plane, rotating images by 90°, 180°, 270° in transvers plane, as well as random elastic deforming training TRUS images and corresponding contour binary masks. For deforming images, as recommended in previous deep learning-based study,[36] we generated smooth deformations using random displacement vectors on a coarse 3 by 3 grid. The displacements were sampled from a Gaussian distribution with 10 pixels standard deviation. Per-pixel displacements are then computed using bicubic interpolation. In our experience, the random elastic deformations of the training data are the key concept to train a segmentation network with limited number of annotated images.

Our proposed method does not need any manual intervention in segmentation stage. We set input patch size as $512 \times 512 \times 4$, which is a group of sequent slices within a TRUS image volume. The patches were automatically extracted by sliding the volume with overlap $0 \times 0 \times 2$. We used mean patch fusion to reconstruct the probability map, and then use a threshold to get the binary mask.

5. CONCLUSIONS

We developed a novel 3D deeply supervised deep learning-based approach to automatically segment the TRUS prostate. A 3D deep supervision strategy has been utilized to address limitations from the size of the training dataset and low contrast challenges in TRUS prostate segmentation. A multidirectional-based contour refinement was used as a post-processing to refine the segmentation results. Experimental validation was performed to demonstrate its clinical feasibility and segmentation accuracy. When considering the whole prostate, the improvement of the proposed method is mostly in terms of the Dice coefficient. In low-contrast regions, the comparison between our method and other method is significant in DSC of both base and apex regions, and in HD of base region. This segmentation technique could be a useful tool for image-guided interventions in prostate cancer diagnosis and treatment. The role of multidirectional-based refinement in the post-processing of TRUS contour is expected to continue to grow as increasingly complex technological challenges and associated robustness concerns need to be addressed.

ACKNOWLEDGMENTS

This research is supported in part by the National Cancer Institute of the National Institutes of Health under Award

Number R01CA215718 (XY), the Department of Defense (DoD) Prostate Cancer Research Program (PCRP) Award W81XWH-13-1-0269 (XY), DoD W81XWH-17-1-0438 (TL), and W81XWH-17-1-0439 (AJ), and Dunwoody Golf Club Prostate Cancer Research Award, a philanthropic award provided by the Winship Cancer Institute of Emory University. We are also grateful for the GPU support from NVIDIA Corporation.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

[a)]Author to whom correspondence should be addressed. Electronic mail: xyang43@emory.edu**;** Telephone: (404)-778-8622.

REFERENCES

1. Martin S, Daanen V, Troccaz J. Atlas-based prostate segmentation using an hybrid registration. *Int J Comput Assist Radiol Surg*. 2008;3:485–492.
2. Sarkar S, Das S. A review of imaging methods for prostate cancer detection. *Biomed Eng Comput Biol*. 2016;7:1–15.
3. Yu YY, Chen YM, Chiu B. Fully automatic prostate segmentation from transrectal ultrasound images based on radial bas-relief initialization and slice-based propagation. *Comput Biol Med*. 2016;74:74–90.
4. Mahdavi SS, Chng N, Spadinger I, Morris WJ, Salcudean SE. Semi-automatic segmentation for prostate interventions. *Med Image Anal*. 2011;15:226–237.
5. Ghose S, Oliver A, Martí R, et al. A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images. *Comput Meth Prog Bio*. 2012;108:262–287.
6. Yan P, Xu S, Turkbey B, Kruecker J. Discrete deformable model guided by partial active shape model for TRUS image segmentation. *IEEE Trans Biomed Eng*. 2010;57:1158–1166.
7. Betrouni N, Vermandel M, Pasquier D, Maouche S, Rousseau J. Segmentation of abdominal ultrasound images of the prostate using a priori information and an adapted noise filter. *Comput Med Imag Grap*. 2005;29:43–51.
8. Kachouie NN, Fieguth P, Rahnamayan S. An elliptical level set method for automatic TRUS prostate image segmentation. *Proc of IEEE International Symposium on Signal Processing and Information Technology*; 2006:191–196. https://doi.org/10.1109/ISSPIT.2006.270795
9. Nouranian S, Mahdavi SS, Spadinger I, Morris WJ, Salcudean SE, Abolmaesumi P. A multi-atlas-based segmentation framework for prostate brachytherapy. *IEEE Trans Med Imaging*. 2015;34:950–961.
10. Yang X, Wu N, Cheng G, et al. Automated segmentation of the parotid gland based on atlas registration and machine learning: a longitudinal MRI study in head-and-neck radiation therapy. *Int J Radiat Oncol Biol Phys*. 2014;90:1225–1233.
11. Zouqi M, Samarabandu J. Prostate segmentation from 2-D ultrasound images using graph cuts and domain knowledge. Proc of Canadian Conference on Computer and Robot Vision; 2008:359–362.
12. Egger J. PCG-Cut: graph driven segmentation of the prostate central gland. *PLoS ONE*. 2013;8:e76645
13. Gong LX, Ng L, Pathak SD, et al. Prostate ultrasound image segmentation using level set-based region flow with shape guidance. *Proc SPIE*. 2005;5747:1648–1657.
14. Yu YY, Cheng JY, Li JZ, Chen WF, Chiu B. Automatic prostate segmentation from transrectal ultrasound images. *Proc of IEEE Biomedical Circuits and Systems Conference;*2014:117–120. https://doi.org/10.1109/BioCAS.2014.6981659
15. Richard WD, Keen CG. Automated texture-based segmentation of ultrasound images of the prostate. *Comput Med Imag Grap*. 1996;20:131–140.
16. Tutar IB, Pathak SD, Gong L, Cho PS, Wallner K, Kim Y. Semiautomatic 3-D prostate segmentation from TRUS images using spherical harmonics. *IEEE Trans Med Imaging*. 2007;25:1645–1654.
17. Qiu W, Yuan J, Ukwatta E, Fenster A. Rotationally resliced 3D prostate TRUS segmentation using convex optimization with shape priors. *Med Phys*. 2015;42:877–891.
18. Nouranian S, Ramezani M, Spadinger I, Morris WJ, Salcudean SE, Abolmaesumi P. Learning-based multi-label segmentation of transrectal ultrasound images for prostate brachytherapy. *IEEE Trans Med Imaging*. 2016;35:921–932.
19. Anas E, Mousavi P, Abolmaesumi P. A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy. *Med Image Anal*. 2018;48:107–116.
20. Yang XF, Fei BW. 3D prostate segmentation of ultrasound images combining longitudinal image registration and machine learning. *Proc SPIE*. 2012;8316:83162O.
21. Yang XF, Schuster D, Master V, Nieh P, Fenster A, Fei BW. Automatic 3D segmentation of ultrasound images using atlas registration and statistical texture prior. *Proc SPIE*. 2011;7964:796432.
22. Yang X, Rossi P, Jani A, Mao H, Curran W, Liu T. 3D transrectal ultrasound (TRUS) prostate segmentation based on optimal feature learning framework. *Proc SPIE*. 2016;9784:97842F.
23. Akbari H, Yang X, Halig LV, Fei B. 3D segmentation of prostate ultrasound images using wavelet transform. *Proc SPIE*. 2011;7962:79622K.
24. Ghose S, Mitra J, Oliver A, et al. A supervised learning framework for automatic prostate segmentation in trans rectal ultrasound images. *Advanced Concepts for Intelligent Vision Systems (Acivs. 2012)*; 2012;7517:190–200.
25. Khellaf F, Leclerc S, Voorneveld JD, Bandaru RS, Bosch JG, Bernard O. Left ventricle segmentation in 3D ultrasound by combining structured random forests with active shape models. *Proc SPIE*. 2018;10574:105740J.
26. Milletari F, Navab N, Ahmadi SA. V-Net: fully convolutional neural networks for volumetric medical image segmentation. *Proceedings of International Conference on 3D Vision*; 2016:565–571. https://doi.org/10.1109/3dv.2016.79
27. Dong X, Lei Y, Wang T, et al. Automatic multiorgan segmentation in thorax CT images using U-net-GAN. *Med Phys*. 2019. 46:2157–2168.
28. Wang B, Lei Y, Tian S, et al. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. *Med Phys*. 2019;46:1707–1718.
29. Wang T, Lei Y, Tang H, et al. A learning-based automatic segmentation and quantification method on left ventricle in gated myocardial perfusion SPECT imaging: a feasibility. *J Nucl Cardiol*. 2019. https://doi.org/10.1007/s12350-019-01594-2
30. Wang T, Lei Y, Tian S, et al. Learning-based automatic segmentation of arteriovenous malformations on contrast CT images in brain stereotactic radiosurgery. *Med Phys*. 2019. https://doi.org/10.1002/mp.13560
31. Yang X, Yu L, Wu L, et al. Recurrent neural networks for automatic prostate segmentation in ultrasound images. *Proc of AAAI*;2017.
32. Ghavami N, Hu YP, Bonmati E, et al. Automatic slice segmentation of intraoperative transrectal ultrasound images using convolutional neural networks. *Proc SPIE*. 2018;10576:1057603.
33. Zhu QK, Du B, Turkbey B, Choyke PL, Yan PK. Deeply-supervised CNN for prostate segmentation. *Proc of International Joint Conference on Neural Networks*;2017:178-184. https://doi.org/10.1109/IJCNN.2017.7965852
34. Zeng Q, Samei G, Karimi D, et al. Prostate segmentation in transrectal ultrasound using magnetic resonance imaging priors. *Int J Comput Assist Radiol Surg*. 2018;13:749–757.
35. Ngiam J, Khosla A, Kim M, Nam J, Lee H, Ng AY.Multimodal deep learning. *Proc of International Conference on International Conference on Machine Learning*; 2011:689–696.
36. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *Proc MICCAI*. 2015;9351:234–241.
37. Postema A, Mischi M, de la Rosette J, Wijkstra H. Multiparametric ultrasound in the detection of prostate cancer: a systematic review. *World J Urol*. 2015;33:1651–1659.

38. Dou Q, Yu LQ, Chen H, et al. 3D deeply supervised network for automated segmentation of volumetric medical images. *Med Image Anal*. 2017;41:40–54.
39. Ding MY, Chiu B, Gyacskov I, et al. Fast prostate segmentation in 3D TRUS images based on continuity constraint using an autoregressive model. *Med Phys*. 2007;34:4109–4125.
40. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6:65–70.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig S1: DSC as a function of weighting parameter ρ and balancing parameter μ in our proposed hybrid loss function.
Fig S2: Convergence of proposed deep supervised V-Net with multi-derivative image and single image input.
Fig S3: Comparison of segmented prostate contours from the proposed deep supervised V-Net with multi-derivative image to single image input.
Fig S4: Comparison of the prostate probability map at each stage.
Fig S5: Batch-based mean DSC convergence of the V-Net and our DS-V-Net.
Fig S6: Segmentation comparison from our DS-V-Net based on three different loss functions.
Fig S7: Inter- and intra-observer reliability of the prostate contours.
Table S1: Default parameter setting.
Table S2: Quantitative metrics comparison of the proposed DSN-V-Net with three different loss functions.
Table S3: DSC of base and apex regions of our proposed algorithm versus state-of-the-art methods.
Table S4: HD of base and apex regions of our proposed algorithm versus state-of-the-art methods.

Medical Image Analysis 67 (2021) 101845

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Biomechanically constrained non-rigid MR-TRUS prostate registration using deep learning based 3D point cloud matching

Yabo Fu[a], Yang Lei[a], Tonghe Wang[a,b], Pretesh Patel[a,b], Ashesh B. Jani[a,b], Hui Mao[b,c], Walter J. Curran[a,b], Tian Liu[a,b], Xiaofeng Yang[a,b,*]

[a] Department of Radiation Oncology, Emory University, 1365 Clifton Road NE, Atlanta, GA 30322, United States
[b] Winship Cancer Institute, Emory University, Atlanta, GA 30322, United States
[c] Department of Radiology and Imaging Sciences, Emory University, Atlanta, GA 30322, United States

ARTICLE INFO

Article history:
Received 20 November 2019
Revised 17 August 2020
Accepted 31 August 2020
Available online 7 October 2020

Keywords:
MR-TRUS
Image registration
Point cloud matching
Finite element
Deep learning

ABSTRACT

A non-rigid MR-TRUS image registration framework is proposed for prostate interventions. The registration framework consists of a convolutional neural networks (CNN) for MR prostate segmentation, a CNN for TRUS prostate segmentation and a point-cloud based network for rapid 3D point cloud matching. Volumetric prostate point clouds were generated from the segmented prostate masks using tetrahedron meshing. The point cloud matching network was trained using deformation field that was generated by finite element analysis. Therefore, the network implicitly models the underlying biomechanical constraint when performing point cloud matching. A total of 50 patients' datasets were used for the network training and testing. Alignment of prostate shapes after registration was evaluated using three metrics including Dice similarity coefficient (DSC), mean surface distance (MSD) and Hausdorff distance (HD). Internal point-to-point registration accuracy was assessed using target registration error (TRE). Jacobian determinant and strain tensors of the predicted deformation field were calculated to analyze the physical fidelity of the deformation field. On average, the mean and standard deviation were 0.94±0.02, 0.90±0.23 mm, 2.96±1.00 mm and 1.57±0.77 mm for DSC, MSD, HD and TRE, respectively. Robustness of our method to point cloud noise was evaluated by adding different levels of noise to the query point clouds. Our results demonstrated that the proposed method could rapidly perform MR-TRUS image registration with good registration accuracy and robustness.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Prostate cancer is the second most common malignant disease in men around the world. Prostate cancer accounts for about 28% of newly diagnosed cancers among men and is the leading cause of death from cancer (Weir et al., 2015). Transrectal ultrasound (TRUS) imaging is the standard imaging modality for cancer diagnosis and image-guided interventions, such as biopsy (Yacoub et al., 2012; Yang et al., 2011; Yang et al., 2015), brachytherapy (Reynier et al., 2004; Yang et al., 2014; Yang et al., 2017) and cryotherapy (Bahn et al., 2012) since TRUS is non-ionizing, widely accessible, inexpensive and real-time. Prostate brachytherapy is regularly used to treat prostate cancer using high-dose-rate (HDR), whole prostate gland irradiation (Challapalli et al., 2012). Dose escalation to the dominant intraprostatic lesion (DIL) after whole prostate irradiation was often used to improve tumor control in the initial site of DIL (Shaaer et al., 2019). However, it is difficult to delineate the tumor and DIL accurately on TRUS images due to the lack of soft tissue contrast.

Multiparametric MRI (mp-MRI) provides detailed anatomical and functional information which is unavailable on TRUS images (Bloom et al., 2018). Due to its superior soft-tissue contrast, mp-MRI has higher sensitivity and specificity for prostate cancer detection and characterization than TRUS (Shaaer et al., 2019; van de Ven et al., 2013). However, MRI is expensive, not as widely accessible as US and, most importantly, impractical to provide real-time image guidance for prostate interventions. It is beneficial to fuse TRUS and mp-MRI images to provide real-time image guidance with improved soft-tissue contrast (Dickinson et al., 2013; Zhang et al., 2015). MR-TRUS image registration could enable targeted biopsy and brachytherapy, which allow precise perineum

* Corresponding author at: Department of Radiation Oncology, Emory University School of Medicine, 1365 Clifton Road NE, Atlanta, GA 30322 United States.
E-mail address: xiaofeng.yang@emory.edu (X. Yang).

https://doi.org/10.1016/j.media.2020.101845
1361-8415/© 2020 Elsevier B.V. All rights reserved.

biopsy needle insertion and brachytherapy catheter placement (Kadoury et al., 2010). In addition, it enables accurate DIL delineation which allows local radiation dose boost to DIL, leading to improved tumor control. Conventionally, MR-TRUS image registration is done manually by physicians. After rigid alignment of MR and US prostates, manual lesion delineation is performed to account for the complex prostate deformation (Shaaer et al., 2019). Prostate deformation could be induced by bladder and rectum filling, patient setup and TRUS probe. Manual registration allows only rigid registration and fails to account for the complex prostate deformation. In addition, manual registration process is physician dependent, time-consuming and irreproducible.

Single modality image registration such as CT-CT and MRI-MRI is often performed by minimizing image intensity differences between the fixed and moving images since images acquired with the same imaging modality have similar image intensities (Fu et al., 2018a). For multi-modality image registration such as CT-MRI registration, there is no direct image intensity correlation. As a result, statistical image texture similarities such as mutual information (MI) are often used in multi-modality image registration (Heinrich et al., 2012; Loeckx et al., 2010). For MRI-TRUS registration, it is very difficult to rely solely on image intensities due to the distinct image appearance between the two. Hence, the prostate shapes segmented from the MR and TRUS images should be utilized to facilitate the registration. Challenges are to segment the prostate with high accuracy and to regularize internal prostate deformation based on only surface-driven deformations. Spatial smoothing is commonly used to regularize the predicted deformation field. However, the over-simplified spatial smoothness constraint does not model the complex prostate deformation well. Accurate motion modelling is needed to account for the large prostate deformation induced by US probe. Therefore, it is important to develop an accurate MR-TRUS image registration method which overcomes the above-mentioned challenges simultaneously, including the lack of effective MR-TRUS image similarity metric and difficulties in accurate prostate segmentations and motion modeling.

2. Related works

Many methods have been proposed for MR-TRUS image registration (Hu et al., 2012; Hu et al., 2018b; Khallaghi et al., 2015a; Khallaghi et al., 2015b; Mitra et al., 2012; Sun et al., 2015; van de Ven et al., 2015; Wang et al., 2016; Wang et al., 2018; Yang et al., 2011; Yang et al., 2015). Generally, these methods can be categorized into three different groups, including intensity-based methods, surface-based methods and deep learning-based methods.

For intensity-based method, MI is often used as the image similarity measure between MR and TRUS. A variety of MI variants such as correlation ratio-based MI (Gong et al., 2017), contextual conditioned MI (Rivaz et al., 2014) have been proposed. However, MI does not model the spatial dependency of local structures well. Later, modality independent neighborhood descriptor (MIND) (Heinrich et al., 2012) was proposed by Heinrich et al. to calculate structural similarity across modalities. MIND accounts for the spatial structural dependency using local self-similarity measures. Sun et al. (2015) used a multi-channel MIND as the image similarity measures for MR-TRUS image registration. Wang et al. (2016) proposed to use MIND as a weighting factor for robust prostate surface point matching. Gabor wavelet was also used for feature extraction from MRI and TRUS images. Ou et al. (2011) proposed to extract multi-scale and multi-orientation Gabor attributes which reflect the anatomical and geometric context around each voxel. The extracted Gabor attributes from MR and TRUS were then matched using a continuously-valued weighting function. Singh et al. (2008) used manually identified landmarks to aid MR-TRUS image registration. Since TRUS provides very limited intraprostatic anatomical information other than the sparse calcifications and cysts, it is very difficult to achieve high MR-TRUS image registration accuracy based solely on image intensity.

For surface-based methods, prostate shapes on MR and TRUS are usually manually segmented in the first step. The shapes of the prostate were then meshed into 3D point cloud for surface matching. Commonly used point cloud matching methods include thin-plate spline robust point matching (TPS-RPM) (Chui and Rangarajan, 2003), iterative closest point (ICP) (Du et al., 2010), coherent point drifting (CPD) (Bernhard et al., 2007) and local feature statistics histogram (LFSH) (Yang et al., 2016). To constrain the deformation of point matching, Gaussian-mixture model (GMM) was used to compute probabilistic point correspondence to avoid binary point correspondence (Jian and Vemuri, 2011). CPD utilized a penalty term to constrain nearby points to move coherently as a group. Deformation of the point cloud could also be regularized by biomechanical models. For instance, Wang et al. developed a personalized statistical deformable model (PSDM) to guide prostate surface point matching (Wang et al., 2016). Then, interpolation-based models such as thin-plate spline (Bookstein, 1989) and elastic body splines (Kohlrausch et al., 2005) were utilized to calculate intraprostatic deformation. Though PSDM used in Wang et al.'s paper is a biomechanical model, intraprostatic deformation still needs to be interpolated from surface point deformation using thin-plate spline. Hu et al. (2008; 2012) built a patient-specific statistical motion model using principal component analysis (PCA) by simulating the prostate-US probe interaction using finite element (FE) method. Different from Wang et al.'s model, Hu et al.'s model includes not only prostate surfaces points, but also intraprostatic points. Many FE models with various boundary conditions need to be performed for the same patient in order to build a specific/personalized motion model, which prevents the patient-specific model from being used in large scale. Later, Hu et al. (2015) proposed a population-based prediction of patient-specific motion model. The idea was to utilize kernel regression analysis to describe a separate patient's prostate motion using a multivariate combination of a pre-built population motion models. Khallaghi et al. proposed a biomechanically constrained surface-based MR-TRUS registration method by explicitly integrating a FE volumetric strain energy function into the surface point matching framework (Khallaghi et al., 2015b).

Deep learning-based methods have achieved the-state-of-art performance in many medical image applications, such as segmentation (Dong et al., 2019; Fu et al., 2018b; Fu et al., 2019) and registration (Fu et al., 2020; Haskins et al., 2019b). One advantage of deep learning-based registration is that it enables rapid prediction of the dense deformation vector field (DVF) in a single or a few forward predictions. Yan et al., (2018) used an adversarial deep learning framework to predict an image similarity metric in terms of target registration error (TRE). In Yan et al.'s work, manually aligned images by experts were treated as the ground truth. A discriminator was used to tell whether an image was distorted using the ground truth deformation or the estimated deformation. Haskins et al. (2019a) predicted a deep-learning based MR-TRUS similarity metric, which outperformed MIND similarity metric in MR-TRUS image registration. Only rigid image registration was considered by both Yan et al. and Huskins et al. Onofrey et al. (2016) proposed to synthesize TRUS images from MR images so that the MR-TRUS registration could be cast into TRUS-TRUS single modality registration. Hu et al. (2018b) trained a weakly supervised network for MR-TRUS image registration using label correspondences. Only MR-TRUS image pairs were required in the inference stage since labels were used in the training stage for loss calculation. Deep learning network learns through gradient back propaga-

tion of the loss with respect to the network learnable weights. The label correspondence loss represents only a portion of the true loss that should be defined on the whole prostate. As a result, the network needs to infer the whole prostate deformation based on an under-constrained loss definition. Hu et al. (2018a) recently proposed an adversarial network to distinguish between the network-predicted DVF and the FE-generated DVF. Their results showed that the adversarial network was able to generate physically plausible deformation with only prostate gland segmentations and without any other deformation smoothness regularization.

