Deformable Mri To Transrectal Ultrasound Registration For Prostate Interventions With Shape-Based Deep Variational Auto-Encoders

Conference Paper · April 2021

DOI: 10.1109/ISBI48211.2021.9434101

4 authors, including:

William Trung Le
University of Montreal Hospital Research Centre
13 PUBLICATIONS **38** CITATIONS

SEE PROFILE

Samuel Kadoury
Polytechnique Montréal
204 PUBLICATIONS **4,586** CITATIONS

SEE PROFILE

2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)
April 13-16, 2021, Nice, France

DEFORMABLE MRI TO TRANSRECTAL ULTRASOUND REGISTRATION FOR PROSTATE INTERVENTIONS WITH SHAPE-BASED DEEP VARIATIONAL AUTO-ENCODERS

Shirin Shakeri[⋆]*, William Le*[⋆†]*, Cynthia Ménard*[†]*, and Samuel Kadoury*[⋆†]

[⋆] MedICAL Laboratory, Polytechnique Montréal, Montréal, Canada
[†] Centre de recherche du CHUM (CRCHUM), Université de Montréal, Canada

2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) | 978-1-6654-1246-9/20/$31.00 ©2021 IEEE | DOI: 10.1109/ISBI48211.2021.9434101

ABSTRACT

Prostate cancer is one of the most prevalent cancers in men, where diagnosis is confirmed through biopsies analyzed with histopathology. A diagnostic T2-w MRI is often registered to intra-operative transrectal ultrasound (TRUS) for effective targeting of suspicious lesions during image-guided biopsy procedures or needle-based therapeutic interventions such as brachytherapy. However, this process remains challenging and time-consuming in an interventional environment. The present work proposes an automated 3D deformable MRI to TRUS registration pipeline that leverages both deep variational auto-encoders with a non-rigid iterative closest point registration approach. A convolutional FC-ResNet segmentation model is first trained from 3D TRUS images to extract prostate boundaries during the procedure. Matched MRI-TRUS 3D segmentations are then used to generate a vector representation of the gland's surface mesh between modalities, used as input to a 10-layer dense variational autoencoder model to constrain the predicted deformations based on a latent representation of the deformation modes. At each iteration of the registration process, the warped image is regularized using the autoencoder's reconstruction loss, ensuring plausible anatomical deformations. Based on a 5-fold cross-validation strategy with 45 patients undergoing HDR brachytherapy, the method yields a Dice score of 85.0 ± 2.6 with a target registration error of 3.9 ± 1.4mm, with the proposed method yielding results outperforming the state-of-the-art, with minimal intra-procedural disruptions.

***Index Terms*—** Prostate cancer, Deep variational auto-encoders, TRUS segmentation, Deformable registration, Non-iterative closest point alignment

1. INTRODUCTION

Prostate cancer (PCa) is the second most prevalent type of cancer in men and fourth overall, with an estimated 1.3 million new cases in 2019 according to the Global Cancer Observatory [1]. It is expected that 1 in 9 men will be diagnosed with PCa in their lifetime. Due to the small size of the prostate gland, a systematic biopsy is typically included in the diagnosis process to confirm the presence and grade of the lesion. Guidance tools are often used to assist clinicians in localizing the needles with respect to the tumor. Transrectal ultrasound (TRUS) is the most commonly used modality for real-time intra-operative imaging of the prostate for this purpose. However identification of tumor targets on TRUS remains challenging due to acoustic properties of the tumor relative to the surounding tissue. To address this issue, pre-treatment magnetic resonance imaging (MRI) is often performed to provide high quality detection of the tumor region. Still, the required image registration step between TRUS and MRI remains challenging. Intensity-based methods struggle under multi-modal conditions, whereas surface-based methods require real-time annotations of the gland on TRUS. Thus, an unsupervised segmentation framework during the TRUS-guided procedure followed by accurate multi-modal image registration is required.

Recently, deep neural network (DNN) approaches were shown to be promising for deformable registration procedures [2]. Compared to statistical deformation models (SDM), DNNs avoid making linear assumptions about the underlying distribution of non-linear deformations and are easily parallelizable for fast online predictions. This makes DNN techniques attractive for real-time intra-operative applications.