Due to the unique challenges of MR-TRUS registrations, it is not optimal to directly apply the commonly used convolutional neural networks (CNN) to MR-TRUS image registration. The TRUS images provide very sparse reliable anatomical markers such as urethra, seminal vesicles, cysts and calcifications. It is very challenging to automatically detect and segment these markers on both MRI and US to establish correspondences. In addition, some cases may not show these anatomical markers due to image noise and artifacts. Therefore, it is not effective to train a CNN to directly predict the dense DVF from the MR-TRUS image intensities. In this study, we proposed a shape-based registration where prostate shapes are represented by 3D point clouds. Traditional CNN can only take regular voxel grids dataset such as images as input and cannot directly consume irregular data structures such as 3D point cloud. Fortunately, Qi et al. (2016; 2017) proposed a deep learning framework, PointNet, for classification and segmentation of 3D point cloud. The PointNet was designed to directly consume irregular geometric dataset such as 3D point clouds. Based on PointNet, Liu et al. (2018) proposed a network called FlowNet3D to predict 3D scene flow from 3D points that were generated from stereo and RGB-D images. Meanwhile, Aoki et al. (2019) proposed a network called PointNetLK to combine PointNet and Lucas & Kanade (LK) algorithm for 3D point cloud rigid registration. Rigid transformations were calculated iteratively in PointNetLK. Inspired by these methods, we proposed a new point cloud matching method for MR-TRUS prostate image registration. Different from FlowNet3D, the proposed ProRegNet utilized a weighted surface and volume point cloud loss to balance surface and intraprostatic point cloud matching. For the ease of description, the proposed MR-TRUS prostate registration network is referred to 'ProRegNet'. Prostate shapes from both MR and TRUS were first automatically segmented using our previously developed deep learning-based segmentation networks. The segmented prostate volumes were then meshed into 3D tetrahedron elements to generate volumetric point clouds. Surface point correspondences were established using non-rigid ICP. FE analysis was performed for every case in our datasets using surface nodal correspondence as FE boundary condition. Dense DVFs were generated based on the results of the FE simulation. The dense DVFs were considered as the ground truth deformation for MR-TRUS prostate registration. Volumetric point cloud correspondences were then established based on the dense DVFs. ProRegNet was trained using the prostate volumetric point clouds and their correspondences, which implicitly encoded biomechanical constraints. Compared to previous studies, the contributions of our study include

(1) A novel MR-TRUS image registration network, ProRegNet, was proposed to directly predict volumetric point cloud motion that deforms the MR prostate shape to the TRUS prostate shape.
(2) Trained with FE-simulated point correspondences, ProRegNet was able to predict volumetric prostate point cloud motion with implicit biomechanical constraints.
(3) Deep learning-based MR and TRUS prostate segmentations were adopted to enable fully automatic MR-TRUS image registration.

3. Materials and methods

The proposed registration framework consists of 1) volumetric point cloud generation, 2) FE-based training datasets preparation, and 3) ProRegNet training and testing. Detailed workflow of the proposed method is shown in Fig. 1. In the first step, volumetric point cloud was generated from automatically segmented MR and TRUS prostate masks using tetrahedron meshing. In the second step, FE models were created to simulate the prostate deformation. Ground truth (GT) volumetric point cloud correspondence was established using the deformation field calculated by the FE models. In the last step, the proposed ProRegNet was trained to perform point cloud matching by directly taking the two prostate volumetric point clouds as input.

3.1. Data acquisition

A total of 50 pairs of T2-weighted MR and TRUS images were collected from 50 prostate cancer patients who underwent HDR brachytherapy at our clinic. The 3D TRUS images were captured with a clinical ultrasound scanner (HI VISION Avius, Hitachi Medical Group, Japan) and a transrectal 7.5 MHz prostate biplane probe (UST-672-5/7.5) in the operating room before catheter placement. To scan over the longitudinal axis, a mechanical SurePoint stepper (Bard Medical, Inc., GA) was used to guide the stepwise movement of the US probe along the longitudinal axis. Multiple parallel scans in the transverse plane were captured from the patient, with a 1mm step size from the apex to the base. The patient was scanned in lithotomy position. The size of a typical prostate is ~60-80 mm in the longitudinal axis. A margin of 5-10 mm anterior and posterior to the prostate was scanned to ensure whole prostate coverage. The T2-weighted MR images were acquired using a 1.5T scanner (MAGNETOM Avanto, Siemens, Germany) with repetition time of 1200ms and echo time of 123ms. MR images were acquired using a body coil. The original voxel sizes of the MR and TRUS images were $0.63 \times 0.63 \times 2.00$ mm^3 and $0.12 \times 0.12 \times 1.00$ mm^3 respectively. For consistency, all MR and TRUS images were resampled to an isotropic voxel size of $0.5 \times 0.5 \times 0.5$ mm^3 prior to image registration. Institutional review board approval was obtained with no informed consent required for this HIPAA-compliant retrospective analysis.

3.2. Volumetric point cloud generation

Prostate can be represented by volumetric point cloud. It is crucial to have accurate prostate segmentation for shape-based MR-TRUS image registration. Ghavami et al. reported that all top five prostate segmentation methods in PROMISE12 challenge were based on CNN (Ghavami et al., 2019; Litjens et al., 2014). The top DSC on whole MRI prostate segmentation was 0.90. Yuan et al. proposed a global optimization approach to delineate 3D endfiring TRUS prostate and achieved a DSC of 0.93 (Yuan et al., 2013) Yang et al. improved TRUS prostate segmentation in brachytherapy with TRUS-CT registration and achieved an average DSC of 0.91 (Yang et al., 2017). In this study, we developed two separate CNN networks for MR and TRUS prostate segmentations. The DSC of our automatic segmentation results were 0.88 ± 0.05 and 0.92 ± 0.03 for MR and TRUS images, respectively. The mean surface distance (MSD) of our automatic segmentation results were 1.02 ± 0.35mm and 0.6 ± 0.23mm for MR and TRUS images, respectively. Our prostate segmentation methods have been published. Readers are referred to Lei et al. (2019) and Wang et al. (2019) for details regarding our segmentation methods.

The segmented prostate masks were then meshed into first order tetrahedron elements for volumetric prostate point cloud generation. First order tetrahedron elements were used to be consis-

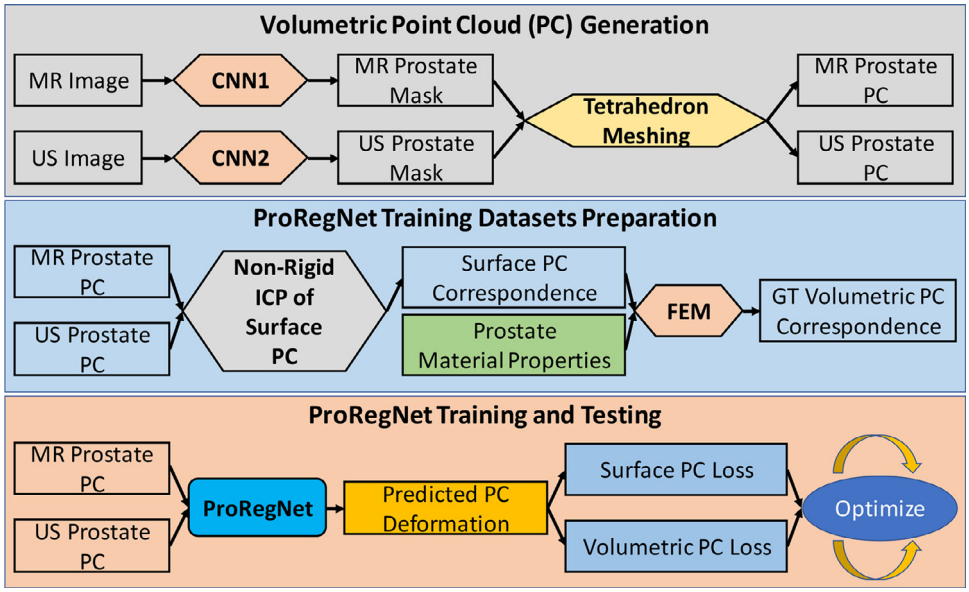


Fig. 1. Flowchart of the proposed method. The proposed framework consists of three major steps which are volumetric point cloud generation, training datasets preparation, network training and testing.

tent with our following FE analysis. Second order tetrahedron elements could be used if the element type in the meshing is consistent with the element type used in the FE. The volumetric MR and TRUS prostate point cloud were subsequently used as the input for the point cloud matching and image registration. Surface triangular meshes were obtained from the tetrahedron meshes after mesh analysis to facilitate calculation of surface-based loss function, which is described in detail later. For robust mesh generation, it is not required for the MR and TRUS mesh to have the same number of vertices, elements, faces or patterns.

3.3. Training datasets generation

For many surface-based methods, the shape of the prostate was represented using 3D surface point cloud. The final voxel-wise dense DVF was generated from surface point cloud deformation using interpolation-based models such as spline interpolations. The interpolation-based models may generate inaccurate intraprostatic deformation due to the lack to internal control points and biomechanical constraints. To address this issue, we aim to train a network that can predict dense volumetric point cloud motion with biomechanical constraints for MR-TRUS registration. To this end, we need to generate dense volumetric point cloud pairs with FE-calculated point correspondences between the dense volumetric point cloud pair for the network training. The patient-specific FE models proposed by Hu et al. (2012) and Wang et al. (2016) cannot be easily applied to our case since the prostate shapes segmented from MR and TRUS images are usually different from each other. Instead of using US probe-prostate interactions as boundary conditions, using nodal displacement between MR and TRUS prostate shapes as boundary conditions ensures that the FE-deformed MR prostate shape match the TRUS prostate shape. Population-based training datasets could be prepared by performing such FE analysis for every case in our datasets. To build FE models with such boundary condition, we need to first establish surface point correspondence. In the next section, surface point registration process is described in detail to establish point correspondence.

3.3.1. Prostate surface registration

Unless otherwise stated, MR prostate point cloud and source point cloud are used interchangeably while TRUS prostate point cloud and target point cloud are used interchangeably. A modified iterative closest point (ICP) based non-rigid surface registration algorithm was used to establish surface point cloud correspondence. The input of the non-rigid ICP algorithm is the prostate surface mesh which include 3D surface point cloud and triangular element connectivity. As an initialization step, the MR and TRUS surface meshes were aligned by overlapping their geometrical centers. Rigid alignment was then performed by optimizing a global bidirectional point cloud correspondence loss (Besl and McKay, 1992). The bidirectional point cloud correspondence includes forward vertex correspondence (source vertex to target vertex) and backward vertex correspondence (target vertex to source vertex).

For computational efficiency, a sparse set of source vertices with even spatial distribution were sampled from the original dense source vertices so that target vertices have a higher spatial density than that of the sampled source vertices. As a result, for the ith vertex on the sampled source surface, there are one nearest neighbor vertex on the target surface and multiple vertices on target surfaces for which their nearest neighbor vertex is the same ith vertex on the sampled source surface. Forward vertex correspondence is defined as the nearest neighbor of one source vertex on the target surface:

$$\overrightarrow{v^t}(i) = \min_{j\in[1,\ N_t]} \left(d\left(v_j^t\ ,\ v_i^s\right)\right) \tag{1}$$

where v_i^srepresents the ith vertex on source surface, v_j^t represents the jth vertex on the target surface, the source and target surfaces have a total number of N_s and N_t vertices. $d(v_j^t\ ,\ v_i^s)$ denote the Euclidean distance between the two spatial vertices v_j^t and v_i^s.

Backward vertex correspondence is defined as the vertices on the target surface for which the same ith vertex on the source surface is their nearest neighbor. Backward vertex correspondence can be expressed as one or more vertices:

$$\overleftarrow{v^t}(i) = \left\{ v_j^t\ v_i^s\ = \min_{\substack{j\in[1,\ N_t]\\ i\in[1,\ N_s]}} \left(d\left(v_j^t\ ,\ v_i^s\right)\right) \right\} \tag{2}$$

Rotation, translation and scaling are solved iteratively by minimizing a weighted forward closest point distance and backward

closest point distance (Audenaert et al., 2019). Scaling was not used in our study as the voxel sizes of the resampled MR and US images were consistent prior to meshing.

$$T(R,c,s) = \min\left(\sum_{i}^{N_s} w_i d^2\left(T\left(v_i^s\right), \overrightarrow{v^t}(i)\right) + \sum_{j}^{N_j} w_j d^2\left(T\left(v_j^s\right), \overleftarrow{v^t}(j)\right)\right) \quad (3)$$

where the first term minimizes the forward closest point distance and the second term minimizes the backward closest point distance. The inclusion of the second term is to introduce extra force to warp the source surface to the target surface with high curvatures since the second term puts more weight on the target surface. Final rigid alignment was achieved using:

$$v' = sRv^s + c \quad (4)$$

To account for the complex prostate deformation, additional non-rigid transformation was needed. The distances between the source and target point cloud were modeled as a sum of N_c Gaussian Radial Basis Functions (G-RBF) (Kim et al., 2016). For computational efficiency, N_c is usually a much smaller subset of N_s The distances vector could be modeled using G-RBF as:

$$d\left(v_i^s\right) = \sum_{k}^{N_c} w_k \varphi\left(v_i^s - c_k\right),\ i\epsilon[1,\ N_s],\ \varphi(r) = e^{-(\varepsilon r)^2} \quad (5)$$

where w_k is the deformation coefficient to be determined, c_k is located at the source surface point location v_i^s . The width of the Gaussian function is controlled by r. ε is the distance between the c_k and v_i^s . The deformation coefficient w_k could be calculated by minimizing:

$$\sum_i^{N_s} d(v_i^s) - \delta(v_i^s)^2 \quad (6)$$

where $\delta(v_i^s)$ is the distance mapping to be approximated using G-RBF. Given the distance mapping $\delta(v^s)$ between the source and target surface during iteration, the deformation coefficient w_k can be calculated as:

$$w = (\emptyset^T\emptyset + \lambda I)^{-1}\emptyset^T\delta(v^s) \quad (7)$$

A final step to refine the surface point matching was performed by locally aligning the two surface point clouds as a weighted local rigid transformation (Audenaert et al., 2019; Li et al., 2008).

3.3.2. Finite element simulation of prostate deformation

FE simulations were performed using the established surface point correspondences as boundary condition to calculate intraprostatic deformations. Commercial FE software (ANSYS 2019 R2, Oxfordshire, UK) was used to construct FE models. The FE models were configured as a static structural job in ANSYS. Throughout our experiments, the initial material properties of prostate were modeled as Young's modulus $E = 5$kPa and Poisson's ratio $\nu = 0.49$, same as that used in Khallaghi et al.'s work (Khallaghi et al., 2015a). Prostate material behavior was simulated by an isotropic and homogeneous hyper-elastic material, represented by Neo-hookean hyper-elastic model (Shahzad et al., 2015). According to continuum mechanics, we have:

$$G = \frac{E}{2(1+\nu)}\ K = \frac{E}{3(1-2\nu)} \quad (8)$$

where G is the initial shear modulus, which was calculated to be 14.9kPa, K is the bulk modulus, which was calculated to be 83kPa. Compressible Neo-hookean material has a strain energy function of (Pence and Gou, 2015):

$$W = C_{10}\left(\overline{I_1} - 3\right) + \frac{1}{D_1}(J_{el} - 1)^2 \quad (9)$$

where C_{10} is the material constant that control the shear behavior, which was calculated to be $C_{10} = G/2 = 7.45$kPa, D_1 is the material constant that control the bulk compressibility, which was calculated to be $D_1 = 2/K = 2.4\ \times 10^{-5}\text{Pa}^{-1}$. $\overline{I_1}$ denotes the first strain invariant, J_{el} denotes the elastic volume strain.

There are around 1000 nodes and 5000 tetrahedral elements in the FE models. A template ANSYS input file was generated for each patient using ANSYS Workbench. Subsequently, these input files were modified to apply the surface point correspondences as nodal displacement boundary conditions. Large deflection checkbox in ANSYS was switched on to account for the material stiffness change due to prostate shape change. The FE models were solved using ANSYS Mechanical APDL with 10 sub-steps with equal time interval. Nodal directional displacements of each sub-step were recorded. One example of the FE meshing and final nodal displacement contour were shown in Fig. 2.

The proposed network took two prostate point clouds generated from MR and TRUS, respectively. The network predicted the MR prostate points motion vector to align the prostate shape of MRI to that of the TRUS. The inputs of the network have two characteristics. First, there is no one-to-one point correspondence between the two point clouds. Second, the numbers of points in the two point clouds are not necessarily the same. Nevertheless, the two point clouds before and after deformation in the FE simulation have such correspondence and the same number of points. Therefore, the network cannot be directly trained using the two point clouds and nodal displacements in the FE simulation. To generate the MR prostate points motion vectors for network training, we first generated voxel-wise DVF images from the FE volumetric nodal displacements using thin-plate-spline interpolation. The voxel-wise DVF images were considered as the ground truth deformation between the MR and TRUS prostate. Then, MR prostate points motion vectors were sampled from the voxel-wise DVF images. For 50 patients, we generated a total of 500 FE results which include ten FE sub-steps per patient. Each training dataset contains a MR prostate volumetric point-cloud set, a TRUS prostate volumetric point-cloud set, and the ground truth motion vector of MR point-cloud set. Each point cloud set contains around 25,000 points.

3.4. ProRegNet

3.4.1. Network design

Fig. 3 shows the goal of ProRegNet, which is to directly match the MR prostate shape to the TRUS prostate shape. Point cloud data structure is represented by a group of spatially distributed 3D points. Compared to image data structure, point cloud data structure is highly irregular. Point cloud data structure has several important properties, which are 1) points in the point cloud are unordered, meaning that the geometrical representation and interpretation should not change regardless of the order of the points in the point cloud. The network should be immune to any permutation of the points in the point cloud; 2) despite being unordered, point cloud is spatially organized. The interaction among points locally or globally forms its own signature that uniquely features the underlying geometric representation; 3) the meaning of the point cloud should be invariant under rigid transformations, such as rotation and translation. A rotated prostate represented by a rotated point cloud with vertices at different spatial locations should have the same meaning as the original prostate. A network that directly consumes point cloud needs to consider all the above-mentioned properties. The vertices number discrepancy between the source point cloud and target point cloud imposes additional challenge to the network design.

The source point cloud has N^svertices, expressed as $v_i^s,\ i \in [1,\ N^s]$. The target point cloud has N^tvertices, expressed as $v_i^t,\ i \in$

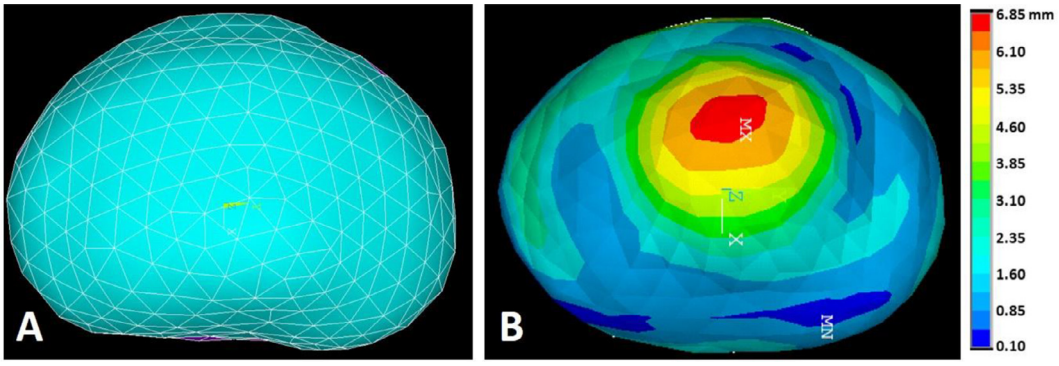


Fig. 2. A: Prostate meshing in FE model, representing MR prostate, B: Contour of prostate displacement magnitude after deformation, representing US prostate.

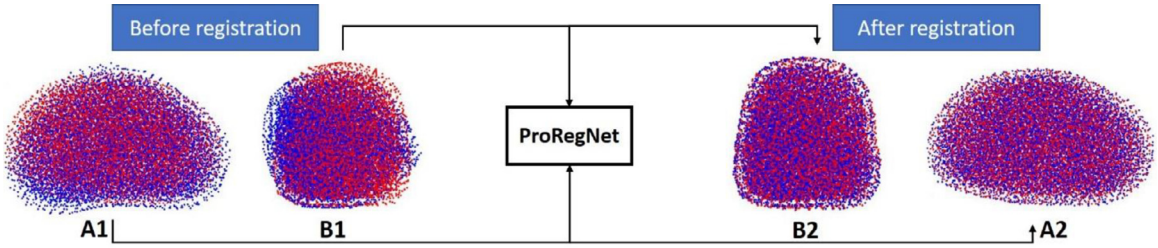


Fig. 3. ProRegNet aims to predict volumetric point cloud motion from a point cloud pair. Red point cloud represents the TRUS prostate shape while blue point cloud represents the MR prostate shape. A1 and A2 show the point clouds in axial plane while B1 and B2 show the point clouds in sagittal plane.

$[1,\ N^t]$. As an initialization step, rigid ICP of surface points were performed to roughly align the point clouds. Then, the centroids of the source and target point clouds were calculated as the geometrical centers and expressed as v_c^s and v_c^t, respectively. The input vectors of the network were then rearranged as a matrix of size $N^s \times 7$ and $N^t \times 7$ for the source and target point clouds, respectively. The i^{th} point in the source point cloud was rearranged as $v_i^s = (x_i^{s\_s}, y_i^{s\_s}, z_i^{s\_s}, x_i^{s\_t}, y_i^{s\_t}, z_i^{s\_t}, m_i^s)$, where $x_i^{s\_s}, y_i^{s\_s}, z_i^{s\_s}$ are the i^{th} point location in the source point cloud relative to v_c^s, $x_i^{s\_t}, y_i^{s\_t}, z_i^{s\_t}$ are the i^{th} point location in the source point cloud relative to v_c^t, and m_i^s is a binary indicator of whether the point belongs to the prostate surface. The m_i^s was introduced to incorporate surface point loss to the network. Similarly, the i^{th} point in the target source point cloud is expressed as $v_i^t = (x_i^{t\_s}, y_i^{t\_s}, z_i^{t\_s}, x_i^{t\_t}, y_i^{t\_t}, z_i^{t\_t}, m_i^t)$, where $x_i^{t\_s}, y_i^{t\_s}, z_i^{t\_s}$ are the i^{th} point location in the target point cloud relative to v_c^s, $x_i^{t\_t}, y_i^{t\_t}, z_i^{t\_t}$ are the i^{th} point location in the target point cloud relative to v_c^t, and m_i^t is a binary indicator of whether the point belongs to the target prostate surface. The reason to spatially encode each point with respect to both v_c^s and v_c^t is to provide two frames of reference for network point cloud motion prediction.

In ProRegNet, N points were randomly sampled from the original point cloud pair. Two types of convolutional layers specific to point cloud were introduced, one for convolution and the other for de-convolution. The point cloud convolutional layer first samples n regions from the input point cloud with farthest point sampling. The centers of the n regions were at v_k'. For each region with radius of r around v_k', features were extracted using function:

$$f_k = \underset{\left\{id\left(v_i^s, v_k'\right) < r\right\}}{\mathrm{MAX}} \left\{h\left(v_i^s\right)\ i\right\} \tag{10}$$

where function $h(\ \cdot\ )$ was realized as a multi-layer perceptron (MLP) network. The MAX operation is element-wise max pooling. Max pooling was used here as a symmetric operation which is suitable for the unordered point cloud and invariant to point cloud rigid transformations. With varying value of radius r, the MLP network could learn multi-scale local point cloud features. For the point cloud de-convolutional layer, features of n input points were computed on specified locations on the target points with n' points, where $n' > n$. The point cloud convolutional and de-convolutional layers solve the problem of end-to-end feature learning of a single point cloud. However, they did not address the problem of point cloud motion prediction between two point-clouds. Liu et al. (2018) proposed a point cloud embedding layer to address this issue. Due to the lack of one-to-one point correspondence, the point cloud embedding layer used multiple softly corresponding points in the target point cloud near a source point within radius of r_e. The point cloud embedding layer learnt to use both feature similarities and spatial relationships to predict a weighted point correspondence.

Fig. 4 shows the architecture of the ProRegNet. N was set to 4096 in our network. Each input point cloud first went through a separate PC conv operator, which includes two-point cloud convolutional layers with $r = 5mm,\ n = 1024,\ c = 64$ and $r = 10mm, n = 256,\ c = 128$, respectively. Then, the learnt features from both point clouds were fed into a point cloud embedding layer with $r_e = 2.5mm,\ c = 256$. The last point cloud conv operator includes two point-cloud convolutional layers with $r = 17.5mm,\ n = 64,\ c = 256$ and $r = 25mm,\ n = 16,\ c = 512$, respectively. The point cloud up-conv operator includes three point-cloud de-convolutional layers to resample the number back to its original size of N.