In this work, we present a registration framework composed of two steps: an automatic convolutional neural network (CNN) for TRUS prostate segmentation, followed by a deep variational autoencoder (DVAE) model used for registration based on a low-dimensional embedding of the segmented prostate surface variations. At registration time, both the MRI and the automatically segmented TRUS surface meshes are used within a non-rigid optimal Iterative Closest Point (NICP) framework, regularized by the TRUS embedding. The framework is evaluated on a clinical dataset of 45 patients with paired 3D MRI and TRUS images.

2. RELATED WORKS

Surface-based registration methods have presented state of the art results in the past decade for fusion-based prostate procedures. Natarajan et al. [3] used elastic matching to align the pre-biopsy 3D MRI with intra-biopsy TRUS volumes. This required fixed manual anatomical landmarks on both the base and apex of the prostate, with both volumes fused using a block matching approach with a 3D B-spline control grid. Hu et al. [4] employed a biomechanical finite element statistical motion model (FMM) trained on the MRI as a deformation constraint, while Onofrey et al. [5] trained a patient-specific statistical deformation model (SDM) on the MRI volume to obtain the set of possible prostate deformations. Khallaghi et al. [6] proposed a method based on the finite element model (FEM). While these works demonstrated reliable multi-modal alignment, they presented highly complex solutions for the intra-procedural workflow.

More recently, several works have explored incorporating deep learning methods for TRUS-MRI registration. Haskins et al. [7] trained a similarity metric using a CNN for an implicit optimization of the registration task. Zeng et al. [8] used a weekly supervised approach inspired by a viscous fluid physical model, combining a CNN and a long-short term memory neural network to predict the dense deformation fields. However these methods still relied on manual

annotations from TRUS images to obtain a target feature for intra-operative registration.

3. MATERIALS AND METHODS

3.1. Clinial Dataset

The clinical dataset used in this work consisted of 45 patients treated for PCa with high dose rate (HDR) brachytherapy. Each patient dataset contained a diagnostic T2-w MRI from a 1.5T scanner and an intra-operative TRUS volume (BK), obtained before catheter placement. MRI scans were T2-weighted 3D variable-flip-angle TSE images with 1mm isotropic voxels. TRUS scans consisted of 94 2D images obtained with a pullback technique, with 0.5mm slice thickness. Prostate segmentations were performed on both modalities by a radiation oncologist, yielding 3D surface meshes with vertex correspondences between 3006 nodes, using a spectral decomposition enabling vertex matching [9]. The gold standard MRI-TRUS registration was available for validation purposes: an expert clinician manually selected the 3-5 fiducial points per patient (cysts, calcifications and anatomical features) in order to perform a B-spline registration, and used as landmarks for internal validation.