3.4.2. Loss functions

The loss of the ProRegNet was a weighted sum of surface point cloud loss and volumetric point cloud loss. Mean square error was used as the loss function between the predicted motion and ground truth motion. The volumetric point cloud loss put equal weight on every point in the point cloud. Surface point cloud loss was added to promote the point correspondence between prostate surface points. Therefore, m_i^s was encoded into v_i^s to distinguish the surface points from the intraprostatic points. Challenges to design a suitable surface loss function include 1) the function needs to be differentiable with respect to point locations; 2) the function needs to be computationally efficient since the loss needs to be

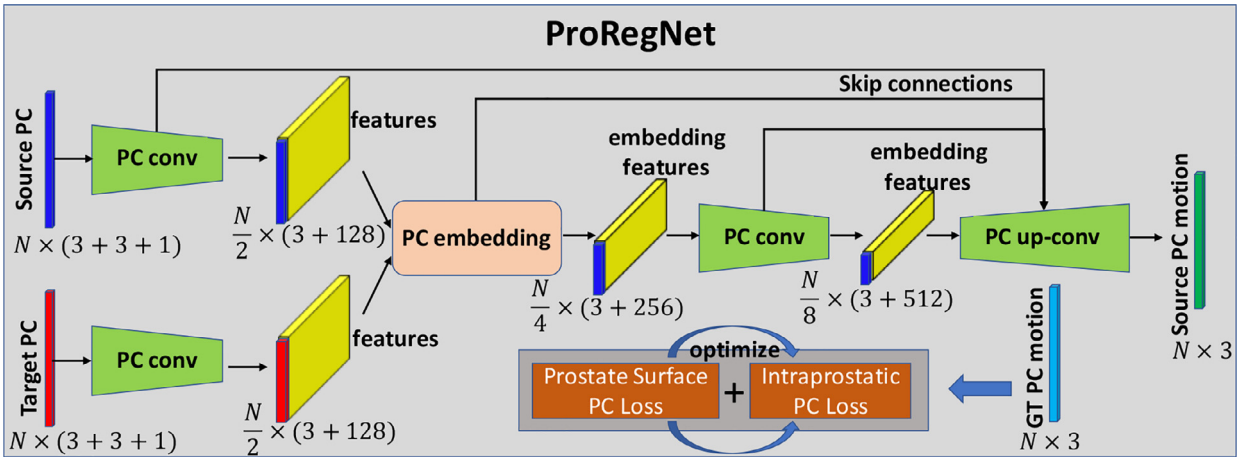


Fig. 4. Network design of ProRegNet. The network consists of PC conv, PC embedding and PC up-conv operations. The loss function is a weighted sum of prostate surface PC loss and intraprostatic PC loss.

calculated repeatedly throughout the optimization process; 3) the function needs to be robust to outliers. Fan et al. (2016) studied two distance metrics, which are the Chamfer distance and Earth Mover's distance. We adopted the Chamfer distance (CD), which is defined as:

$$d_{CD}\left(v^{s'}, v^{t'}\right) = \sum_{i \in N^{s'}} \underset{j \in N^{t'}}{\text{MIN}} \left(d^2\left(v_i^{s'}, v_j^{t'}\right)\right) + \sum_{j \in N^{t'}} \underset{i \in N^{s'}}{\text{MIN}} \left(d^2\left(v_j^{t'}, v_i^{s'}\right)\right) \quad (11)$$

where $v_i^{s'}$ *and* $v_j^{t'}$ represent the surface points from the source point cloud and target point cloud. $N^{s'}$ and $N^{t'}$ are the number of surface points for source and target, respectively. For each point, the Chamfer distance algorithm finds the nearest neighbor in the other point cloud and adds the squared distance of each points together. The Chamfer distance is continuous and piecewise smooth with respect to point locations. The final loss for ProRegNet is:

$$L = \text{AVG}\left\{(V_p - V_{GT})^2\right\} + 10 \cdot d_{CD}\left(v^{s'}, v^{t'}\right) \quad (12)$$

where the first term is the mean square error between the predicted motion vector and the ground truth motion vector. The second term is the surface distance loss. The weighting factor was set to 10 because the ratio between the number of volumetric and surface points was around 10.

3.4.3. Training and testing

ProRegNet was implemented in Python 3.6 and TensorFlow on a NVIDIA TITAN XP GPU with 12GB of memory. Adam gradient optimizer with learning rate of 1e-3 was used in optimization. Ten FE results per patient were used to generate nine training datasets by pairing up the first step with each subsequent sub-step. In total, we have 450 training datasets for 50 patients. Five-fold cross validation was used to train and test the network. During training and testing, 4096 points were first randomly sampled from the original point cloud. For 4096 volumetric points, there were around 400 surface points. Since the network was trained to be invariant to point permutations, the sampled 4096 points were randomly shuffled ten times in testing stage. Final motion results were obtained by averaging the ten motion predictions. Voxel-wise DVF image was obtained from the network motion prediction using thin-plate-spline interpolation. MR image was then deformed using the voxel-wise DVF to register with the TRUS image.

The training datasets preparation is time-consuming. However, once trained, the network can make the point matching prediction very rapidly in less than 3 seconds. The ProRegNet took around 2 hours to train. The prostate can be segmented on MR and US images using our previously trained segmentation networks in less than 2 seconds (Lei et al., 2019; Wang et al., 2019). The whole MR-TRUS image registration procedures, including segmentation, meshing, network prediction and DVF image calculation took 3 to 5 mins on average.

3.5. Comparison methods

To demonstrate the superiority of the proposed method, it was compared with another two MR-TRUS image registration methods. The first method was a surface-based method, which performed prostate surface registration same as Section 3.3.1. Voxel-wise DVF image was then obtained from the surface point cloud motion using thin-plate-spline interpolation. The other method was to use CNN for MR-TRUS image registration. From the 450 point-cloud training datasets, we created another 450 image training datasets. Each image training dataset includes a MR prostate mask, a TRUS prostate mask and a ground truth DVF image obtained using FE models. Fig. 5. shows the network structure of the CNN network. The CNN has 12 convolutional layers and 3 max pooling layers. Bicubic resampler was used to resample the predicted DVF to its original image size. Deformed MR masks were generated using spatial transformers. DICE coefficient loss, DVF mean absolute error (MAE) loss and DVF regularization smoothness loss were minimized for network training. To provide spatial information to the input mask images, three spatially-encoded masks in the x, y, z directions were generated from one input mask. There was a total of 6 spatially-encoded masks, three masks for MR and another three masks for TRUS. For fair comparison, the CNN was trained in the same way as ProRegNet. For the ease of description, the surface-based method was referred to 'SurfReg'. The CNN-based method is referred to 'CnnReg'.

3.6. Evaluation of registration accuracy

Target registration error (TRE) were calculated to evaluate the registration accuracy. Landmarks such as centers of small nodules, cysts, calcifications were carefully selected manually by an experienced physician. For each patient, around 3 to 5 landmarks were selected. DSC coefficient was calculated to evaluate the volume overlap ratio between the deformed MR and TRUS prostate. To evaluate the surface similarity between the deformed MR and TRUS prostate, we used Hausdorff distance (HD) and mean surface distance (MSD). The HD measures the local maximum distance between the deformed MR prostate shape and the TRUS prostate shape. The MSD is the mean surface distance between the two surfaces. Jacobian determinants (JD) and normal strains E_{xx}, E_{yy}, E_{zz}

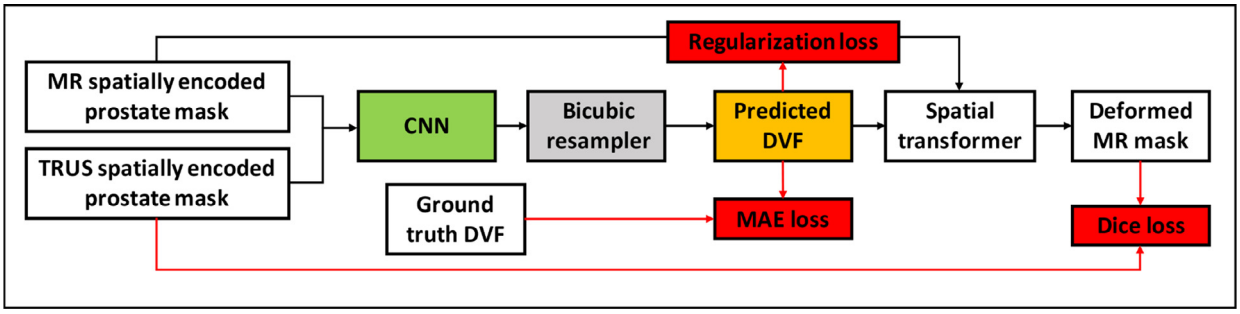


Fig. 5. Network design of CnnReg. The CNN network consists of 12 convolutional layers and 3 max pooling layers. The loss function contains DVF regularization loss, DVF mean absolute error loss and Dice loss.

Table 1
DSC, MSD, HD and TRE for SurfReg, CnnReg and ProRegNet. Values: Mean±Std.

	DSC	MSD *(mm)*	HD *(mm)*	TRE *(mm)*
SurfReg (1)	0.98±0.01	0.58±0.04	1.86±1.24	2.70±1.15
CnnReg (2)	0.96±0.03	0.73±0.19	2.51±1.21	2.48±1.05
ProRegNet (3)	0.94±0.02	0.90±0.23	2.96±1.00	1.57±0.77
p-value (1 vs. 3)	0.11	0.02	0.09	0.005
p-value (2 vs. 3)	<0.001	<0.001	<0.001	0.002

of the deformation field were calculated to assess the physical fidelity of the predicted deformation. The Jacobian determinant is defined as:

$$D = \det(\varepsilon + I) \tag{13}$$

where ε is the strain tensor calculated from the deformation field, I is the identity matrix, $\det(\cdot)$ is the determinant of the matrix.

4. Results

4.1. Registration accuracy

To evaluate the registration accuracy, we calculated eight metrics, including DSC, MSD, HD, TRE, JD, E_{xx}, E_{yy} and E_{zz}, for SurfReg, CnnReg and ProRegNet, respectively. The results were compared and analyzed. Table 1 shows the mean and standard deviation of DSC, MSD, HD and TRE for the three different methods. Boxplot of these metrics were shown in Fig. 6 and Fig. 7. The central mark in the boxplot is the median, the top and bottom edges of the box indicate the 75th and 25th percentiles, respectively. The dashed line in boxplot extends to the most extreme data points that are not outliers. Outliers are shown as red '+' symbol. As evidenced by its DSC, MSD and HD values, SurfReg has the best performance on surface matching and volume matching among the three methods. This is reasonable since SurfReg focused solely on the surface point matching. As described in Section 3.3.1, SurfReg utilized G-RBF to calculate general deformation for rough point matching and a weighted rigid deformation for local fine point matching. Therefore, SurfReg enforced the two prostate shapes to match well both globally and locally. CnnReg generated the second-best surface matching and volume matching performance in terms of DSC, MSD and HD. However, volume overlapping (DSC) and surface distance (MSD, HD) metrics are not accurate or adequate to evaluate the point-to-point MR-TRUS registration accuracies. The reasons are 1) intraprostatic deformation is not considered in DSC, MSD or HD; 2) exact shape matching could sometimes generate adverse effect due to inaccurate local segmentations as a result of either poor contrast, low slice resolution or tissue ambiguities. To evaluate the intraprostatic registration accuracy, TRE were calculated and reported in Table 1. Paired student t-test was performed to analyze the statistical significance for all four metrics. Contrary to DSC, MSD and HD values, the TRE values suggested that ProRegNet was the most accurate among the three methods, with an average improvement of 0.91mm over CnnReg and 1.13 mm over SurfReg. Multiple comparisons were made with the same null hypothesis using DSC, MSD, HD. Bonferroni correction suggests a threshold significance level of 0.05/3=0.017 to account for the familywise type I error. The p-values (1 vs. 3) in Table 1 are greater than 0.017, which suggests that SurfReg has the same performance as ProRegNet in terms of DSC, MSD and HD.

Registration results of two cases are shown in Fig. 8 and Fig. 9 in the transverse, coronal and sagittal planes. In Fig. 8 and Fig. 9, first column shows the original TRUS and MR images. Second column shows the images after rigid registration in checkboard and red-green fusion. Third, fourth and fifth columns show the registration results for SurfReg, CnnReg and ProRegNet, respectively. Blue

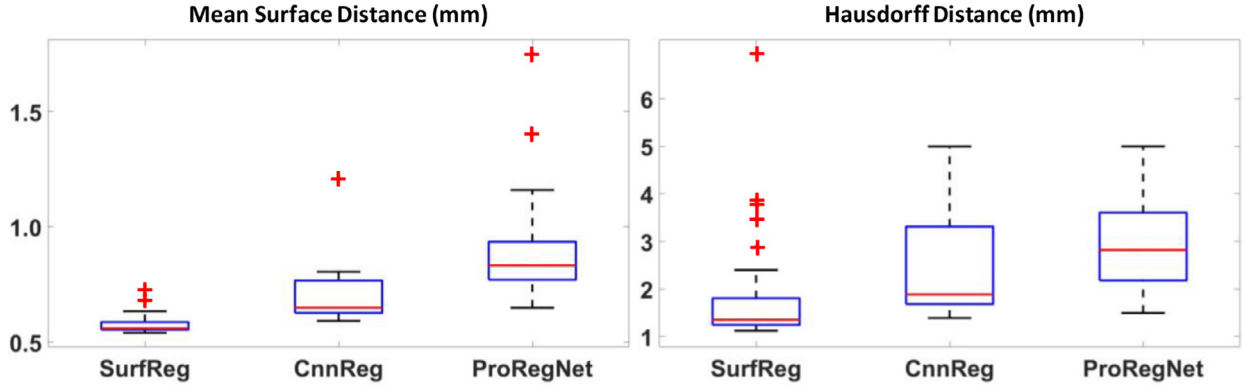


Fig. 6. Boxplots of MSD and HD for SurfReg, CnnReg and ProRegNet. The central mark in the boxplot is the median, the top and bottom edges of the box indicate the 75th and 25th percentiles, respectively. The dashed line in boxplot extends to the most extreme data points that are not outliers. Outliers are shown as red '+' symbol.

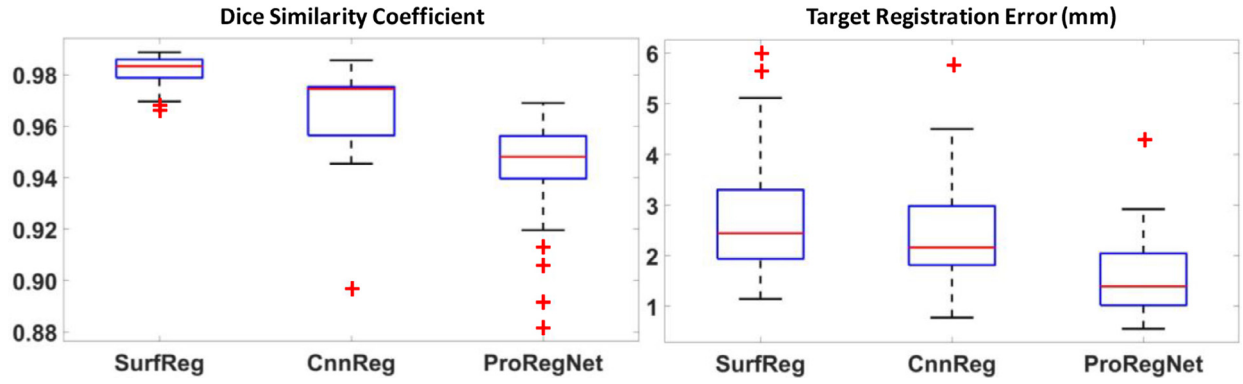


Fig. 7. Boxplots of DSC and TRE for SurfReg, CnnReg and ProRegNet. The central mark in the boxplot is the median, the top and bottom edges of the box indicate the 75th and 25th percentiles, respectively. The dashed line in boxplot extends to the most extreme data points that are not outliers. Outliers are shown as red '+' symbol.

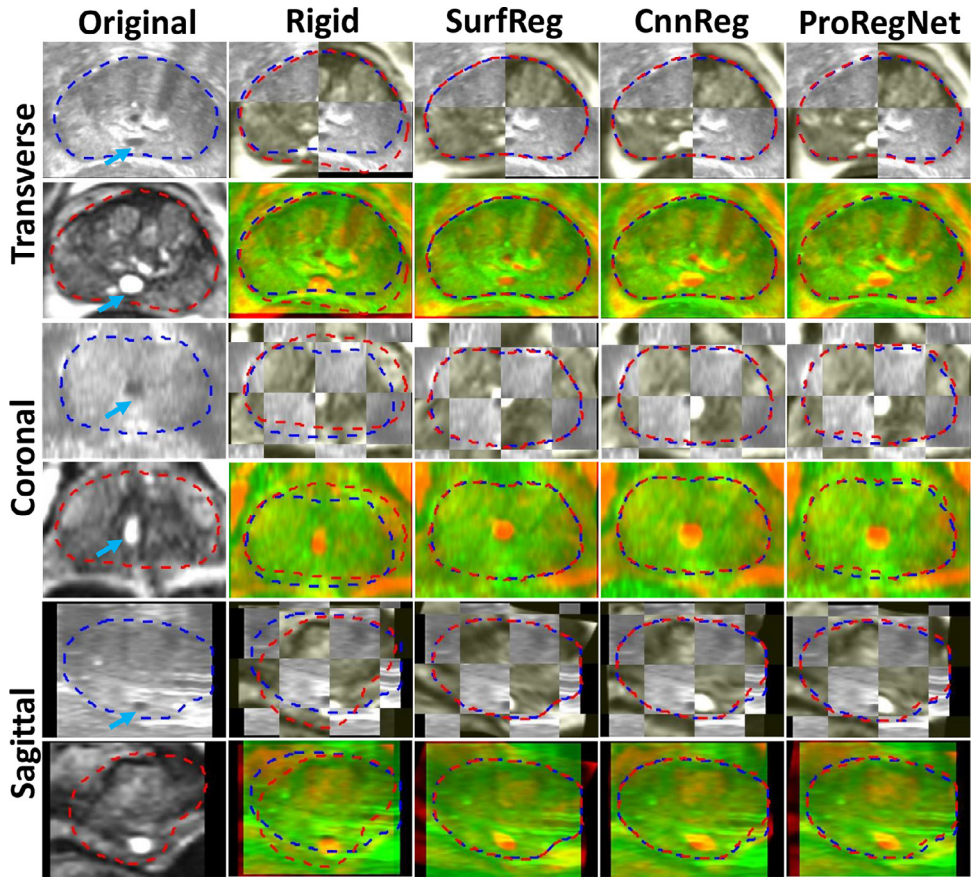


Fig. 8. Example image slices from one case. Results are shown in the transverse, coronal and sagittal planes. First column shows the original TRUS and MR images. Second column shows the images after rigid registration in checkboard and red-green fusion. Third, fourth and fifth columns show the registration results for SurfReg, CnnReg and ProRegNet, respectively. Blue dashed contour indicates the TRUS prostate shape while red dashed contour indicates MR prostate shape. The cyst within prostate were highlighted by arrows on the original images.

dashed contour indicates TRUS prostate shape while red dashed contour indicates MR prostate shape. The cyst within prostate was highlighted by arrows. Arrows were not plotted in fused images to avoid occlusion. The checkboards were placed so that the checkboard edges crossed the cyst of interest for better visualization. Fig. 8 and Fig. 9 shows that the prostate shape matching after registration was the best for SurfReg than the other two methods. However, as illustrated by the arrows in Fig. 8 and Fig. 9, ProRegNet has the best cyst matching among the three methods. This phenomenon could be explained by the fact that SurfReg lacks biomechanical constraint for intraprostatic deformation regularization. For SurfReg, TPS interpolation was used to calculate intraprostatic deformation, leading to under-constrained deformation. Despite trained using ground truth DVFs, the CnnReg only achieved limited improvement over SurfReg in terms of TRE. The results show that CnnReg could not efficiently learn how to perform registration from the spatially-encoded prostate mask pair. This could be explained by the fact that MR-TRUS masks lack enough intensity-based features for the registration task. As a result, CnnReg which is good at extracting intensity-based textural features from images failed to significantly outperform SurfReg. On the contrary, ProRegNet has managed to learn the underlying point matching pattern directly from the point cloud pair and its correspondences. Despite of some contour discrepancies between the

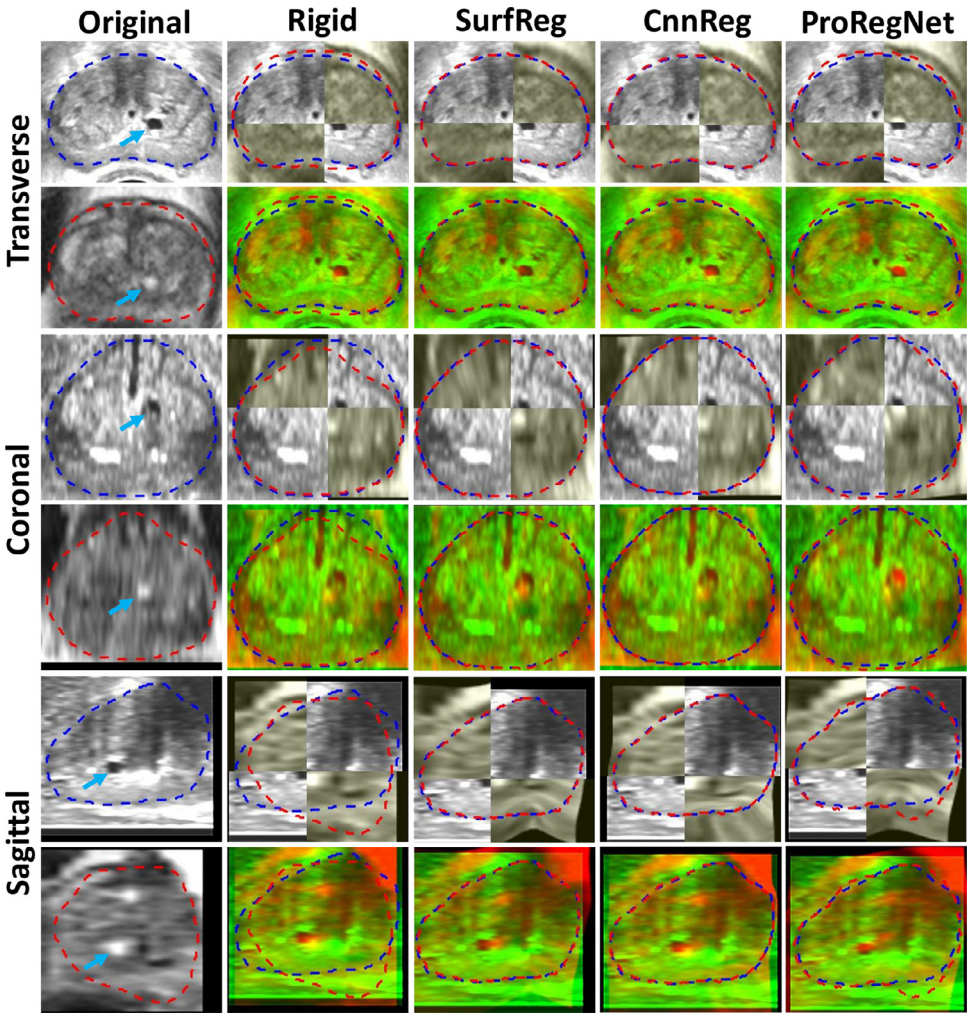


Fig. 9. Example image slices from another case. Results are shown in the transverse, coronal and sagittal planes. First column shows the original TRUS and MR images. Second column shows the images after rigid registration in checkboard and red-green fusion. Third, fourth and fifth columns show the registration results for SurfReg, CnnReg and ProRegNet, respectively. Blue dashed contour indicates the TRUS prostate shape while red dashed contour indicates MR prostate shape. The cyst within prostate were highlighted by arrows on the original images.

deformed MR shape and TRUS shape, the ProRegNet generated the best cyst matching among the three methods. Therefore, we claim that it is sometimes necessary to tolerate certain level of prostate surface mismatch and preserve the underlying biomechanical constraints, given possible inaccurate local contours. Similar concept was found in Khallaghi et al.'s work (Khallaghi et al., 2015a), where the authors proposed to use statistical shape models to account for prostate contour inaccuracies for consistent surface registration.

Strain tensor images were obtained from the predicted DVF. JD, E_{xx}, E_{yy} and E_{zz} images were calculated from the strain tensor images to evaluate the physical fidelity of the predicted DVF. JD of the predicted DVF calculated by Eq. 13 represents the volumetric strain. JD value greater than one means volumetric expansion while JD value less than one means volumetric shrinkage. Negative JD value means deformation singularity, indicating 'folding' in the DVF. Fig. 10 shows that the ranges of JD are all positive, which means the predicted DVF is free from singularities for all three methods. Since SurfReg enforced strong surface matching, SurfReg was more likely to generate inaccurate JD near the surface, as evidenced by the arrows in Fig. 10. Though trained using the ground truth DVF, CnnReg generated JD images that were not smooth enough near the prostate surface. In comparison, ProRegNet generated JD images that were in best agreement to the ground truth. To evaluate the JD error quantitatively, we calculated the mean absolute error between the ground truth and the predicted JD within the prostate mask for all three methods. Results were reported in Table 2. The proposed ProRegNet had the smallest mean and standard deviation error for JD, E_{xx}, E_{yy} and E_{zz}. P-values were calculated using two-sample t-test for (ProRegNet, SurfReg) and (ProRegNet, CnnReg). P-values show that the improvement of ProRegNet over SurfReg was statistically significant for JD, E_{yy} and E_{zz}. However, the improvement of ProRegNet over CnnReg was statistically significant for only E_{zz}. This could be explained by the fact that both the ProRegNet and CnnReg were trained using the ground truth DVF, resulting in statistically insignificant differences for JD, E_{xx} and E_{yy}. Nevertheless, Fig. 10 suggests that the Cn-

Table 2
Mean absolute error between the ground truth and prediction for Jacobian determinant, normal strain in the x, y, z directions. Values: Mean±Std.

	JD	E_{xx}	E_{yy}	E_{zz}
SurfReg (1)	0.18±0.04	0.10±0.02	0.12±0.03	0.10±0.02
CnnReg (2)	0.17±0.05	0.10±0.01	0.10±0.02	0.12±0.04
ProRegNet (3)	0.14±0.02	0.09±0.01	0.09±0.01	0.09±0.01
p-value (1 vs 3)	**0.01**	0.05	**0.01**	**0.01**
p-value (2 vs 3)	0.09	0.14	0.19	**0.02**

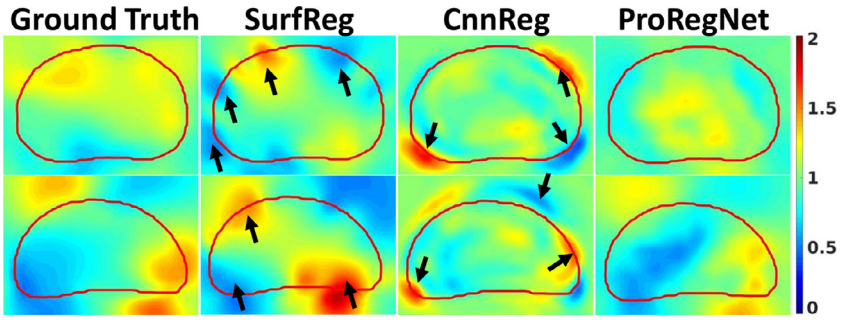


Fig. 10. Example slices of JD comparison for SurfReg, CnnReg and ProRegNet from two different cases. Red contour indicates the TRUS prostate shape. Arrows indicate high JD errors.

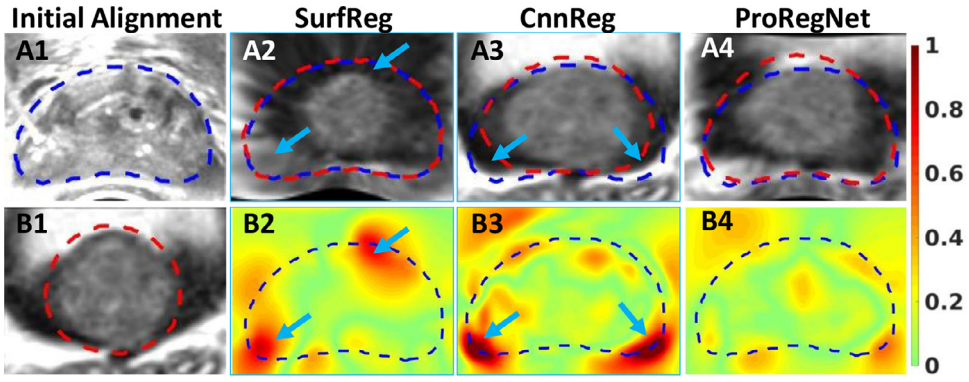


Fig. 11. A1: TRUS image. B1: MR image. A2-A4: Deformed MR images after registration by SurfReg, CnnReg and ProRegNet, respectively. B2-B4: JD absolute error for SurfReg, CnnReg and ProRegNet, respectively. Blue contours indicate the TRUS prostate shapes while the red contours indicate MR prostate shapes. Poor registration was observed at regions where JD error was large, as indicated by arrows.

nReg had significant JD error, especially outside the prostate mask, which was not included in the JD error calculation in Table 2.

Fig. 11 shows one example slice before and after registration for all three methods. Fig. 11 (A2-A4) shows the deformed MR, with the TRUS contours shown in blue and deformed MR contours shown in red. Fig. 11 (B2-B4) shows the absolute JD error with same TRUS contours shown in blue. We have found out that the SurfReg sometimes generated unrealistically deformed MR image after registration, such as Fig. 11(A2). This phenomenon was more likely to happen near the prostate surface for patients whose prostate has undergone significant shape changes during TRUS. Such phenomenon did not occur for any patient using ProRegNet. Indicated by the arrows in Fig. 11 (A3), there may be large contour shape discrepancies after registration for CnnReg. We aligned the JD error image vertically with the deformed MR images to analyze the correlation. As indicated by the arrows, poor registration results occurred at locations where JD errors were large. Fig. 11 (B4) shows that JD error was minimal for ProRegNet. Fig. 11 demonstrated that the biomechanical constraints learnt by ProRegNet was effective in regularizing the deformation prediction.

Besides the numerical evaluations, qualitative assessments of registration uncertainties were also performed by one experienced medical physicist with more than 10 years of experience on MR-TRUS fusion. To assess the uncertain levels of registration results, we have followed the guidelines recommended by AAPM Task Group 132 (Brock et al., 2017), which defines uncertainty level of 0 to 4 as 'whole scan aligned', 'locally aligned', 'usable with risk of deformation', 'usable for diagnosis only', 'alignment not acceptable', respectively. Fig. 12 shows the grouped bar plot of the uncertainty levels for all three methods. The mean uncertainty levels were 2.08, 2.04 and 1.50 for SurfReg, CnnReg and ProRegNet, respectively. The unacceptable cases were caused by large prostate shape discrepancies at regions such as lateral prostate, prostate apex and base, resulting in unrealistically stretched or compressed prostate.

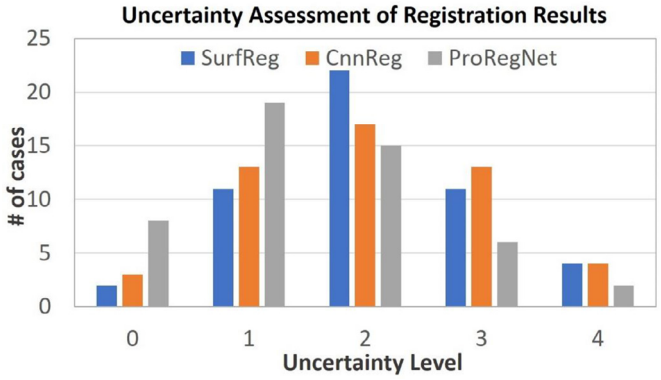


Fig. 12. Qualitative expert uncertainty assessments of MR-TURS image registration results.

4.2. Robustness of ProRegNet to inaccurate segmentation

In practice, the prostate shapes might be corrupted due to inaccurate segmentations. We studied the robustness of ProRegNet to inaccurate segmentations and meshing perturbations by adding random noise to the locations of the points in both MR and TRUS point cloud. The original point clouds were perturbed by adding noise with four levels of amplitude on selected points with five percentages. The four spatial noise amplitude levels were [-2.5, 2.5], [-5.0, 5.0], [-7.5, 7.5] and [-10.0, 10.0] mm with random uniform distribution. The five percentages were 20%, 40%, 60%, 80% and 100%. Experiments of registration were conducted on these

Table 3
Noise sensitivity of ProRegNet in terms of TRE in mm. Values: Mean±Std. Statistical significance is indicated in bold.

Noise Amplitude (mm)	Point Percentage				
	20%	40%	60%	80%	100%
[-2.5,2.5]	1.51±0.62	1.54±0.59	1.53±0.65	1.55±0.59	1.53±0.58
[-5,5]	1.54±0.56	1.53±0.59	1.57±0.56	1.55±0.60	1.57±0.59
[-7.5,7.5]	1.65±0.61	1.69±0.62	1.67±0.57	1.70±0.63	1.73±0.66
[-10,10]	1.74±0.58	1.79±0.65	1.87±0.60	1.94±0.77	1.97±0.79

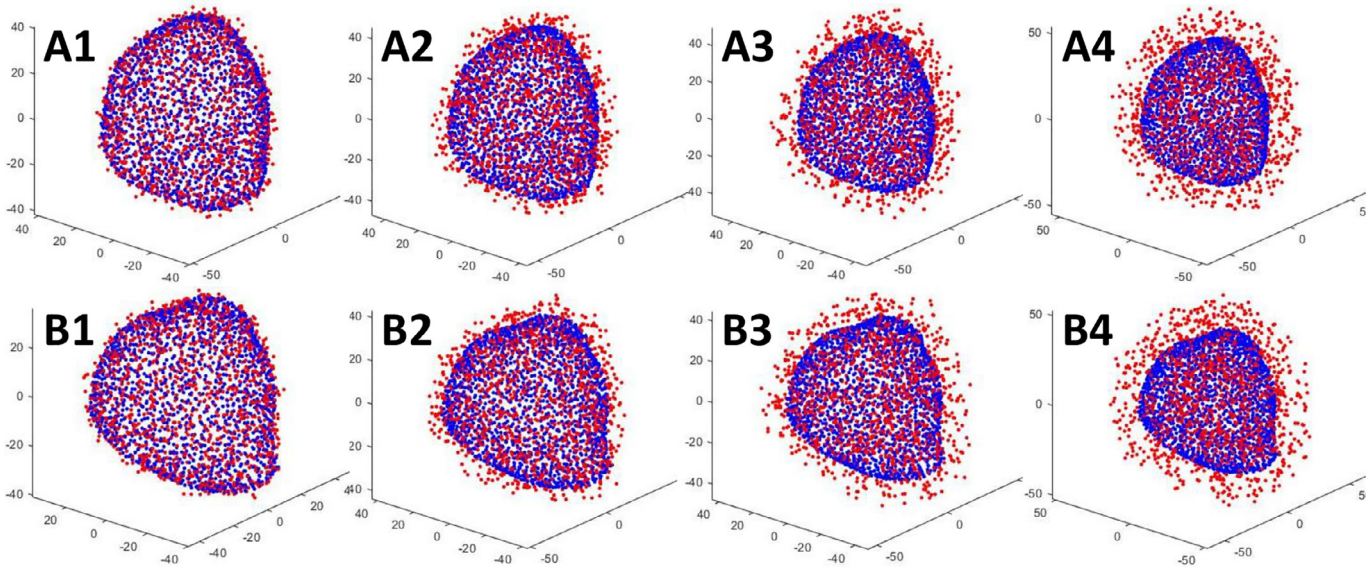


Fig. 13. Point clouds contaminated with different level of noise. Blue point clouds represent original shapes while red point clouds represent noisy shapes. A1-A4: TRUS prostate shapes with noise amplitude of ±2.5, ±5, ±7.5, ±10 mm on 100% of points, respectively. B1-B4: MR prostate shapes with noise amplitude of ±2.5, ±5, ±7.5, ±10 mm on 100% of points, respectively.

Table 4
P-values between TREs calculated using original shapes and TREs calculated using noisy shapes. Statistical significance is indicated in bold.

Noise Amplitude (mm)	Point Percentage				
	20%	40%	60%	80%	100%
[-2.5,2.5]	0.91	0.78	0.84	0.71	0.84
[-5.0,5.0]	0.78	0.81	0.58	0.72	0.60
[-7.5,7.5]	0.29	0.18	0.21	0.18	0.12
[-10.0,10.0]	0.09	0.05	**0.01**	**0.01**	**0.01**

noisy prostate shapes. TRE was calculated using the same landmarks and reported in Table 3. It can be observed from Table 3 that the TRE generally increased from 1.51mm to 1.97mm as the noise level increased from 2.5mm on 20% points to 10mm on 100% points.

To further study the statistical significance of the TRE results between using original shapes and using noisy shapes, we calculated the p-values using two-sample t-test. The p-values were shown in Table 4. For visual comparison, we plotted noisy shapes with its original shapes on both TRUS and MR prostate shapes in Fig. 13. Blue point clouds represent original shapes while red point clouds represent noisy shapes. Fig. 13 (A1-A4) show TRUS prostate shapes with noise amplitude of ±2.5, ±5, ±7.5, ±10 mm on 100% of points, respectively. Fig. 13 (B1-B4) show MR prostate shapes with noise amplitude of ±2.5, ±5, ±7.5, ±10 mm on 100% of points, respectively. Table 4 shows that only shapes of Fig. 13 (A4 and B4) in Fig. 13 caused the ProRegNet to generate TRE values that were statistically significantly different than the original shapes. Both Tables 3 and 4 demonstrated that our ProRegNet was able to perform robustly under severe noisy contamination on the original point clouds. One reason for the good robustness to noise is that the point clouds of prostate shapes in the training datasets had no specific spatial patterns. The spatial pattern of the point clouds could easily be destroyed by the added noise, which would significantly degrade the network performance. The robustness of ProRegNet could also be attributed to the fact that noise was added uniformed to all the points which did not perturb the overall shapes of the prostate much. Nevertheless, this proved that the ProRegNet was able to retrieve important shape representative features and discard inaccurate outlier points.

Since the automatically prostate segmentation may have errors, we performed an additional experiment to investigate the robustness of our method to such segmentation errors. For the ease of evaluation, we have chosen the case in Fig. 9 as the experiment subject as the cyst in this case is clearly visible. Manual contour was carefully prepared for this patient. MR-TRUS registration was performed using the manual segmentation and compared to that using the automatic segmentation results. In Fig. 14, we can see that there are very minimal differences between the registration results. The registration results using manual contours appear to be slightly better as judged by the cyst matching indicated by the arrows in Fig. 14.

5. Discussions

In this study, a novel deep learning-based point cloud matching method was proposed for MR-TRUS image registration. Due to the vast appearance difference between MR and TRUS images, image

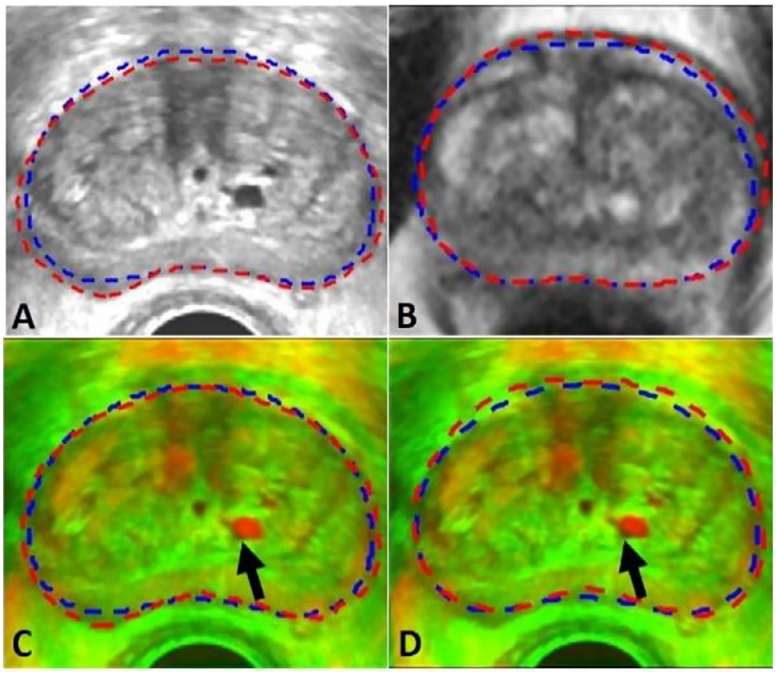


Fig. 14. A-B: Manual contours and automatic contours on TRUS and MR images, respectively. Manual contours are shown in red and automatic contours are shown in blue. C-D: Fused images of registration results using manual contours and automatic contours, respectively.

intensity-based similarity metrics have very limited success in MR-TRUS image registration. To bypass the need for an accurate similarity metric, the segmented prostate shapes were utilized to drive the registration process. The prostate shapes are the reliable features that are available in both MR and TRUS images. However, it is difficult to propagate the prostate surface-driven deformation to intraprostatic deformation. Intraprostatic deformation was usually calculated from surface-driven deformation using interpolation-based models, such as radial basis functions, free-form deformations and Gaussian-basis functions. These interpolation-based models may result in implausible interior deformation field due to the lack of intraprostatic control points and deformation regularization. To address this issue, we have used volumetric point could with many intraprostatic points. The proposed ProRegNet was able to directly consume irregular 3D point cloud and predict the motion between the MR and TRUS volumetric point clouds without iterations. Biomechanical models have been proven to be superior to interpolation-based models (Khallaghi et al., 2015a; Khallaghi et al., 2015b; van de Ven et al., 2015). However, it is impractical to build a patient-specific FE model for every patient who underwent TRUS in the operating room given that FE modeling is very time-consuming and resource-intensive. Recently, Liang et al. (2018) published a paper on a first attempt to use deep learning as a fast and accurate surrogate for the time-consuming FE modeling on aorta stress calculation. Inspired by this work, we proposed to train a deep learning network to replace FE modeling that deforms the MR prostate shape to TRUS prostate shape. A total of 500 training datasets, 10 datasets per patient, were generated by FE modeling of MR-TRUS deformation for 50 patients. The proposed ProRegNet was trained to learn the biomechanical constraints that were encoded in the FE-generated point correspondence.

It is important to ensure the segmentation accuracy of the prostate for surface-driven registrations. Manual contour is too time-consuming to be performed for every patient in real time. In addition, human contour is prone to error under time-constraint and fatigue. To address this issue, we proposed to use deep learning-based methods for prostate segmentation in the first step. The deep learning-based methods could achieve an average DSC of 0.86 and 0.92 for MR and TRUS, respectively. To avoid large segmentation errors, physicians are encouraged to check and make necessary corrections prior to registration. We have found out that the most extreme outlier case in Fig. 6 and Fig. 7 for ProRegNet was from the same patient. To further analyze the cause, we plotted the prostate shapes of this outlier case in Fig. 15 (B1 and B2). For comparison purpose, another case with good registration accuracy was shown in Fig. 15 (A1 and A2). The cause of the outlier could be two-fold, 1) the large shape discrepancy could be caused by segmentation inaccuracies, especially at the apex and base slices along the transverse axis. Though the prostate boundary is usually more visible in MR than TRUS, MR prostate segmentation is prone to error because MR usually has a large slice thickness. The large slice thickness would introduce errors especially at the apex and base of the prostate since the prostate may end between slices. To avoid such an outlier, we could ask physicians to check the contours and make necessary corrections quickly prior to registration; 2) the prostate shape in TRUS has undergone large deformation compared to MR, leading to significant prostate shape difference. The small training datasets could not cover this extreme case, resulting in poor generalization to this extreme case. We could collect more datasets to include such cases into the training datasets in the future.

In this study, we utilized nodal displacement boundary conditions instead of US probe-prostate interaction modeling to drive the prostate deformation. The US probe-prostate interaction modelling may be advantageous at modelling the actual US imaging process. However, it is very difficult to accurately define other FE boundary conditions that require not only precise surrounding organs modelling such as shapes, stiffnesses but also proper interactions handling between them such as nodal coupling, friction or frictionless sliding. In addition, the level of elastic support the prostate gets from its surrounding tissues directly affects how much rigid translation the prostate could undergo when the prostate is pressed by the US probe. More importantly, the FE simulation driven by different US probe-prostate tissue inter-

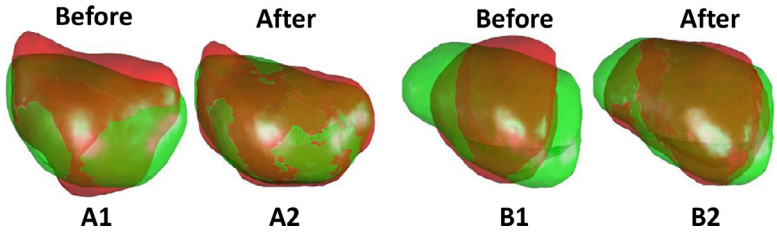


Fig. 15. MR-TRUS prostate shapes before and after ProRegNet registration in the sagittal plane. Green shape represents TURS prostate shape while red shape represents MR prostate shape. A1-A2: Prostate shapes before and after registration for one good case. B1-B2: Prostate shapes before and after registration for the outlier case.

actions could hardly deform the MR prostate in a way that exactly replicate the prostate shape of the US. This is undesired in our study since our goal is to replicate the prostate shape correspondence between MR and US, before and amid the US imaging, so that we can predict the prostate motion from the two shapes. On the contrary, nodal displacement boundary condition could ensure that the FE-deformed MR prostate shape matches that of the US prostate shape. Being aware that the nodal displacement correspondence generated by surface registration could lead to errors in FE simulation, we only generated training datasets based on the surface registration method to bootstrap our biomechanical modelling and network training. Once the network was trained, one-to-one point correspondence was not required to predict the motion vector field of one volumetric point cloud to match the other. It was further shown in Fig. 14 that the proposed method is tolerable to certain levels of prostate segmentation errors.

We used the sum of volumetric points motion error and surface Chamber distance as loss function. The volumetric points motion error which measures the differences between the predicted motion and the FE-generated motion enables the network to learn the biomechanically constrained motion patterns. The surface Chamber distance encourages the network to match prostate shapes at the same time. We have trained another model without the surface distance loss to test its importance. The DSC, MSD, HD and TRE for the network trained without surface distance loss were 0.93±0.02, 0.97±0.17 mm, 3.40±1.06 mm and 1.61±0.55 mm, respectively. All four metrics were worse without surface distance loss, demonstrating the usefulness of surface distance loss.

One limitation of this study is that only one set of material parameters was used in the FE modeling. Hu et al. (2012) simulated the prostate using a range of parameters. Wang et al. (2016) obtained the prostate material parameter using US shear wave elastography. The accuracy of the US elastography needs to be carefully studied before its precision can be trusted. All tissues were assumed to behave as isotropic, homogeneous and elastic material in both studies, which is another source of error in biomechanical modeling. It is difficult to justify the importance of absolute accuracy of patient-specific material elasticity given other uncertainties, such as material anisotropy, heterogeneity and boundary condition precision. On the other hand, it is reasonable for Wang et al. to measure individual material elasticity since they aimed to build a patient-specific model. On the contrary, we aimed to build a population-based model using FE. Fu et al. (2013) reported that soft tissue is highly inhomogeneous with nondeterministic material parameters. Soft tissue of the same subject could vary across different days. Hence, it is reasonable to use the mean prostate elasticity for population-based simulations.