3.2. Proposed Model

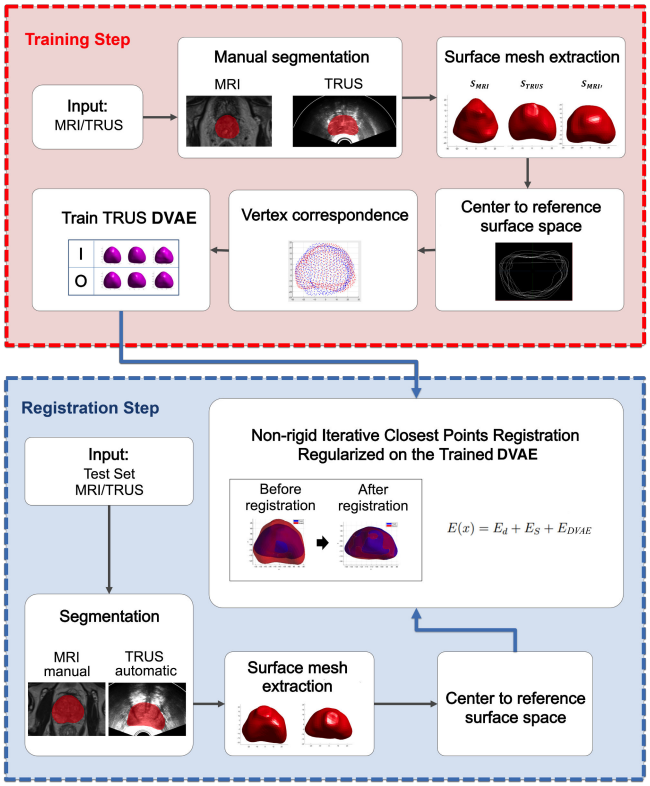

Fig. 1. **Top** Offline training phase of the deep variational auto-encoder using MRI-aligned TRUS prostate surface meshes, with S'_{MRI} as the deformed mesh. **Bottom** Online registration phase, performing segmentation using FC-ResNet trained on the 2D TRUS slices. At prediction time, the FC-ResNet predicted prostate meshes which are used to obtain the latent space embedding for used in the NICP algorithm, generating a deformed prostate surface.

The proposed automatic registration framework shown in Figure 1 consists of two steps: an offline phase, training the deep neural networks, and an online registration step based on NICP for deformable MRI to TRUS registration, constrained by the DVAE embedding in its loss function, and using the deep segmentation predictions of the FC-ResNet. Thus, two separate deep neural networks models are proposed: one for automatic segmentation of the prostate in intra-operative TRUS images, and one for learning the deep deformation constraints of the multi-modal registration step.

3.2.1. *TRUS segmentation of the prostate*

In clinical practice, manual annotations of the prostate on the live ultrasound is time consuming and tedious, adding significant time to the procedure. To address this issue, a FC-ResNet is proposed for automatic 2D segmentation of the organ. The architecture shown in Figure 2 is inspired by [10] in which a U-Net [11] is augmented with residual skip connections from ResNet [12] around each convolutional block. Figure 3 shows the block designs. These simple

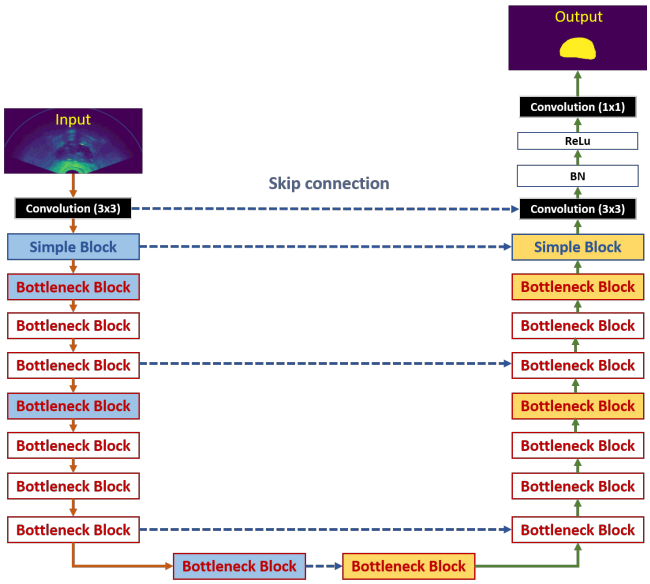

Fig. 2. FC-ResNet TRUS segmentation model.

blocks consist of a Batch Normalization layer, a ReLU activation layer and a 3×3 convolutional layer with optional max-pooling and up-sampling, in the encoder and decoder blocks respectively. A bottleneck block design from ResNet [10] is also employed in certain layers to increase network capacity. In these, each 3×3 convolutional layer is replaced with a 1×1, 3×3, then 1×1 convolutional layer. A 20% dropout layer is included at the end of every block to prevent overfitting. The model also features 4 long skip connections before each downsampling step, helping to preserve fine grained edge details in the predicted prostate mask.