Our method was trained and tested using datasets from brachytherapy which has a similar US probe-tissue interaction, US image acquisition and needle/catheter placement orientation as perineum biopsy. On the other hand, end-fired biopsy widely used today has different processes, which may cause different prostate deformation pattern. Therefore, the proposed method may not be readily applicable to end-fired biopsy. The segmentation networks were trained and tested using images from our department. They are likely to generate sub-optimal segmentation results on images from different commercial systems. In the future, we plan to investigate our networks using multi-institutional datasets from different imaging systems.

We have investigated our method via comparison with the traditional surface-based methods and CNN-based method using eight different metrics. To study the robustness of our method to noise, many experiments were conducted with prostate shapes contaminated with different level of noise. Though we have investigated our method from different perspectives, there are many other experiments that we could conduct to further improve our understanding. Currently, we paired only the first step with the subsequent sub-steps of FE results to prepare the training datasets. In the future, we will try to pair up any two sub-steps of the FE results in the training dataset. In this manner, we could have 2250 training datasets instead of the current 450 training datasets, which should increase the network's ability to generalize well on new patients. In addition, the network could learn the point matching as a multi-step temporal procedure by pairing up each sub-step with its subsequent sub-step. Despite not available across all patients, corresponding anatomical landmarks inside the prostate have been demonstrated to be helpful in MR-TRUS image registration. We plan to develop an automatic algorithm to detect and establish landmarks correspondences prior to registration to further improve the registration accuracy.

6. Conclusions

A novel deep learning-based 3D point cloud matching method was proposed for MR-TRUS image registration. Population-based FE models were conducted to simulate prostate deformations between MR and TRUS. The FE-generated deformation field was used to train the proposed network to apply biomechanical constraint on the 3D point matching. Our results demonstrated that the proposed method was able to accurately and rapidly register MR and TRUS prostate images for image-guided prostate intervention.

Disclosures

The authors declare no conflicts of interest.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships th t could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yabo Fu: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. **Yang Lei:** Conceptualization, Methodology, Software. **Tonghe Wang:** Data curation, Writing

- review & editing. **Pretesh Patel:** Data curation. **Ashesh B. Jani:** Data curation. **Hui Mao:** Writing - review & editing. **Walter J. Curran:** Project administration, Funding acquisition. **Tian Liu:** Project administration, Funding acquisition. **Xiaofeng Yang:** Resources, Conceptualization, Methodology, Supervision, Project administration, Funding acquisition, Writing - review & editing.

Acknowledgements

This research is supported in part by the National Cancer Institute of the National Institutes of Health under Award Number R01CA215718 (XY), the Department of Defense (DoD) Prostate Cancer Research Program (PCRP) Award W81XWH-17-1-0438 (TL) and W81XWH-17-1-0439 (AJ) and Dunwoody Golf Club Prostate Cancer Research Award (XY), a philanthropic award provided by the Winship Cancer Institute of Emory University.

References

Audenaert, E.A., Van Houcke, J., Almeida, D.F., Paelinck, L., Peiffer, M., Steenackers, G., Vandermeulen, D., 2019. Cascaded statistical shape model based segmentation of the full lower limb in CT. Comput. Methods Biomech. Biomed. Eng. 22, 644–657.

Bahn, D., de Castro Abreu, A.L., Gill, I.S., Hung, A.J., Silverman, P., Gross, M.E., Lieskovsky, G., Ukimura, O., 2012. Focal cryotherapy for clinically unilateral, low-intermediate risk prostate cancer in 73 men with a median follow-up of 3.7 years. Eur. Urol. 62, 55–63.

Bernhard, S., John, P., Thomas, H., 2007. Non-rigid point set registration: Coherent Point Drift. In: Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference. MITP, p. 1.

Besl, P.J., McKay, N.D., 1992. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14, 239–256.

Bloom, J.B., Gold, S.A., Hale, G.R., Rayn, K.N., Sabarwal, V.K., Bakhutashvili, I., Valera, V., Turkbey, B., Pinto, P.A., Wood, B.J., 2018. "Super-active surveillance": MRI ultrasound fusion biopsy and ablation for less invasive management of prostate cancer. Gland Surg. 7, 166–187.

Bookstein, F.L., 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans. Pattern Anal. Mach. Intell. 11, 567–585.

Brock, K.K., Mutic, S., McNutt, T.R., Li, H., Kessler, M.L., 2017. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. Med. Phys. 44, e43–e76.

Challapalli, A., Jones, E., Harvey, C., Hellawell, G.O., Mangar, S.A., 2012. High dose rate prostate brachytherapy: an overview of the rationale, experience and emerging applications in the treatment of prostate cancer. Br. J. Radiol. 85 Spec No 1, S18–S27.

Chui, H., Rangarajan, A., 2003. A new point matching algorithm for non-rigid registration. Comput. Vision Image Underst. 89, 114–141.

Dickinson, L., Hu, Y., Ahmed, H.U., Allen, C., Kirkham, A.P., Emberton, M., Barratt, D., 2013. Image-directed, tissue-preserving focal therapy of prostate cancer: a feasibility study of a novel deformable magnetic resonance-ultrasound (MR-US) registration system. BJU Int. 112, 594–601.

Dong, X., Lei, Y., Wang, T., Thomas, M., Tang, L., Curran, W.J., Liu, T., Yang, X., 2019. Automatic multiorgan segmentation in thorax CT images using U-net-GAN. Med. Phys. 46, 2157–2168.

Du, S., Zheng, N., Ying, S., Liu, J., 2010. Affine iterative closest point algorithm for point set registration. Pattern Recognit. Lett. 31, 791–799.

Fan, H., Su, H., Guibas, L.J., 2016. A Point Set Generation Network for 3D Object Reconstruction from a Single Image. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2463–2471.

Fu, Y., Lei, Y., Wang, T., Curran, W.J., Liu, T., Yang, X., 2020. Deep learning in medical image registration: a review. Phys. Med. Biol..

Fu, Y., Liu, S., Li, H.H., Li, H., Yang, D., 2018a. An adaptive motion regularization technique to support sliding motion in deformable image registration. Med. Phys. 45, 735–747.

Fu, Y., Mazur, T.R., Wu, X., Liu, S., Chang, X., Lu, Y., Li, H.H., Kim, H., Roach, M.C., Henke, L., Yang, D., 2018b. A novel MRI segmentation method using CNN-based correction network for MRI-guided adaptive radiotherapy. Med. Phys. 45, 5129–5137.

Fu, Y., Wu, X., Thomas, A.M., Li, H.H., Yang, D., 2019. Automatic large quantity landmark pairs detection in 4DCT lung images. Med. Phys. 46, 4490–4501.

Fu, Y.B., Chui, C.K., Teo, C.L., 2013. Liver tissue characterization from uniaxial stress–strain data using probabilistic and inverse finite element methods. J. Mech. Behav. Biomed. Mater. 20, 105–112.

Ghavami, N., Hu, Y., Gibson, E., Bonmati, E., Emberton, M., Moore, C.M., Barratt, D.C., 2019. Automatic segmentation of prostate MRI using convolutional neural networks: Investigating the impact of network architecture on the accuracy of volume measurement and MRI-ultrasound registration. Med. Image Anal. 58, 101558.

Gong, L., Wang, H., Peng, C., Dai, Y., Ding, M., Sun, Y., Yang, X., Zheng, J., 2017. Non–rigid MR-TRUS image registration for image-guided prostate biopsy using correlation ratio-based mutual information. Biomed. Eng. Online 16 (8).

Haskins, G., Kruecker, J., Kruger, U., Xu, S., Pinto, P.A., Wood, B.J., Yan, P., 2019a. Learning deep similarity metric for 3D MR–TRUS image registration. Int. J. Comput. Assist. Radiol. Surg. 14, 417–425.

Haskins, G., Kruger, U., Yan, P., 2019b. Deep Learning in Medical Image Registration: A Survey. ArXiv 1903.02026.

Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, S.M., Schnabel, J.A., 2012. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. Med. Image Anal. 16, 1423–1435.

Hu, Y., Ahmed, H.U., Taylor, Z., Allen, C., Emberton, M., Hawkes, D., Barratt, D., 2012. MR to ultrasound registration for image-guided prostate interventions. Med. Image Anal. 16, 687–703.

Hu, Y., Gibson, E., Ahmed, H.U., Moore, C.M., Emberton, M., Barratt, D.C., 2015. Population-based prediction of subject-specific prostate deformation for MR-to-ultrasound image registration. Med. Image Anal. 26, 332–344.

Hu, Y., Gibson, E., Ghavami, N., Bonmati, E., Moore, C.M., Emberton, M., Vercauteren, T., Noble, J.A., Barratt, D.C., 2018a. Adversarial deformation regularization for training image registration neural networks. MICCAI.

Hu, Y., Modat, M., Gibson, E., Li, W., Ghavami, N., Bonmati, E., Wang, G., Bandula, S., Moore, C.M., Emberton, M., Ourselin, S., Noble, J.A., Barratt, D.C., Vercauteren, T., 2018b. Weakly-supervised convolutional neural networks for multimodal image registration. Med. Image Anal. 49, 1–13.

Hu, Y., Morgan, D., Ahmed, H.U., Pendse, D., Sahu, M., Allen, C., Emberton, M., Hawkes, D., Barratt, D., 2008. A statistical motion model based on biomechanical simulations for data fusion during image-guided prostate interventions. Med. Image Comput. Comput. Assist. Interv. 11, 737–744.

Jian, B., Vemuri, B.C., 2011. Robust point set registration using gaussian mixture models. IEEE Trans. Pattern Anal. Mach. Intell. 33, 1633–1645.

Kadoury, S., Yan, P., Xu, S., Glossop, N., Choyke, P., Turkbey, B., Pinto, P., Wood, B.J., Kruecker, J., 2010. Realtime TRUS/MRI Fusion Targeted-Biopsy for Prostate Cancer: A Clinical Demonstration of Increased Positive Biopsy Rates. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 52–62.

Khallaghi, S., Sanchez, C.A., Rasoulian, A., Nouranian, S., Romagnoli, C., Abdi, H., Chang, S.D., Black, P.C., Goldenberg, L., Morris, W.J., Spadinger, I., Fenster, A., Ward, A., Fels, S., Abolmaesumi, P., 2015a. Statistical biomechanical surface registration: application to MR-TRUS fusion for prostate interventions. IEEE Trans. Med. Imaging 34, 2535–2549.

Khallaghi, S., Sánchez, C.A., Rasoulian, A., Sun, Y., Imani, F., Khojaste, A., Goksel, O., Romagnoli, C., Abdi, H., Chang, S., Mousavi, P., Fenster, A., Ward, A., Fels, S., Abolmaesumi, P., 2015b. Biomechanically constrained surface registration: application to MR-TRUS Fusion for prostate interventions. IEEE Trans. Med. Imaging 34, 2404–2414.

Kim, Y., Na, Y.H., Xing, L., Lee, R., Park, S., 2016. Automatic deformable surface registration for medical applications by radial basis function-based robust point–matching. Comput. Biol. Med. 77, 173–181.

Kohlrausch, J., Rohr, K., Stiehl, H.S., 2005. A New class of elastic body splines for nonrigid registration of medical images. J. Math. Imaging Vision 23, 253–280.

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. Med. Phys. 46, 3194–3206.

Li, H., Sumner, R.W., Pauly, M., 2008. Global correspondence optimization for non–rigid registration of depth scans. Comput. Graphics Forum 27, 1421–1430.

Liang, L., Liu, M., Martin, C., Sun, W., 2018. A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis. J. R. Soc. Interface 15.

Litjens, G., Toth, R., van de Ven, W., Hoeks, C., Kerkstra, S., van Ginneken, B., Vincent, G., Guillard, G., Birbeck, N., Zhang, J., Strand, R., Malmberg, F., Ou, Y., Davatzikos, C., Kirschner, M., Jung, F., Yuan, J., Qiu, W., Gao, Q., Edwards, P.E., Maan, B., van der Heijden, F., Ghose, S., Mitra, J., Dowling, J., Barratt, D., Huisman, H., Madabhushi, A., 2014. Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. Med. Image Anal. 18, 359–373.

Liu, X., Qi, C.R., Guibas, L.J., 2018. FlowNet3D: Learning Scene Flow in 3D Point Clouds. CVPR.

Loeckx, D., Slagmolen, P., Maes, F., Vandermeulen, D., Suetens, P., 2010. Nonrigid image registration using conditional mutual information. IEEE Trans. Med. Imaging 29, 19–29.

Mitra, J., Kato, Z., Marti, R., Oliver, A., Llado, X., Sidibe, D., Ghose, S., Vilanova, J.C., Comet, J., Meriaudeau, F., 2012. A spline-based non-linear diffeomorphism for multimodal prostate registration. Med. Image Anal. 16, 1259–1279.

Onofrey, J.A., Oksuz, I., Sarkar, S., Venkataraman, R., Staib, L.H., Papademetris, X., 2016. MRI-TRUS Image Synthesis with Application to Image-Guided Prostate Intervention. Springer International Publishing, Cham, pp. 157–166.

Ou, Y., Sotiras, A., Paragios, N., Davatzikos, C., 2011. DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. Med. Image Anal. 15, 622–639.

Pence, T.J., Gou, K., 2015. On compressible versions of the incompressible neo-Hookean material. Math. Mech. Solids 20, 157–182.

Qi, C., Su, H., Mo, K., Guibas, L.J., 2016. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 77–85.

Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. NIPS.

Reynier, C., Troccaz, J., Fourneret, P., Dusserre, A., Gay-Jeune, C., Descotes, J.L., Bolla, M., Giraud, J.Y., 2004. MRI/TRUS data fusion for prostate brachytherapy. Preliminary results.. Med. Phys. 31, 1568–1575.

Rivaz, H., Karimaghaloo, Z., Fonov, V.S., Collins, D.L., 2014. Nonrigid registration of ultrasound and MRI using contextual conditioned mutual information. IEEE Trans. Med. Imaging 33, 708–725.

Shaaer, A., Davidson, M., Semple, M., Nicolae, A., Mendez, L.C., Chung, H., Loblaw, A., Tseng, C.L., Morton, G., Ravi, A., 2019. Clinical evaluation of an MRI-to-ultrasound deformable image registration algorithm for prostate brachytherapy. Brachytherapy 18, 95–102.

Shahzad, M., Kamran, A., Siddiqui, M.Z., Farhan, M., 2015. Mechanical characterization and FE modelling of a hyperelastic material.

Singh, A.K., Kruecker, J., Xu, S., Glossop, N., Guion, P., Ullman, K., Choyke, P.L., Wood, B.J., 2008. Initial clinical experience with real-time transrectal ultrasonography-magnetic resonance imaging fusion-guided prostate biopsy. BJU Int. 101, 841–845.

Sun, Y., Yuan, J., Qiu, W., Rajchl, M., Romagnoli, C., Fenster, A., 2015. Three-dimensional nonrigid MR-TRUS registration using dual optimization. IEEE Trans. Med. Imaging 34, 1085–1095.

van de Ven, W.J., Hu, Y., Barentsz, J.O., Karssemeijer, N., Barratt, D., Huisman, H.J., 2015. Biomechanical modeling constrained surface-based image registration for prostate MR guided TRUS biopsy. Med. Phys. 42, 2470–2481.

van de Ven, W.J., Hulsbergen-van de Kaa, C.A., Hambrock, T., Barentsz, J.O., Huisman, H.J., 2013. Simulated required accuracy of image registration tools for targeting high-grade cancer components with prostate biopsies. Eur. Radiol. 23, 1401–1407.

Wang, B., Lei, Y., Tian, S., Wang, T., Liu, Y., Patel, P., Jani, A.B., Mao, H., Curran, W.J., Liu, T., Yang, X., 2019. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. Med. Phys. 46, 1707–1718.

Wang, Y., Cheng, J.Z., Ni, D., Lin, M., Qin, J., Luo, X., Xu, M., Xie, X., Heng, P.A., 2016. Towards Personalized Statistical Deformable Model and Hybrid Point Matching for Robust MR-TRUS Registration. IEEE Trans. Med. Imaging 35, 589–604.

Wang, Y., Zheng, Q., Heng, P.A., 2018. Online Robust Projective Dictionary Learning: Shape Modeling for MR-TRUS Registration. IEEE Trans. Med. Imaging 37, 1067–1078.

Weir, H.K., Thompson, T.D., Soman, A., Moller, B., Leadbetter, S., 2015. The past, present, and future of cancer incidence in the United States: 1975 through 2020. Cancer 121, 1827–1837.

Yacoub, J.H., Verma, S., Moulton, J.S., Eggener, S., Aytekin, O., 2012. Imaging-guided prostate biopsy: conventional and emerging techniques. Radiographics 32, 819–837.

Yan, P., Xu, S., Rastinehad, A.R., Wood, B.J., 2018. Adversarial Image Registration with Application for MR and TRUS Image Fusion. Springer International Publishing, Cham, pp. 197–204.

Yang, J., Cao, Z., Zhang, Q., 2016. A fast and robust local descriptor for 3D point cloud registration. Inf. Sci. 346-347, 163–179.

Yang, X., Akbari, H., Halig, L., Fei, B., 2011. 3D non-rigid registration using surface and local salient features for transrectal ultrasound image-guided prostate biopsy. Proc. SPIE Int. Soc. Opt. Eng. 7964, 79642v.

Yang, X., Rossi, P., Mao, H., Jani, A.B., Ogunleye, T., Curran, W.J., Liu, T., 2015. A MR-TRUS registration method for ultrasound-guided prostate interventions. Proc. SPIE Int. Soc. Opt. Eng. 9415.

Yang, X., Rossi, P., Ogunleye, T., Marcus, D.M., Jani, A.B., Mao, H., Curran, W.J., Liu, T., 2014. Prostate CT segmentation method based on nonrigid registration in ultrasound-guided CT-based HDR prostate brachytherapy. Med. Phys. 41, 111915.

Yang, X., Rossi, P.J., Jani, A.B., Mao, H., Zhou, Z., Curran, W.J., Liu, T., 2017. Improved prostate delineation in prostate HDR brachytherapy with TRUS-CT deformable registration technology: A pilot study with MRI validation. J. Appl. Clin. Med. Phys. 18, 202–210.

Yuan, J., Qiu, W., Rajchl, M., Ukwatta, E., Tai, X., Fenster, A., 2013. Efficient 3D endfiring TRUS prostate segmentation with globally optimized rotational symmetry. In: 2013 IEEE conference on computer vision and pattern recognition, pp. 2211–2218.

Zhang, S., Jiang, S., Yang, Z., Liu, R., 2015. In: 2D Ultrasound and 3D MR Image Registration of the Prostate for Brachytherapy Surgical Navigation, 94. Medicine, Baltimore, p. e1643.

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER

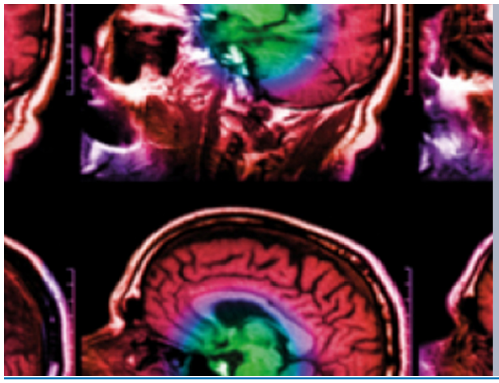
Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy

Recent citations

- Head and neck multiorgan auto segmentation on CT images aided by synthetic MRI
Yingzi Liu *et al*

To cite this article: Qiulan Zeng *et al* 2020 *Phys. Med. Biol.* **65** 135002

View the article online for updates and enhancements.



IPEM | IOP
Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics, biomedical engineering and related subjects.

Start exploring the collection–download the first chapter of every title for free.

This content was downloaded from IP address 163.246.0.131 on 08/12/2020 at 20:28

IOP Publishing Phys. Med. Biol. 65 (2020) 135002 https://doi.org/10.1088/1361-6560/ab8cd6

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

RECEIVED
9 December 2019

REVISED
13 April 2020

ACCEPTED FOR PUBLICATION
24 April 2020

PUBLISHED
22 June 2020

PAPER

Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy

Qiulan Zeng[1], Yabo Fu[1], Zhen Tian[1,2], Yang Lei[1], Yupei Zhang[1], Tonghe Wang[1,2], Hui Mao[2,3], Tian Liu[1,2], Walter J Curran[1,2], Ashesh B Jani[1,2], Pretesh Patel[1,2] and Xiaofeng Yang[1,2]

[1] 1Department of Radiation Oncology, Emory University, Atlanta, Georgia, United States of America
[2] 2Winship Cancer Institute, Emory University, Atlanta, Georgia, United States of America
[3] 3Department of Radiology and Imaging Sciences, Emory University, Atlanta, Georgia, United States of America

E-mail: xiaofeng.yang@emory.edu

Keywords: deformable image registration, weakly supervised method, prostate, MRI-TRUS, deep learning

Supplementary material for this article is available online

Abstract

Registration and fusion of magnetic resonance imaging (MRI) and transrectal ultrasound (TRUS) of the prostate can provide guidance for prostate brachytherapy. However, accurate registration remains a challenging task due to the lack of ground truth regarding voxel-level spatial correspondence, limited field of view, low contrast-to-noise ratio, and signal-to-noise ratio in TRUS. In this study, we proposed a fully automated deep learning approach based on a weakly supervised method to address these issues. We employed deep learning techniques to combine image segmentation and registration, including affine and nonrigid registration, to perform an automated deformable MRI-TRUS registration. To start with, we trained two separate fully convolutional neural networks (CNNs) to perform a pixel-wise prediction for MRI and TRUS prostate segmentation. Then, to provide the initialization of the registration, a 2D CNN was used to register MRI-TRUS prostate images using an affine registration. After that, a 3D UNET-like network was applied for nonrigid registration. For both the affine and nonrigid registration, pairs of MRI-TRUS labels were concatenated and fed into the neural networks for training. Due to the unavailability of ground-truth voxel-level correspondences and the lack of accurate intensity-based image similarity measures, we propose to use prostate label-derived volume overlaps and surface agreements as an optimization objective function for weakly supervised network training. Specifically, we proposed a hybrid loss function that integrated a Dice loss, a surface-based loss, and a bending energy regularization loss for the nonrigid registration. The Dice and surface-based losses were used to encourage the alignment of the prostate label between the MRI and the TRUS. The bending energy regularization loss was used to achieve a smooth deformation field. Thirty-six sets of patient data were used to test our registration method. The image registration results showed that the deformed MR image aligned well with the TRUS image, as judged by corresponding cysts and calcifications in the prostate. The quantitative results showed that our method produced a mean target registration error (TRE) of 2.53 $\pm$ 1.39 mm and a mean Dice loss of 0.91 $\pm$ 0.02. The mean surface distance (MSD) and Hausdorff distance (HD) between the registered MR prostate shape and TRUS prostate shape were 0.88 and 4.41 mm, respectively. This work presents a deep learning-based, weakly supervised network for accurate MRI-TRUS image registration. Our proposed method has achieved promising registration performance in terms of Dice loss, TRE, MSD, and HD.

© 2020 Institute of Physics and Engineering in Medicine

1. Introduction

High-dose-rate (HDR) brachytherapy has become a popular treatment modality for prostate cancer. Conventional transrectal ultrasound (TRUS)-guided prostate HDR brachytherapy could benefit significantly if the dominant intraprostatic lesions defined by multiparametric magnetic resonance imaging (MRI) can be incorporated into TRUS to guide HDR catheter placement.