3.2.2. *Deformable MRI-TRUS registration with deep VAE*

To learn the distribution of plausible non-linear prostate deformations, a DVAE is trained on the MRI-TRUS dataset. The input data consists of manually segmented TRUS prostates from which the triangular surface meshes are extracted using the marching cube algorithm, which are then centered to a common MRI space. Figure 4

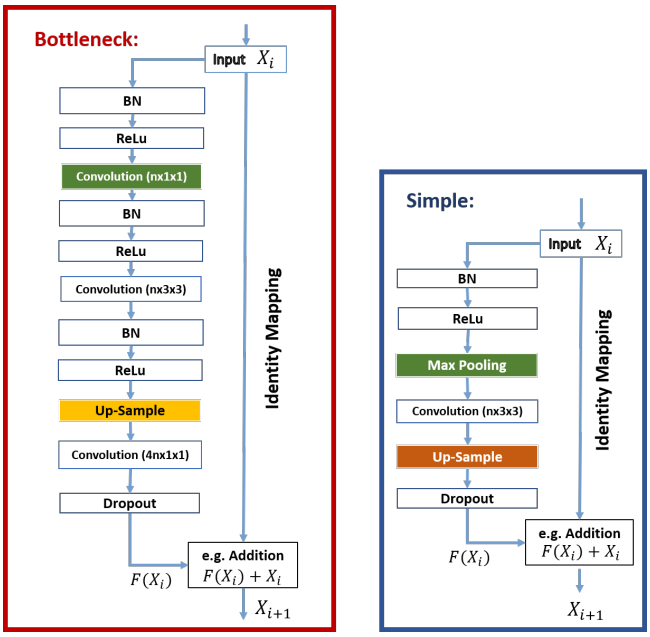

Fig. 3. Convolutional blocks used in the FC-ResNet segmentation network.

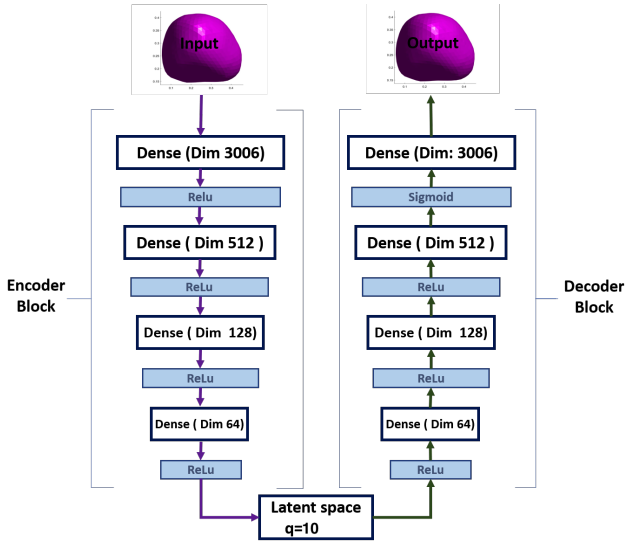

Fig. 4. Deep variational autoencoder model to learn modes of prostate deformation.

presents the feed-forward architecture of the DVAE: the encoder and decoder are 4 symmetrically distributed dense layers with ReLU activations. A Sigmoid activation is used instead of the final ReLU, thus allowing to learn the non-linear deformations of the prostate. The training objective seeks to minimize the surface mesh reconstruction error. The compressed form of the 3006-vector input to the 10-vector embedding serves to encode a latent representation of plausible TRUS prostate deformations.

At registration time, the manually contoured MRI and the automatically segmented TRUS (S_{MRI} and S_{TRUS}), predicted by the FC-ResNet, are converted to surface meshes and mappped onto a common surface space. The NICP algorithm then iteratively warps S_{MRI} to S_{TRUS}, by minimizing the following cost function using an alternating approach [13]:

$$E(X) = E_d + E_S + E_{DVAE} \tag{1}$$

with E_d as a sum of squared distances between the source MRI and the target TRUS vertices, defined as:

$$E_d(X) := \sum_{v_i \in V} dist^2(T, X_i v_i), \tag{2}$$

where $X = \{x_1, x_2, x_3, ...x_n\}$ is a 3x4 affine transformation matrices for each vertex in set V. Furthermore, $v_i = [x, y, z, 1]^T$ is the template vertices and $dist(T, v)$ represents the distance between vertex v and the nearest vertex on S_{TRUS}.