The biggest challenge of MRI-TRUS registration is that it lacks robust image similarity measurement. In other words, a reliable image-intensity-based statistical correlation between MRI and TRUS images is not available. Thus, intensity-based models, which are based on optimizing image similarity (Brock *et al* 2017), show inferior performance for MRI-TRUS registration (Hu *et al* 2018b). Due to the lack of an image similarity measurement, it is also difficult to derive the driving force solely based on image intensity similarities for a physics-based model (Broit 1981, Sotiras *et al* 2013). In addition, supervised deep learning techniques are not feasible for MRI-TRUS registration due to the unavailability of ground truth deformation. To tackle this issue, research nowadays focuses on three popular approaches: knowledge-based models (Ferrant *et al* 2001, Mohamed *et al* 2002, van de Ven *et al* 2015, Fleute and Lavallee 1999, Rueckert *et al* 2003, Ashraf *et al* 2006), weakly supervised or unsupervised deep learning models (de Vos *et al* 2017, Yang *et al* 2017, Hu *et al* 2018a, 2018b, Lei *et al* 2019a), and the combination of those two approaches (Hu *et al* 2018).

Regarding the knowledge-based models, Mohamed *et al* (2002) firstly incorporated biomechanical model into SDMs to simulate prostate deformation. The main idea is to firstly use finite element analysis (FEA) to generate a series of deformations, which later are used to construct a statistical model, bypassing the similarity measure problem. However, one major issue in the biomechanical model is the uncertainty of biomechanical parameter setting. Hu *et al* (2011, 2012) attempted to address this issue by randomly sampling the biomechanical parameters from physically plausible parametric ranges. Further, Wang *et al* (2016) utilized patient-specific tissue parameters obtained with ultrasound elastography. Nonetheless, due to its unavailability in many hospitals and large variability in tissue stiffness measurements, the use of ultrasound elastography seems inconvenient and impractical for image registrations with respect to clinical practice.

Meanwhile, some researchers have been attempting to apply deep learning techniques for MRI-TRUS registration. Hu *et al* proposed a weakly supervised dense correspondence learning method with a fully convolutional neural network (FCN) (Hu *et al* 2018a, 2018b). The idea of this method is to use labels to represent anatomical structures (Zhou 2018), and then train a registration neural network to register those labels; in the inference stage, the MR and TRUS images are input into the trained registration neural network to predict the deformation occurring during prostate cancer interventions.

This weakly supervised method avoids the intensity-based similarity measure. In addition, their method is fully automated without registration initialization. However, the registration accuracy of their results was rather limited. In their work, they used whole image pairs as the input. The size of the training samples, which are high dimensional, is small and may be insufficient. This relatively small training data size could cause the network to be susceptible to overfitting. To alleviate this problem, for training the neural network, they labeled more than 4000 pairs of diverse anatomical landmarks from 111 pairs of T2-weighted MR and 3D TRUS images (Hu *et al* 2018a). However, the manual labeling process is time-consuming and laborious. Even with so many anatomical landmarks, the registration results (a median target registration error of 4.2 mm on a landmark) are far from meeting the clinical requirement of 1.9 mm, which is required to correctly grade 95% of the aggressive tumor components (van de Ven *et al* 2013). Another reason for the limited registration accuracy may be the labeling uncertainty. Though a large number of manual labels have been obtained for training, many of them such as cysts and calcification deposits, were not readily identifiable or even unreliably labeled. The inaccurate labels may result in the degradation of the performance of the neural network prediction . Furthermore, the training labels overlapped with each other, which may lead to the degradation of the performance of the registration.

Apart from the above-mentioned methods, recently, Hu *et al* (2018) combined the knowledge-based models with deep learning techniques to perform the MRI-TRUS image registration. In this study, FEA was used to produce a series of deformations, and then an adversarial convolutional neural network (CNN) was designed to train for image registration. The results do not seem very accurate, partly due to the above-mentioned limitation in the biomechanical model, the network architecture they employed, and the limited dataset.

In addition to the difficulties of designing accurate and efficient methods, the great challenges of implementing robust and reliable evaluation metrics exist for MRI-TRUS image registration (Brock *et al* 2017, Paganelli *et al* 2018). Generally, qualitative and quantitative validation can be integrated to evaluate the overall registration process (Paganelli *et al* 2018). For qualitative validation, we can use split screen and checkerboard displays, image overlay displays, difference image displays, and contour/structure mapping displays (Paganelli *et al* 2018). For quantitative validation, the following metrics are available: target

registration error (TRE), mean distance to agreement, the Dice similarity coefficient, the Jacobian determinant, consistency, etc. For the details of each definition of the metrics, please refer to (Paganelli *et al* 2018).

We make the following contributions: (1) We proposed a novel fully automated MRI-TRUS image registration method. Specifically, we combined three major networks: one for prostate contour segmentation, one for affine registration, and the third one for nonrigid registration, into a workflow of MRI-TRUS registration. In the inference stage of segmentation, a pair of MRI-TRUS images was required; in the inference stage of affine and nonrigid registration, the pair of MRI-TRUS labels obtained from the segmentation step were treated as inputs into the registration network to generate a dense displacement field (DDF). The workflow is as follows: we firstly utilized two separate FCNs to extract the hierarchical feature maps to segment the MRI-TRUS prostate labels. To provide a better global registration initialization, we employed a 2D CNN to perform affine registration. A 3D UNET-like network was applied afterward to achieve fine local nonrigid registration. (2) We conducted a comparison qualitatively and quantitively among our method, Hu's fullyautomated registration method with a composite neural (CN) network (Hu *et al* 2018a, 2018b), and a point matching (PM) method. Additionally, we also compared two variants of our method, one using only pairs of MRI-TRUS prostate labels as inputs of registration networks and the other one using only pairs of MRI-TRUS images as inputs.

The remaining content is organized into four sections as follows: section 2 introduces our proposed method; section 3 presents the experimental results; section 4 provides a discussion, followed by a conclusion in section 5.

2. Methods

Our method combines three major networks into the workflow of MRI-TRUS registration. These networks are described as follows: first, we utilized two separate whole-volume-based 3D FCNs to perform prostate segmentation on MRI and TRUS images; second, we applied a 2D CNN to train an affine registration network using the segmented MRI and US prostate labels; third, we employed a 3D UNET-like network to train a nonrigid registration network using the same MRI-TRUS labels. All the above-mentioned networks are trained separately.

2.1. FCN for segmentation

FCNshave demonstrated promising performances in medical image automated segmentation. Dense pixel-wise prediction enables the FCN to have end-to-end predictions from the whole image in a single forward pass. We separately trained two FCNs for MRI and TRUS segmentation. Manual prostate labels of MRI and TRUS images were used as the learning targets for the two FCNs. A 3D supervision mechanism was integrated into the FCN's hidden layers to facilitate informative features extraction. We combined a binary cross-entropy loss and a batch-based Dice loss into a hybrid loss function for deeply supervised training. The implementation details of the FCN have been presented in our previous studies (Wang *et al* 2019, Lei *et al* 2019b). Once trained, the FCNs were able to rapidly segment the prostate from whole volume MRI and TRUS images of a new patient.

2.2. Affine registration

Magnetic resonance (MR) and TRUS labels were designated as moving and fixed labels, respectively, since we aim to use the warped preoperative MRI, which matches the intra-operative TRUS, as image guidance during the procedure. According to research (van de Ven *et al* 2015, Wang *et al* 2016), it is very helpful to set a good rigid or affine registration as an accurate initialization prior to subsequent nonrigid registration. Therefore, we designed a 2D CNN which takes 3D image as inputs to predict 12 affine transformation parameters to warp the original MR images and labels, as an initialization step of nonrigid registration. The automatic affine registration aimed to optimize the Dice similarity coefficient. The Dice similarity coefficientrepresents the overlap between the binary fixed and warped labels. Its definition is given as follows:

$$C_{Dice}(A, B) = \frac{2\sum_{i=1}^{I} a_i \cdot b_i}{\sum_{i=1}^{I} a_i + \sum_{i=1}^{I} b_i}, \quad (1)$$

where the a and b denote the binary value (0 or 1) for a voxel in a moving label and fixed label, respectively, and I is the total number of voxels over the label.

The architecture is shown in figure 1. The MRI and TRUS images are concatenated as 'two-channel' images as inputs. The tensor shape of the input is [4, 120, 108, 100], where 4 is the batch size, 120 and 108 are the number of voxels in the x- and y-directions, respectively, and 100 is the number of voxels in the

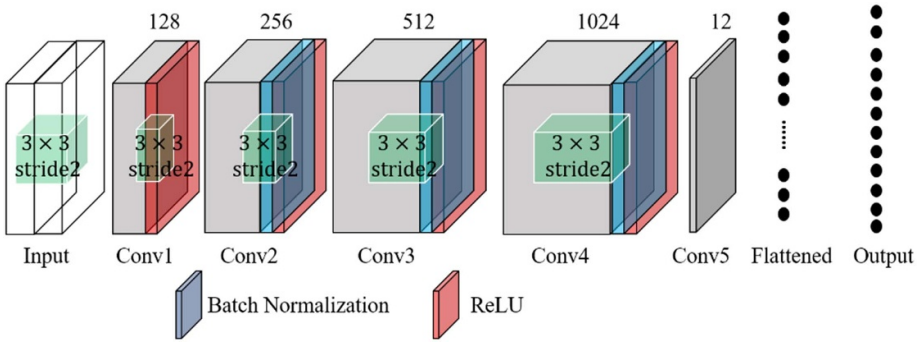


Figure 1. CNN for affine registration.

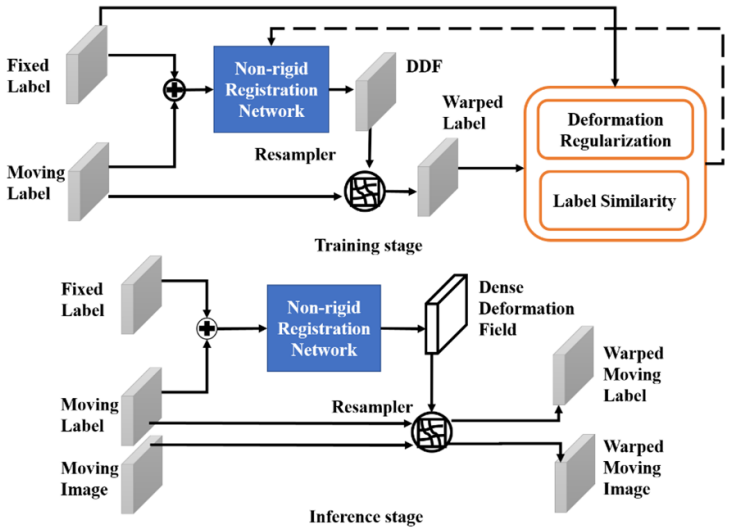


Figure 2. Nonrigid registration using labels as inputs into the networks.

z-direction, representing the number of channels. The number of filters used for the five convolutional layers was 128, 256, 512, 1024, and 12 respectively. These convolutional layers are followed by a flatten layer and a fully connected layer. The first four convolutional layers are followed by a batch normalization and a rectified linear unit as the activation. All the convolutional filters with a size of 3×3 and a stride of 2. The number of units for the fully connected layer is 12, which equals that of the 3D affine transformation parameters.

2.3. 3D UNET-like network for nonrigid registration

We compared two frameworks for nonrigid registration. One framework is to use only MRI-TRUS images as network inputs, similar to that of (Hu *et al* 2018b), and the other framework is to use pairs of MRI-TRUS labels as the input, as shown in figure 2. The nonrigid registration network architecture (Hu *et al* 2018b) is illustrated in figure 3. The MRI-TRUS labels are concatenated and input into a network similar to a U-NET, which has four down-sampling blocks followed by four up-sampling blocks. Compared to U-NET, this nonrigid neural network is more densely connected. In addition, it has two more types of summation-based residual shortcuts: one type is the standard residual network shortcutting two sequential convolution layers in each block; the other is the trilinear additive up-sampling layers, which shortcut the deconvolution layers in adjacent up-sampling blocks and are added onto the deconvolution layers (Wojna *et al* 2019).

The Dice loss was employed as part of label similarity cost function. The Dice loss could encourage the network to have a high volume overlap between the warped MR prostate label and the fixed TRUS prostate label. However, a high volume overlap does not necessarily translate into good surface matching. Therefore, we have proposed including surface PM as an additional loss function to encourage prostate shape matching. The surface-based loss function is defined as the distance between the specific prostate surface points on the

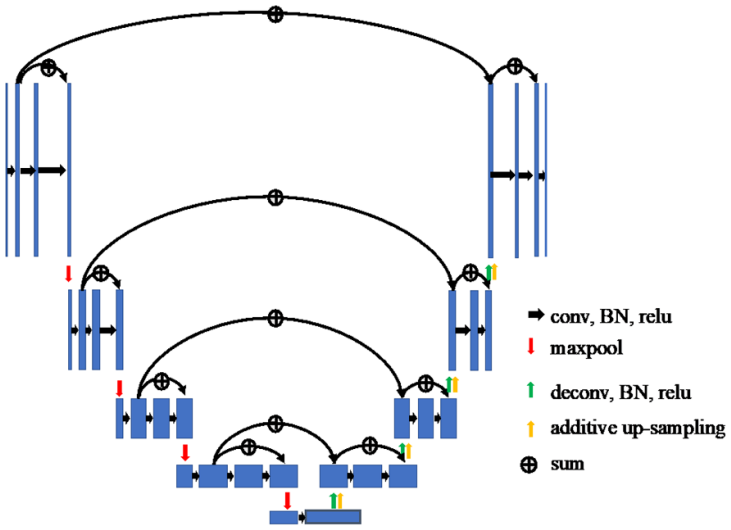


Figure 3. Illustration of the nonrigid neural network.

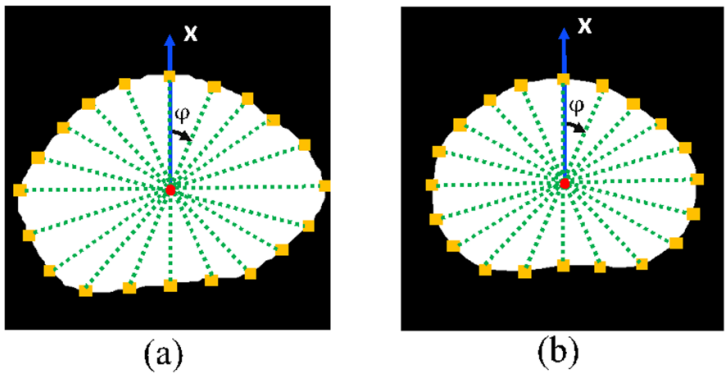


Figure 4. Prostate surface point (yellow square) projection at the *X*-*Y* plane at multiple directions on the (a) MRI and (b) TRUS images. Red dots indicate centroids of the prostate; green dashed lines denote line directions.

registered MRI and the corresponding surface points on the TRUS in certain directions. This cost function can be given as follows:

$$C_{surface}(A,\ B)\ = \frac{1}{L}\sum_{i=1}^{L}\left(d\left(P_i^A, P_i^B\right)\right), \tag{2}$$

where P_i^A and P_i^B indicate the surface point in the i-direction in image A or B, and L is the total number of directions, $\mathrm{d}(\cdot)$ is the L_2 norm Euclidean distance. In our work, we consider 40 different directions (the projection on the x-y plane is shown in figure 4). Using (θ, φ) to denote the inclination and azimuth, respectively, the θ and φ of these lines can be given as follows:

$$\begin{cases} \theta \in \{45^\circ, -45^\circ\} \\ \varphi = n * 18^\circ, n \in \{1, \ldots, 20\} \end{cases}. \tag{3}$$

Radial projection was used to find the corresponding prostate surface points in these directions. We have observed that 40 surface control points are adequate given the smooth prostate deformation. More surface points can provide very marginal benefits (see supplementary material at stacks.iop.org/PMB/65/135002/mmedia). Therefore, we empirically chose 40 surface points.

Since image registration is an ill-posed problem, regularization is necessary (Sotiras *et al* 2013). In our work, we added bending energy as a regularization term in the loss function, to smooth the deformation field

and penalize non-regular Jacobian values for the transformations (Christensen and Johnson 2001). The formula of the bending energy is given as follows:

$$C_{smooth} = \frac{1}{V}\int_0^X \int_0^Y \int_0^Z \left[\begin{array}{l} \left(\frac{\partial^2 T}{\partial x^2}\right)^2 + \left(\frac{\partial^2 T}{\partial y^2}\right)^2 + \left(\frac{\partial^2 T}{\partial z^2}\right)^2 \\ +2\left(\frac{\partial^2 T}{\partial x \partial y}\right)^2 + 2\left(\frac{\partial^2 T}{\partial x \partial z}\right)^2 + 2\left(\frac{\partial^2 T}{\partial y \partial z}\right)^2 \end{array} \right] dxdydz, \quad (4)$$

where T is the transformation, V is penalty of non-regular Jacobian values, and C represents the cost. We used finite central difference to approximate the second order partial derivatives terms in equation (4).

Given the Dice loss, surface-based loss, and the bending energy loss, we can write the total cost function as

$$C_{total} = -C_{Dice} + C_{surface} + \lambda C_{smooth}, \quad (5)$$

where λ is the smooth regularization weight.

2.4. Dataset

Experiments were conducted on a dataset of 36 pairs of T2-weighted MR and TRUS images collected from 36 prostate cancer patients who have been treated with HDR brachytherapy. TRUS data were acquired with a Hitachi HI VISION with a voxel size of $0.12 \times 0.12 \times 2.0\,mm^3$. The T2-weighted MR images were obtained using a Siemens Avanto 1.5 T scanner (Spin-echo sequence with a repetition time/echo time of: 1200 ms/123 ms, flip angle 150°, voxel size $1 \times 1 \times 1\,cm^3$ with each slice of 256×256 pixels, and pixel bandwidth 651 Hz), and then resampled to the same sizes and resolutions as those of the TRUS images. Both the original MR and TRUS images were reconstructed into a 3D volume and resampled to $0.5 \times 0.5 \times 0.5\,mm^3$ isotropic voxels by a third order spline interpolation. The manual prostate labels of the TRUS and MRI, represented by binary masks, were contoured by one and three radiologists, respectively, using VelocityAI 3.2.1 (Varian Medical Systems, Palo Alto, CA). In addition, the TRUS and MR labels were resampled to $0.5 \times 0.5 \times 0.5\,mm^3$.

Our proposed methods were implemented in TensorFlow with a 3D image augmentation layer from an open-source code in NiftyNet (Gibson *et al* 2018). The augmentation generated 300 times more training datasets. Each network was trained with a 12 GB NVIDIA Quadro TITAN Linux general-purpose graphic process unit.

2.5. Experiments

We used leave-one-out cross-validation for the registration network. Regarding the training of the network, we utilized the Adam optimizer with a learning rate of 10^{-5}. In addition, a trilinear resampled module was implemented, and a slip boundary condition was applied on the boundary of the grid. The initial values for all network parameters, except those in the final displacement prediction layers, were assigned using an Xavier initializer (Glorot and Bengio 2010). All the applicable hyper-parameters were kept same between our proposed method and the to-be-compared ones, unless otherwise stated.

We performed quantitative and qualitative evaluation of image registration accuracy. For quantitative evaluation, we employed the following metrics: target registration error (TRE), Dice loss, mean surface distance (MSD), and Hausdorff distance (HD) (Huttenlocher *et al* 1993, Litjens *et al* 2014). The target registration error (TRE) is defined as root-mean-square distance over all paris of landmarks in the MRI-TRUS images for each patient. Such landmarks include anatomical structures such as urethra, calcifications and cysts, etc In this study, an experienced radiologist carefully selected reliable landmarks on MR and TRUS images for TRE calculation. The MSD measures the average surface distance between two surfaces, and HD is defined as the greatest of all the distances from a point on one surface to the closest point on another surface. In contrast, the qualitative evaluation was mainly based on the visualization of contour overlays and image fusion. For example, in fusion images, a set of landmarks were denoted on registered MRI and US images, and a well alignment of the landmarks counterparts indicates good inner prostate registration.

3. Results

3.1. Registration performance

In this subsection, we will compare the quantitative results and qualitative results with our datasets from our method using MRI-TRUS labels as inputs for affine and nonrigid registration (SR-L), our method using MRI-TRUS images as inputs for affine and nonrigid registration (SR-I), Hu's fully automated method with a CN network (Hu *et al* 2018a, 2018b), and the point matching (PM) method.

Table 1. The MSD and HD values for the CN network, PM, and our segmentation-registration methods (SR-I and SR-L).

	MSD (mm)					HD (mm)				
	Mean	Std	Median	IQR	p-values	Mean	Std	Median	IQR	p-values
Affine	1.14	0.38	1.05	0.40	<0.001	5.49	1.70	5.42	1.81	<0.001
CN	2.31	0.64	2.28	0.93	<0.001	8.38	1.94	7.97	2.74	<0.001
PM	0.37	0.12	0.33	0.16	<0.001	3.41	1.54	3.20	2.31	<0.001
SR-I	**1.05**	**0.29**	**1.03**	**0.33**	**<0.001**	**4.86**	**1.49**	**4.66**	**1.30**	**0.050**
SR-L	**0.88**	**0.20**	**0.84**	**0.27**	**—**	**4.41**	**1.05**	**4.21**	**1.48**	**—**

Table 2. The measured Dice and TRE values for affine registrationthe CN) network, PM, and our segmentation-registration methods (SR-I and SR-L).

	DICE					TRE (mm)				
	Mean	Std	Median	IQR	p-values	Mean	Std	Median	IQR	p-values
Affine	0.88	0.04	0.89	0.05	<0.001	2.93	1.20	2.61	1.34	0.05
CN network	0.76	0.06	0.76	0.04	<0.001	6.04	3.14	6.02	5.14	<0.01
PM	0.96	0.02	0.96	0.03	<0.001	4.46	2.09	3.86	3.07	<0.01
SR-I	**0.89**	**0.03**	**0.90**	**0.05**	**<0.001**	**2.85**	**1.72**	**2.43**	**1.81**	**0.330**
SR-L	**0.91**	**0.02**	**0.90**	**0.03**	**—**	**2.53**	**1.39**	**2.38**	**1.41**	**—**

3.1.1. Quantitative results

Our segmentation neural networks generated Dice losses of 0.88 ± 0.05 and 0.92 ± 0.03 for the MRI and for TRUS segmentation, respectively (Wang *et al* 2019, Lei *et al* 2019b). On the training dataset, using SR, the total loss, Dice similarity loss, surface loss, and regularization loss converged at around 100 epochs with approximate values of 0.165, 0.08, 0.08, and 0.005, respectively. On the validation dataset, these loss functions converged at around 100 epochs with values of 0.196, 0.09, 0.100, and 0.006, respectively. Table 1 compares the registration results in terms of MSD and HD (Huttenlocher *et al* 1993, Litjens *et al* 2014), obtained with the CN, PM, SR-I, and SR-L methods. The MSD and HD from our proposed SR-L method are 0.88 and 4.41 mm, respectively, and these values obtained with SR-I method are 1.05 and 4.86 mm, respectively. On the other hand, the MSD and HD generated by the CN network are 2.31 and 8.38 mm; whereas these values obtained with PM are 0.37 and 3.41 mm. We also present the results after only affine registration: the MRD and HD are 1.14 and 5.49 mm, respectively. In addition, we applied paired Wilcoxon signed-rank tests to compare the leave-one-out cross-validation results between affine/CN/PM/SR-I and SR-L, as given by the fourth sub-column for each metric in table 1. The results show that both the resulting p-values of the pairs of SR-CN, and SR-PM are smaller than the significant level. Consequently, our proposed SR-L method is significantly better than the affine/CN/SR-I method, but not as good as PM in terms of a surface-based metric.

Table 2 presents the results in terms of Dice and TRE. SR-L achieves a mean Dice of 0.91 ± 0.02, compared to that of 0.89 ± 0.03 from SR-I, 0.76 ± 0.06 from the CN network, and 0.96 ± 0.02 from PM. In addition, SR-L obtained a mean TRE of 2.53 ± 1.39 mm, compared to that of 2.85 ± 1.72 mm from SR-I, 6.04 ± 3.14 mm from the CN network, and 4.46 ± 2.09 mm obtained with PM. Also, the median values and the first- and third quartiles are also listed. For example, SR-L generates a median Dice of 0.90 with the interquartile range (IQR) being 0.03, and a median TRE of 2.38 mm with the IQR being 1.41 mm. More detailed results are given in table 2 and figure 5. Again, paired Wilcoxon signed-rank tests have been done and show that both the p-values in terms of and the TRE of the pairs SR-L and CN network, and SR-L and PM are smaller than the significant level. As a result, we can conclude that in terms of both surface-based and volume-based metrics, our proposed method is significantly better than the CN network method, but not comparable to PM; whereas regarding TRE, our proposed method is significantly better than both the CN network and PM methods. The minimal possible TRE is dependent on the image voxel sizes.

We performed additional experiments to test the effectiveness of our surface loss terms. Without the surface loss term, the mean MSD, HD, Dice, and TRE were 1.10, 4.90, 0.89, and 2.86 mm, respectively. The slightly worse metrics results without surface loss show the efficacy of the surface loss.

3.1.2. Qualitative results

Figures 6 and 7 present a qualitative visual assessment of the results on the test dataset. Each row represents a patient case. In figure 6, from the left to the right column, we show the transverse slice in the original MR image, registered MR images from the CN network, PM, SR-I, SR-L methods, and the TRUS image. The dotted lines indicate the corresponding prostate gland contour. In addition, figure 7 demonstrates their

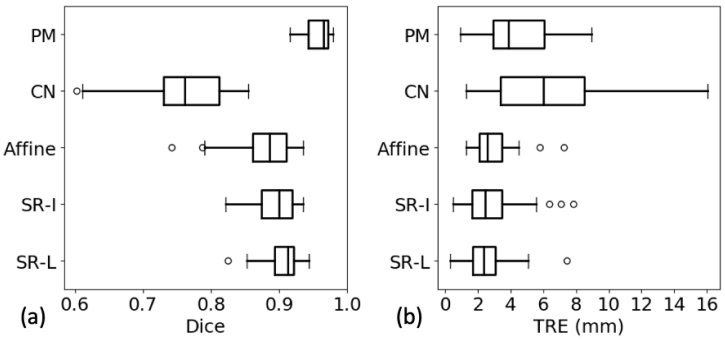


Figure 5. Boxplots of the leave-one-out registration results obtained with affine registration, the SR-I, SR-L, CN network, and PM methods. (a) Dice similarity coefficient in [0, 1], (b) TRE in mm.

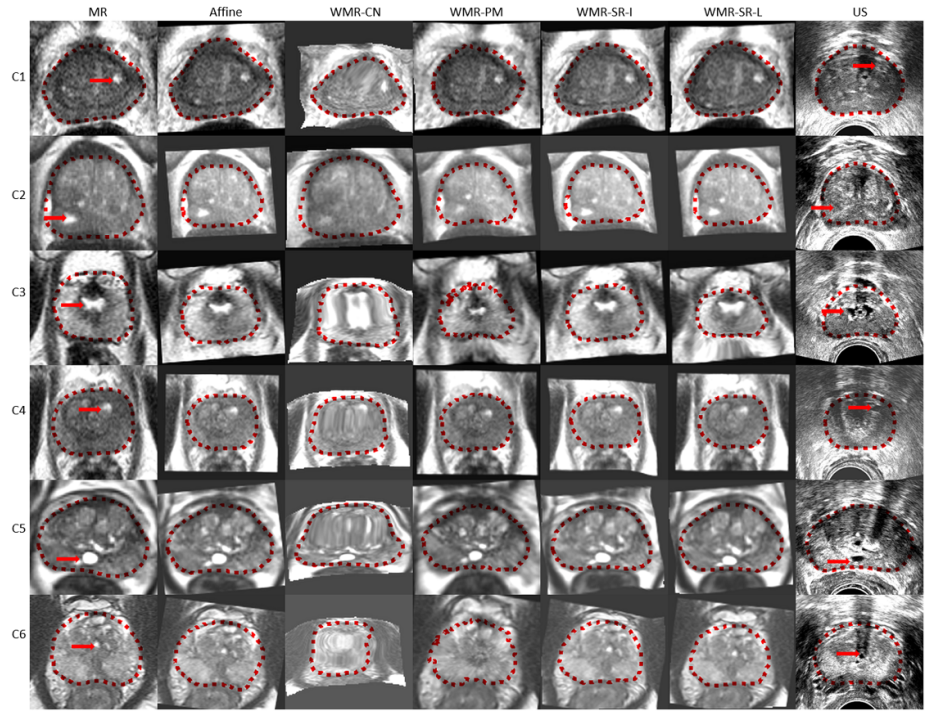


Figure 6. Example of the registered results from six cases (each row represents the results from one case). From the left to the right column in the figure, we show the transverse slice in original MR image, warped MR images from the affine, CN network (WMR-CN), PM (WMR-PM), SR-I (WMR-SR-I), and SR-L (WMR-SR-L) methods, and the TRUS image. The dotted lines indicate the corresponding prostate gland contour, and the red arrows point out the pairs of landmarks on the MRI-TRUS images.

corresponding fusion registered MRI-TRUS images from the affine, CN network, PM, SR-I, and SR-L methods. The yellow arrow points to some identified landmarks in the test dataset. 108 landmarks were manually selected for validation. The warped MR prostate labels obtained by PM are most similar to the labels on the TRUS images, whereas the warped labels obtained with the CN networkhave the least resemblance with the TRUS labels. In addition, figures 6 and 7 show that not only does the CN network generate warped labels that highly disagree with the fixed labels, but the CN network also produces blurred warped MR images. This registration performance implies a physically implausible deformation caused by the CN network. Apart from the unsatisfied registered results from the CN network presented in this work, results reported by Hu *et al* (2018a, 2018b) seem also non-realistic. In contrast to PM and the CN network, our proposed SR methods can generate realistic warped MR images with relatively good alignment between MRI-TRUS labels. In addition, the well-aligned landmark indicates the efficacy of our proposed method.

Furthermore, figure 8 demonstrates the results of other validation metrics that are based on DDF. From the top to bottom rows, we plotted the warped intensity MRI, Jacobian determinants, and the magnitudes of DDF, obtained with the CN network (first column), PM (second column), SR-I (third column), and SR-L (fourth column), respectively. No negative Jacobian determinants were found in SR (the minimum is 0.04

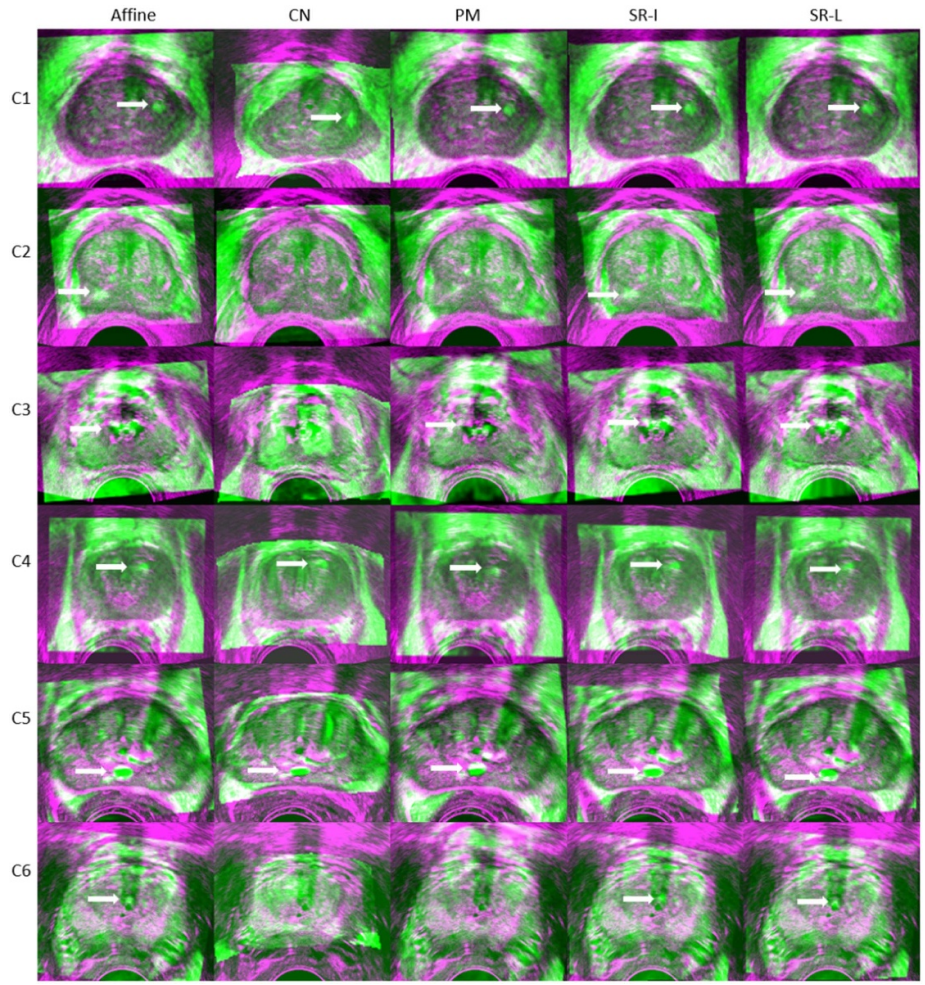


Figure 7. Fusion of registered MRI-TRUS images from the affine, CN network, PM, SR-I, and SR-L methods. The arrows denote the landmarks which are well aligned post registration.

across all patient cases), supporting that the DDF by SR is physically plausible. On the contrary, zero Jacobian determinants appeared in the CN network, implying that physically implausible deformation may exist.

4. Discussion

In this paper, we proposed a new MRI-TRUS image registration framework, following a step-by-step procedure including segmentation, affine registration, and nonrigid registration. Specifically, the framework consists of three major networks: one for prostate contour segmentation, one for affine registration, and the third one for nonrigid registration, and these networks were built into a pipeline to automate the workflow of MRI-TRUS registration. In the inference stage, pairs of MRI-TRUS images are required without the need for manual contour segmentation or manual initialization. Experiments on 36 patients indicate that our proposed method has promising registration performance in terms of registration accuracy.

For comparison purposes, in table 3, we summarized and compared our results in terms of and the TRE with the published results obtained with some other previously proposed methods (Hu *et al* 2012, 2018, 2018b, Khallaghi *et al* 2015, Sun *et al* 2015, van de Ven *et al* 2015, De Silva *et al* 2017) which avoided using intensity-based registration. We skip the results obtained with intensity-based methods due to their poor performance in registering MR and TRUS images. For example, Hu *et al* (2018b) reported that all nine intensity-based methods that they tested produced median TREs larger than 24 mm and median Dice lower than 0.77. Thus, table 3 only presents previously proposed methods based on biomechanical models, statistical deformation models, and deep learning models without intensity-based loss functions.

For methods based on biomechanical models or statistical deformation models (Hu *et al* 2012, Khallaghi *et al* 2015, Sun *et al* 2015, van de Ven *et al* 2015, De Silva *et al* 2017), which are not fully automated, an expected-TRE range of 1.4–2.8 mm was reported, and the experiments were conducted on 8–29 cases. The expected TRE from these methods seem small and satisfactory. Nonetheless, these methods usually rely on

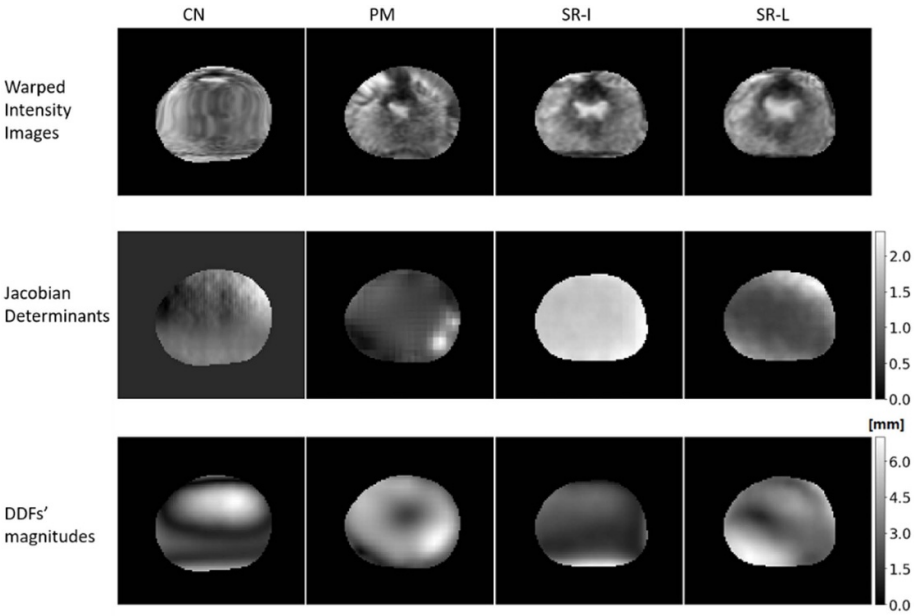


Figure 8. Inspection of the warped MRI (first row), Jacobian determinants (second row), and DDF's magnitudes (third row) from the CN network (first column), PM (second column), SR-I (third column), and SR-L (fourth column).

Table 3. Summary of MRI-TRUS registration results.

Registration method	DICE	TRE (mm)	No. of cases	Initialization method
Hu *et al* (2012)	n/a	2.4 (median)	8	Manual landmarks
Khallaghi *et al* (2015)	n/a	2.4 (mean)	19	Gland centroid
van de Ven *et al* (2015)	n/a	2.8 (median)	10	Rigid surface registration
Sun *et al* (2015)	0.86 (median)	1.8 (median)	20	Manual landmarks
Wang *et al* (2016)	n/a	1.4 (mean)	18	Rigid surface registration
De Silva *et al* (2017)	n/a	2.3 (mean)	29	Learned motion model
Hu *et al* (2018a)	0.88 (median)	4.2 (median)	76	n/a
Hu *et al* (2018b)	0.82 (median)	6.3 (median)	76	n/a
Composite network	0.76 (median)	6.0 (median)	36[a]	n/a
PM method	0.96 (median)	3.9 (median)	36[a]	Affine surface registration
Our method—image inputs	0.89 (mean)	2.9 (mean)	36[a]	n/a
Our method—labels inputs	**0.91 (mean)**	**2.5 (mean)**	36[a]	n/a

[a]Testing with our dataset.

manual or customized optimization initialization. For instance, to the best of the authors' knowledge, based on the biomechanical model, Wang *et al* (2016) reported the lowest expected TRE. However, their method relies on the rigid surface registration and requires prostate surface segmentation. In addition to the dependence on the initialization, biomechanical models have their clinical limitations due to the requirement of patient-specific biomechanical simulation data, which are not readily available for practical clinical application. As a result, these methods may suffer from inconvenience or infeasibility for practical clinic application.

As for the PM method, although it generates better results in terms of surface-based metrics such as Dice, MSD, and HD than those by our method, it does not mean that PM enables a good match inside the prostate, which can be indicated by the TRE. The PM method has the following drawbacks: (1) it takes only the surface points into account without considering inner prostate matching; as a result, it is likely to have non-smooth deformations inside the prostate gland for moving images, as shown in figure 6. (2) Since the PM method applies strong local shape alignment, it may be oversensitive to erroneous prostate segmentation. On the contrary, our method can yield much better results than PM regarding the TRE, implying superior registration to PM. This indicates that the bending energy in our network is better than the interpolation used in PM.

In contrast to the above-mentioned non-fully automated methods, recently, Hu *et al* proposed a fully automated deep learning-based method with a CN network. This method does not require any initialization or pre- or intra-procedural segmentation. With our dataset, the CN network generated a relatively large

expected TRE of 6.0 mm. This is probably due to the incompetence of their network architecture for the extraction of hierarchical feature from the 3D contexts. This method does not have any segmentation prior to the registration. However, during the registration, their network tries to minimize a cost function based on the constructed segmentation that is determined by the generated DDF and the 'ground-truth' segmentation. In this way, in the inference stage, without the ground-truth segmentation, the relation between the inputs and cost function is too weak to allow backpropagation. In other words, this network may suffer from the ill-posedness issue, which impedes its ability to extract hierarchical features for generating accurate DDF.

Compared to methods in other works, our novel method is fully automated, accurate, and has sound feasibility for practical clinical applications. The benefits of our method are as follows. Firstly, by integrating the deep learning image segmentation technique with the FCN into our registration workflow, we were able to extract hierarchical features to explicitly obtain the prostate labels in the MRI-TRUS, which is one of the key steps for successful registration. Secondly, to provide better registration initialization for nonrigid registration, we trained a separate CNN for affine registration. We argue that these two benefits enable our method higher registration accuracy than that of another fully automated method (Hu *et al* 2018a, 2018b). The improved registration accuracy indicates a step forward toward the clinical requirement. Thirdly, similar to that work (Hu *et al* 2018a, 2018b), to overcome the lack-of-similarity ground-truth issue, we used the weakly supervised method that leverages labels to represent the anatomical structures, along with bending energy regularization. Our method yielded smoother deformation within the contour than that by PM. Fourth, our work does not rely on FEA for model training, thereby avoiding the measurement of patient-specific tissue elastic parameters.

Nevertheless, there are several limitations to our proposed methods. Firstly, although the intensity-based similarity measure is bypassed, the TRE was relatively large since the network was not fully knowledgeable about the inner prostate deformation pattern. Obtaining reliable multimodal image mapping relations on the voxel level will help improve the multimodal image registration accuracy. This may be done by the biomechanical model or the feature-based image registration method. Secondly, we used Dice loss and a surface-based loss to maximize the similarity between the MRI and TRUS labels. On addition to that, the registration model may benefit a lot by including the physically correspondence loss into the loss function. Thirdly, we used bending energy as the regularization to achieve a topologically smooth deformation field. Again, it may help for the prediction of physically plausible deformation by incorporating biomechanical regularization terms into the cost function, apart from bending energy regularization. Finally, we have a limited dataset. A limited dataset may cause neural networks to suffer from overfitting. To reduce overfitting, we implemented data augmentation. With data augmentation, the mean Dice after nonrigid registration we obtained for the testing dataset is 0.91, compared to that of 0.87 without data augmentation. However, data augmentation based on the existing dataset could only mitigate the problem. To reduce overfitting and increase the network's generalizability, we plan to collect more datasets in our future work to help improve the performance for application in clinic practice.

5. Conclusion

This work presents a novel fully automated and accurate network with a weakly supervised method for MRI-TRUS image registration. The MRI-TRUS registration problem is difficult due to the considerably distinctive modalities of the two matching images, which in turn lead to little useful image features of prostates for the registration process. Our approaches are capable of solving this difficult problem by using a weakly supervised method to bypass the intensity similarity measurements, and by using deep learning to learn the complex image features for consistent prostate segmentation. What is more, we make the first attempt at integrating the automatic deep learning-based segmentation method into the registration procedure. By doing this, we not only enable a fully automated registration procedure, but also obtain accurate registration results. In addition, our proposed registration method can potentially be extended to other intervention procedures of soft tissue organs such as liver and breast. Future research aims to study the combination of our deep learning-based model with the biomechanical model or the feature-based image registration method to improve the registration accuracy and to make it available for a wider range of applications.

Acknowledgments

This research is supported in part by the National Cancer Institute of the National Institutes of Health under Award Number R01CA215718 (XY), the Department of Defense (DoD) Prostate Cancer Research Program (PCRP) Award W81XWH-17-1-0438 (TL) and W81XWH-17-1-0439 (AJ) and the Dunwoody Golf Club

Prostate Cancer Research Award (XY), a philanthropic award provided by the Winship Cancer Institute of Emory University.