The stiffness term (E_s) constrains the function according to the weighted distance of adjacent vertices (X_i and X_j as the adjacent points) as in Eq. (3). Also, $G = diag(1, 1, 1, \gamma)$ is the weight variable. The stiffness term is also penalized under the Frobenius norm:

$$E_s(X) := \sum_{i,j \sim \varepsilon} \|(X_i - X_j)G\|_2^F \tag{3}$$

with E_{DVAE} as a constraint on the deformation which fits in the DVAE (Eq. (4)). Specifically, at each iteration, one candidate deformation is computed to produce an intermediate TRUS surface S. This is fed to the encoder to produce the resulting embedding q (i.e. $encoder(S) = q$). Here, the parameter q (representing the compact representation of the deformation field) can be optimized such that the $decoder(q)$ is close to S based on the DVAE loss function:

$$E_{DVAE} = \|decoder(q)_i - S_i\|_2 \tag{4}$$

where, i refers to the vertex index. As suggested by Umetani et al.[14], in order to minimize the cost function Eq.(4), each element of q is projected between 0 and 1. This constraint on q works as a regularizer on each candidate deformation in our iterative registration process.

4. RESULTS AND DISCUSSION

The FC-ResNet segmentation model was evaluated using a 5-fold cross-validation on the TRUS images. The samples were split 20%, 30% and 50% for the test, validation and training sets, respectively, and underwent isometric resampling, normalization, as well as data augmentation: horizontal and vertical flipping, sheering, rotation, and warping. The Adam optimizer was used with a learning rate of 1×10^{-8}. Figure 5 shows sample segmentation results on TRUS images. The visual contours accurately match the gold-standard manual annotations. The Dice overlap score was used to assess the TRUS segmentation model performance, with results obtained subdivided per region. The base of the prostate yielded a Dice score of $84.9 \pm 2.3\%$. The best performance was in the mid-gland (89.3%) and the worst at the apex (72.2%). For the overall organ, the segmentation model achieved a Dice score of 82.1%.

Figure 6 presents registration results for a sample patient from the test set. The average registration time for all cases was.3s. The observed deformation fields interpolated from the NICP dense deformation field show that the registration process manages to accurately deform the edges of the prostate, consistent with the type of deformations due to pressure on the organ from the transrectal ultrasound probe.

Table 1. Comparison of the proposed SDM-NICP and DVAE-NICP approaches, using the automatic FC-ResNet segmentation on unseen TRUS image. Evaluation is based on target registration errors (TRE), mean squared distance (MSD), and Dice score. The results are reported in each sub-region separately.

	SDM-NICP			DVAE-NICP		
	Apex	Mid	Base	Apex	Mid	Base
MSD (mm)	0.75 $\pm$0.26	0.30 $\pm$0.11	0.32 $\pm$0.27	0.65 $\pm$0.26	0.19 $\pm$0.13	0.27 $\pm$0.27
TRE (mm)	6.42 $\pm$1.82	4.63 $\pm$1.73	5.73 $\pm$1.89	4.52 $\pm$1.60	3.57 $\pm$1.48	3.86 $\pm$1.12
Dice	72.5 $\pm$ 3.5	83.2 $\pm$ 2.7	80.1 $\pm$ 2.9	82.4 $\pm$ 2.8	88.8 $\pm$ 2.1	85.04 $\pm$ 3.0

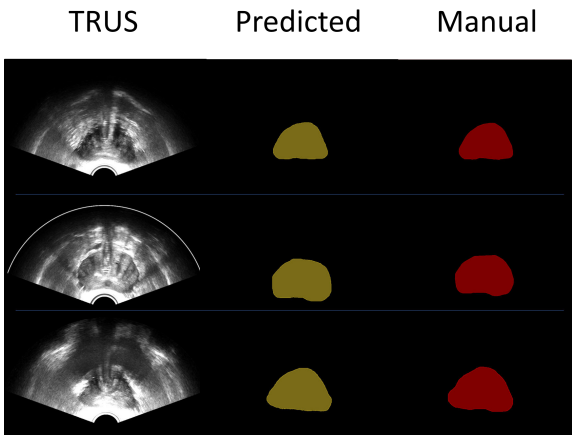

Fig. 5. Sample segmentation results on the test set for various regions of the prostate on TRUS compared with the ground-truth annotation.