ORCID iD

Xiaofeng Yang https://orcid.org/0000-0001-9023-5855

References

Ashraf M, Zacharaki E I, Shen D G and Davatzikos C 2006 Deformable registration of brain tumor images via a statistical model of tumor-induced deformation *Med. Image Anal.* **10** 752–63

Brock K K, Mutic S, Mcnutt T R, Li H and Kessler M L 2017 Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM radiation therapy committee task group no. 132 *J. Med. Phys.* **44** e43–e76

Broit C 1981 Optimal registration of deformed images *PhD Thesis* University of Pennsylvania

Christensen G E and Johnson H J 2001 Consistent image registration *IEEE Trans. Med. Imaging* **20** 568–82

De Silva T, Cool D W, Yuan J, Romagnoli C, Samarabandu J, Fenster A and Ward A D 2017 Robust 2-D–3-D registration optimization for motion compensation during 3-D TRUS-guided biopsy using learned prostate motion data *IEEE Trans. Med. Imaging* **36** 2010–20

de Vos B D, Berendsen F F, Viergever M A, Staring M and Isgum I 2017 End-to-end unsupervised deformable image registration with a convolutional neural network ed M Cardoso *et al Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA 2017, ML-CDS 2017 (Lecture Notes in Computer Science*, vol 10553) (Berlin: Springer) pp 204–12

Ferrant M, Nabavi A, Macq B, Jolesz F A, Kikinis R and Warfield S K 2001 Registration of 3-D intraoperative MR images of the brain using a finite-element biomechanical model *IEEE Trans. Med. Imaging* **20** 1384–97

Fleute M and Lavallee S 1999 Nonrigid 3-D/2-D registration of images using statistical models *Medical Imaging Computing-Assisted Intervention MICCAI'99 (Lecture Notes in Computer Science,* vol 1679) (Berlin: Springer) pp 138–47

Gibson E *et al* 2018 NiftyNet: a deep-learning platform for medical imaging *Comput. Method Prog. Biol.* **158** 113–22

Glorot X and Bengio Y 2010 Understanding the difficulty of training deep feedforward neural networks *Proc. of the 13th Int. Conf.on Artificial Intelligence and Statistics* ed Y W Teh and M Titterington (Chia Laguna Resort, Sardinia, Italy, 13–15 May) (http://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf) pp 249–56

Hu Y P, Ahmed H U, Taylor Z, Allen C, Emberton M, Hawkes D and Barratt D 2012 MR to ultrasound registration for image-guided prostate interventions *Med. Image Anal.* **16** 687–703

Hu Y P, Carter T J, Ahmed H U, Emberton M, Allen C, Hawkes D J and Barratt D C 2011 Modelling prostate motion for data fusion during image-guided interventions *IEEE Trans. Med. Imaging* **30** 1887–900

Hu Y P, Gibson E, Ghavami N, Bonati E, Moore C M, Emberton M, Vercauteren T, Noble J A and Barratt D C 2018 Adversarial deformation regularization for training image registration neural networks *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (Lecture Notes in Computer Science,* vol 11070) (Berlin: Springer) pp 774–82

Hu Y P, Modat M, Gibson E, Ghavami N, Bonmati E, Moore C M, Emberton M, Noble J A, Barratt D C and Vercauteren T 2018a Label-driven weakly-supervised learning for multimodal deformable image registration *IEEE Int. Symp. on Biomedical Imaging (Washington, DC, 4–7 April 2018)* (Piscataway, NJ: IEEE) pp 1070–4

Hu Y P *et al* 2018b Weakly-supervised convolutional neural networks for multimodal image registration *Med. Image Anal.* **49** 1–13

Huttenlocher D P, Klanderman G A and Rucklidge W J 1993 Comparing images using the Hausdorff distance *IEEE Trans. Pattern Anal. Mach. Intell.* **15** 850–63

Khallaghi S *et al* 2015 Statistical biomechanical surface registration: application to MR-TRUS fusion for prostate interventions *IEEE Trans. Med. Imaging* **34** 2535–49

Lei Y, Fu Y, Harms J, Wang T, Curran W J, Liu T, Higgins K and Yang X 2019a 4D-CT Deformable image registration using an unsupervised deep convolutional neural network. ed D Nguyen, L Xing, S Jiang *Artificial Intelligence in Radiation Therapy. AIRT 2019 (Lecture Notes in Computer Science*, vol 11850) (Berlin: Springer) pp 26–33

Lei Y, Tian S, He X, Wang T, Wang B, Patel P, Jani A B, Mao H, Curran W J and Liu T J 2019b Ultrasound prostate segmentation based on multidirectional deeply supervised V-net *Med. Phys.* **46** 3194–206

Litjens G, Toth R, van de Ven W, Hoeks C, Kerkstra S, van Ginneken B, Vincent G, Guillard G, Birbeck N and Zhang J J 2014 Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge *Med. Image Anal.* **18** 359–73

Mohamed A, Davatzikos C and Taylor R 2002 A combined statistical and biomechanical model for estimation of intra-operative prostate deformation *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (Lecture Notes in Computer Science,* vol 2489) (Berlin: Springer) pp 452–60

Paganelli C, Meschini G, Molinelli S, Riboldi M and Baroni G J 2018 Patient-specific validation of deformable image registration in radiation therapy: overview and caveats *Med. Phys.* **45** e908-e22

Rueckert D, Frangi A F and Schnabel J A 2003 Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration *IEEE Trans. Med. Imaging* **22** 1014–25

Sotiras A, Davatzikos C and Paragios N 2013 Deformable medical image registration: a survey *IEEE Trans. Med. Imaging* **32** 1153–90

Sun Y, Yuan J, Qiu W, Rajchl M, Romagnoli C and Fenster A 2015 Three-dimensional nonrigid MR-TRUS registration using dual optimization *IEEE Trans. Med. Imaging* **34** 1085–95

van de Ven W J M, Hu Y P, Barentsz J O, Karssemeijer N, Barratt D and Huisman H J 2015 Biomechanical modeling constrained surface-based image registration for prostate MR guided TRUS biopsy *Med. Phys.* **42** 2470–81

van de Ven W J M, Hulsbergen-van de Kaa C A, Hambrock T, Barentsz J O and Huisman H J 2013 Simulated required accuracy of image registration tools for targeting high-grade cancer components with prostate biopsies *Eur. Radiol.* **23** 1401–7

Wang B *et al* 2019 Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation *Med. Phys.* **46** 1707–18

Wang Y, Cheng J Z, Ni D, Lin M Q, Qin J, Luo X B, Xu M, Xie X Y and Heng P A 2016 Towards personalized statistical deformable model and hybrid point matching for robust MR-TRUS registration *IEEE Trans. Med. Imaging* **35** 589–604

Wojna Z, Ferrari V, Guadarrama S, Silberman N, Chen L-C, Fathi A and Uijlings J J 2019 The devil is in the decoder: classification, regression and GANs *Int. J. Comput. Vis.* **127** 1694–706

Yang X, Kwitt R, Styner M and Niethammer M 2017 Quicksilver: fast predictive image registration - A deep learning approach *Neuroimage* **158** 378–96
Zhou Z H 2018 A brief introduction to weakly supervised learning *Natl. Sci. Rev.* **5** 44–53

EMORY
WINSHIP
CANCER
INSTITUTE
A Cancer Center Designated by
the National Cancer Institute

Reliability of Doppler Blood Flow Evaluation of Neurovascular Bundle Vessels in Patients Receiving Prostate Radiotherapy

Xiuxiu He, Xiaofeng Yang, Ashesh Jani, James Sohn, Pretesh Patel, Walter Curran and Tian Liu

Department of Radiation Oncology
Winship Cancer Institute
Emory University

Purpose

Develop a **Doppler ultrasound imaging technology** to evaluate the blood flow of Neurovascular Bundles in patients receiving prostate cancer radiotherapy.

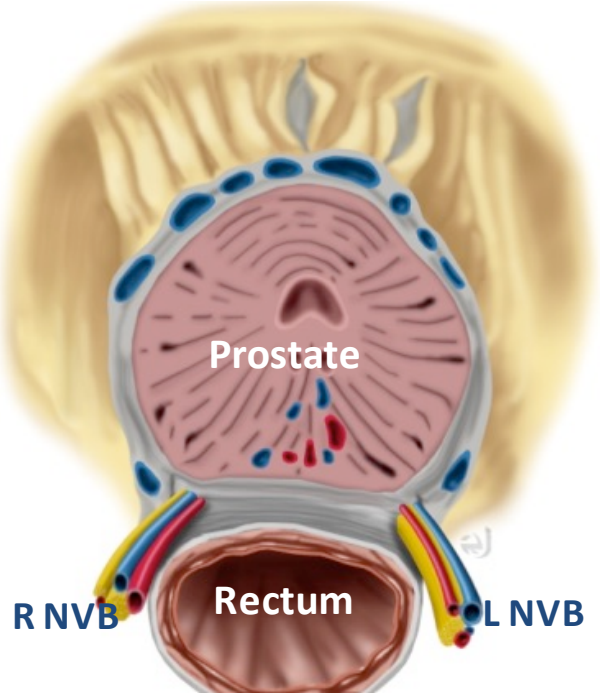
Background

- Erectile dysfunction (ED) is the most common complication of prostate radiotherapy.

- The most important cause is believed to be radiation damage to the neurovascular bundle (NVB), most commonly arteriogenic.

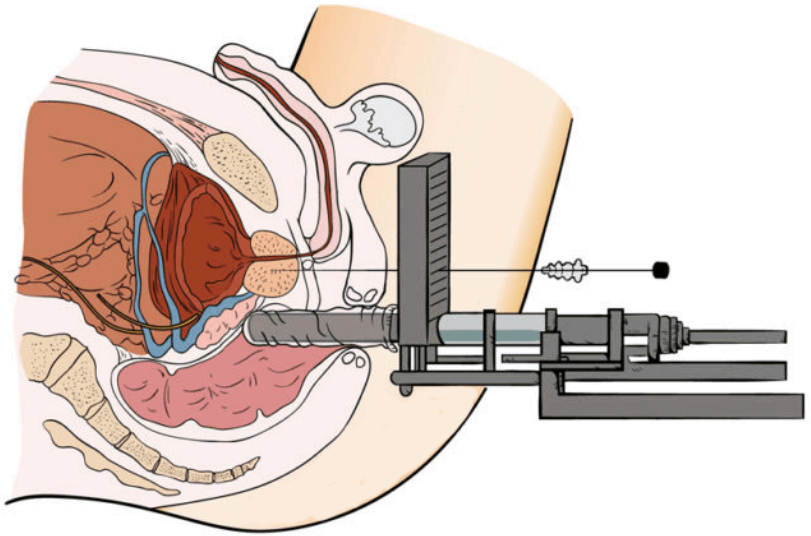
- Technique and data on the measurements of the NVB function are lacking.

Why Neurovascular Bundle?



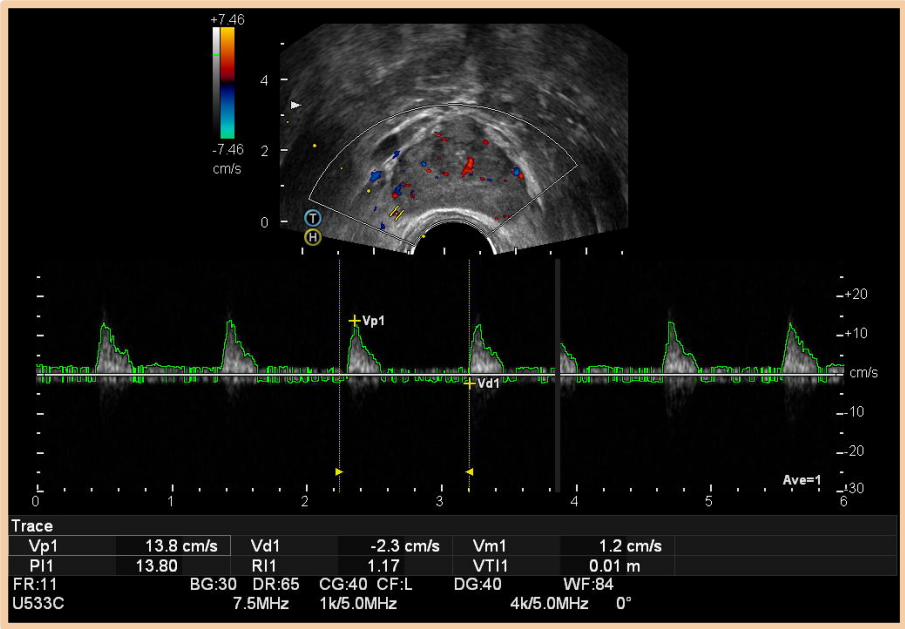
- Posterolateral aspect of prostate
- The structure most correlated with treatment-related ED
- Hypothesized to be the main cause of treatment-related ED
- Primary mechanism of action of major pharmacological agents

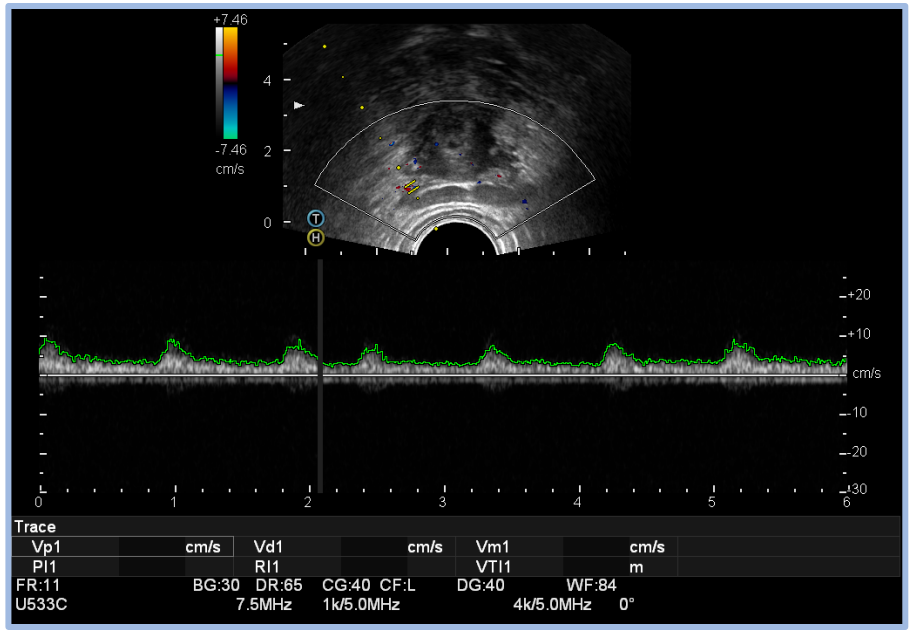
Image Acquisition



- 8 prostate cancer patients
  - HITACHI Avius
  - 7.5MHz transrectal probe
  - Lithotomy position
  - mechanical stepper
- 3D B-mode images
  - Parallel axial (transverse) scans are captured from base to the apex with 1 mm/2mm step size
- Doppler scans

Doppler Blood Flow of NVB & ED



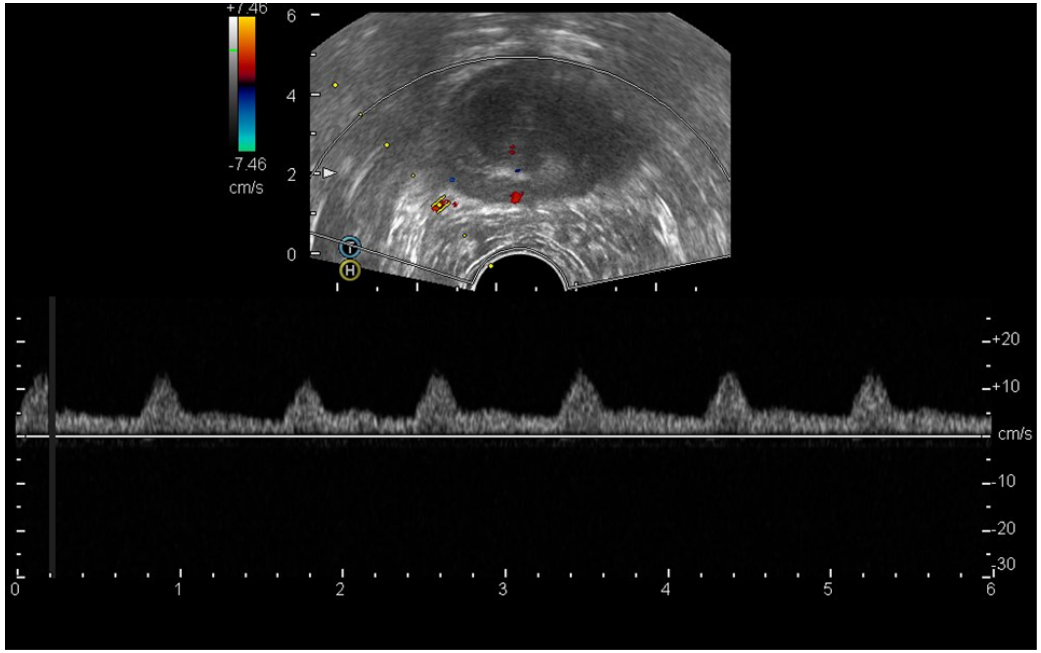


Normal Erectile Function	Erectile Dysfunction

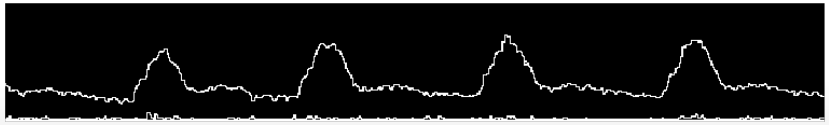
PSV (cm/s)	**13.80**	**9.31**
EDV (cm/s)	**2.25**	**3.37**
RI	**0.84**	**0.64**

International Index of Erectile Function (**IIEF**) Questionnaire

Pulse Waveform Extraction



Doppler ultrasound image

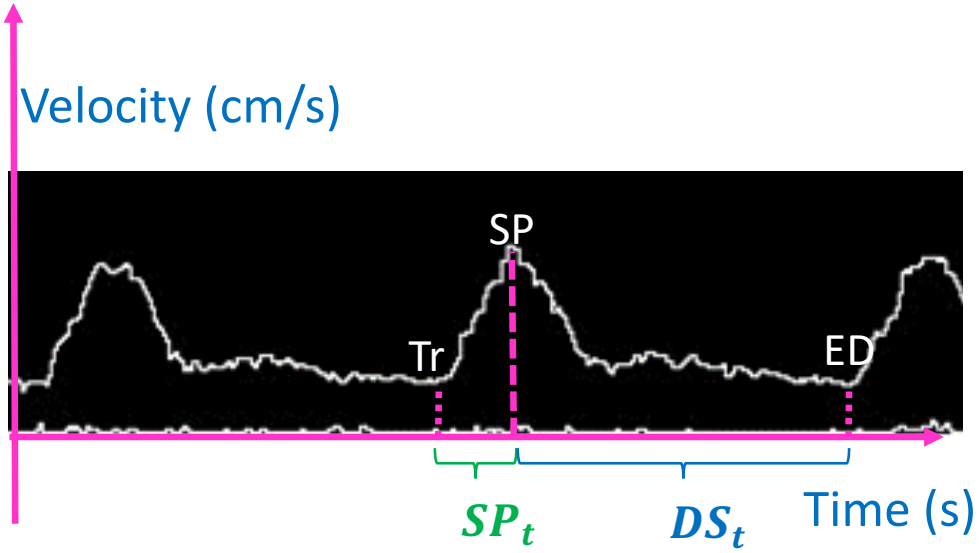


Pulse waveform

Waveform Morphological Feature Extraction

Definitions:

- Systolic peak: $PSV = SP_v$
- End of diastolic: $EDV = ED_v$
- Trough: Tr
- Mean velocity: $v_m = \frac{1}{T}\int_T vdt$
- Total time: time from Tr to ED
- Resistance index: $RI = \frac{PSV - EDV}{PSV}$
- Pulsatile index: $\mathrm{PI} = \frac{PSV - EDV}{v_m}$
- Upstroke time ratio: $RT_1 = \frac{SP_t}{T}$
- Downstroke time ratio: $RT_2 = \frac{DS_t}{T}$
- Time ratio between downstroke and upstroke time: $RT_3 = \frac{DS_t}{SP_t}$
- Scaled downstroke and upstroke time difference: $RT_4 = \frac{DS_t - SP_t}{T}$



Results

– Among the 8 patients

	Mean (μ)	Standard Deviation(σ)
PSV (cm/s)	10.11	4.45
EDV (cm/s)	1.39	0.95
Trv (cm/s)	1.35	0.95
Mean Velocity (cm/s)	4.09	1.47
Total Time (s)	0.88	0.13
Resistance Index	0.85	0.093
Pulsatile Index	2.16	0.78
RT1	0.17	0.067
RT2	0.82	0.057
RT3	0.21	0.10
RT4	0.67	0.127

– Comparison between the scans identified 4 reliable PW parameters: Total Time, Resistance Index, RT2, and RT4.

Conclusions

- We have developed a Doppler ultrasound imaging technique, integrated with image processing and pulse wave morphological features analysis to characterize blood flow of NVB.
- Our preliminary data have demonstrated its feasibility and identified reliable Doppler pulse wave features.
- We are acquiring more data and looking at the correlation between Doppler pulse wave measurements and Erectile Function in prostate cancer patients.

Acknowledgment

Emory Winship Cancer Institute Pilot Grant

W81XWH-17-1-0438
W81XWH-17-1-0439
W81XWH-13-1-0269

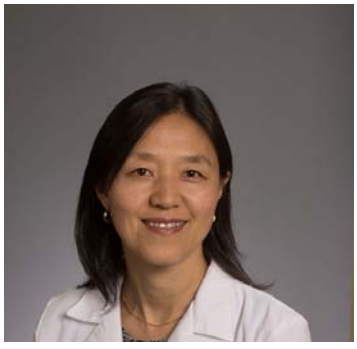
RO1CA215718

Prostate Cancer

Preserving Sexual Function in Men Undergoing Radiation Treatment for Prostate Cancer

Posted September 11, 2020

Tian Liu, Ph.D. and Ashesh Jani, M.D., Emory University



One of the most common complications experienced by men undergoing radiotherapy for prostate cancer is decreased sexual function. Radiation can damage blood vessels and nerves surrounding the prostate, resulting in erectile dysfunction over time and highlighting the need for new nerve-sparing therapies. Some researchers theorize that this side effect could be due to injury to the neurovascular bundles (NVB) surrounding the prostate, which are difficult to visualize with current imaging techniques. Because

Dr. Tian Liu

Dr. Ashesh Jani

sexual dysfunction can occur in up to 50% of men undergoing radiotherapy, there is a need for non-invasive imaging techniques that could spare the NVB, thereby increasing quality of life for prostate cancer patients.

With support from a FY16 Idea Development Award (Established Investigator – Partnering PI Option), Dr. Tian Liu and Dr. Ashesh Jani at Emory University are developing and testing a multimodality imaging platform to measure NVB injury caused by radiation. Dr. Liu's team is currently developing the technology, which will combine magnetic resonance imaging (MRI) and ultrasound, to measure radiation-related NVB injury (Figure 1). Dr. Jani has been leading a clinical study to validate results of the platform and correlate findings with standard clinical endpoints to evaluate erectile dysfunction. The team will also conduct a feasibility study to investigate the ability of this novel technology to improve sexual potency through NVB-sparing radiotherapy.

To date, the research team has created a unique prostate phantom with NVB components for pre-clinical testing and develop a method to analyze Doppler waveforms obtained from ultrasound imaging. This Doppler technology provided more accurate blood flow measurements surrounding the prostate. During radiotherapy treatment planning, clinicians use a process called segmentation to accurately determine prostate boundaries to localize radiotherapy and spare healthy tissues. To overcome the time-consuming process of manual segmentation of the prostate during transrectal ultrasound (TRUS), Co-Investigator, Dr. Xiaofeng Yang, and the team developed a multidirectional deep-learning method to automatically segment the prostate during TRUS-guided radiotherapy, which was published in *Medical Physics.*[1] In addition, the team developed a weakly supervised learning-based MRI-TRUS image registration which could quickly fuse MRI with TRUS to help the NVB localization.[2] A total of 28 subjects have been recruited for the clinical study, 18 (64%) of which are African American men to provide a more diverse set of patient outcomes. Additionally, an imaging and outcomes database has been developed to correlate the team's images with patient reported outcomes and clinical data, in hopes to further optimize and validate the platform in patients.

Results from this study could allow physicians to prevent NVB injury due to radiotherapy and potentially adjust radiation dosage during radiotherapy treatment planning. This work could also lead to new nerve-sparing therapies for prostate cancer patients that would preserve sexual function and increase quality of life.

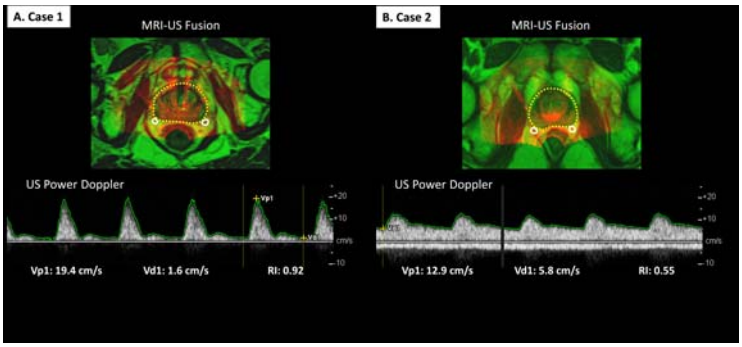


Figure 1. Multimodality MRI-US NVB imaging platform. Case 1: Power Doppler shows NVB blood flow of 59-year old patient with normal sexual function. Case 2: Lower NVB blood flow in 61-year old patient with erectile dysfunction.

References:

1. Lei Y, Tian S, He X, et al. 2019. Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net. Med Phys 46(6):3194-3206.
2. Zeng Q, Fu Y, Tian Z, et al. 2020. Label-driven magnetic resonance imaging (MRI) -transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy. Phys Med Biol. 65(13):135002.

Link:

Public and Technical Abstracts: Multimodality Imaging Platform for Neurovascular Bundle Sparing Prostate Radiotherapy to Preserve Sexual Function
(https://cdmrp.army.mil/search.aspx?LOG_NO=PC160820)

Top of Page

Last updated Friday, September 11, 2020

Note: Documents in Portable Document Format (PDF) require Adobe Acrobat Reader to view, download Adobe Acrobat Reader (http://get.adobe.com/reader/).

CDMRP © 2015

1077 Patchel Street
Fort Detrick, MD 21702-5024

(301) 619-7071

cdmrpwebmaster@webcdmrp.org (mailto:cdmrpwebmaster@webcdmrp.org)

About Us

The CDMRP originated in 1992 via a Congressional appropriation to foster novel approaches to biomedical research in response to the expressed needs of its stakeholders-the American public, the military, and Congress.

(https://www.facebook.com/TheCDMRP) (http://twitter.com/CDMRP)

(https://www.youtube.com/user/CDMRP) (/rss/funding_opportunities.xml)