Table 1 presents the quantitative results of the registration experiments. A baseline method is used as comparison based on the works of Onofrey et al. [5], in which a statistical deformation model (SDM) combined with NICP is used (SDM-NICP). The proposed DVAE-constrained NICP method demonstrates both the lowest overall MSD (0.37 ± 0.22mm) and TRE (3.9 ± 1.4mm) measures for the prostate, as well as the highest whole organ Dice overlap ($85.0 \pm 2.6\%$). These results show a significant improvement ($p < 0.05$) to overall scores from the SDM-NICP method. In the apex region, the DVAE-constrained NICP shows both lower MSD (0.65 ± 0.26mm) and TRE (4.52 ± 1.60mm), while the proposed model yields the higher Dice (82.4%). In the mid-gland, the best performance was obtained using the DVAE-constrained NICP method, with a MSD of 0.19mm, a TRE of 3.57mm and a Dice score of 88.8%.

5. CONCLUSION

In this work, we proposed a framework for deformable MRI to TRUS image registration of the prostate using shape-based deep variational auto-encoders. A deep VAE model was trained to learn the distribution of plausible non-linear deformations observed before and during therapy due to probe pressure, used to constrain the NICP registration algorithm. To meet the requirements of rapid contouring in an intra-operative setting, an FC-ResNet CNN was trained to automatically segment the prostate in TRUS. The proposed model showed improvements in Dice score over previous state-of-the-art methods, improvements in overall MSD and TRE measures in the base, mid-gland and the apex region. This framework could therefore provide a tool for online registration during tracked TRUS procedures, without needing manual prostate annotations on TRUS, saving valuable time in the operating room.

6. COMPLIANCE WITH ETHICAL STANDARDS

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the local Institutional Review Board.

7. ACKNOWLEDGMENTS

This research has been funded in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). There are no conflicts of interests.

8. REFERENCES

[1] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L Siegel, Lindsey A Torre, and Ahmedin Jemal, "Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] Grant Haskins, Uwe Kruger, and Pingkun Yan, "Deep learning in medical image registration: a survey," *Machine Vision and Applications*, vol. 31, no. 1, pp. 1–18, 2020.

[3] Shyam Natarajan, Leonard S Marks, Daniel JA Margolis, Jiaoti Huang, Maria Luz Macairan, Patricia Lieu, and Aaron Fenster, "Clinical application of a 3d ultrasound-guided prostate biopsy system," in *Urologic oncology: seminars and original investigations*. Elsevier, 2011, vol. 29, pp. 334–342.

[4] Yipeng Hu, Hashim Uddin Ahmed, Zeike Taylor, Clare Allen, Mark Emberton, David Hawkes, and Dean Barratt, "Mr to ultrasound registration for image-guided prostate interventions," *Medical image analysis*, vol. 16, no. 3, pp. 687–703, 2012.

[5] John A Onofrey, Lawrence H Staib, Saradwata Sarkar, Rajesh Venkataraman, Cayce B Nawaf, Preston C Sprenkle, and Xenophon Papademetris, "Learning non-rigid deformations for robust, constrained point-based registration in image-guided mr-trus prostate intervention," *Medical image analysis*, vol. 39, pp. 29–43, 2017.

[6] Siavash Khallaghi, C Antonio Sánchez, Abtin Rasoulian, Yue Sun, Farhad Imani, Amir Khojaste, Orcun Goksel, Cesare Romagnoli, Hamidreza Abdi, Silvia Chang, et al., "Biomechanically constrained surface registration: Application to mr-trus fusion for prostate interventions," *IEEE transactions on medical imaging*, vol. 34, no. 11, pp. 2404–2414, 2015.

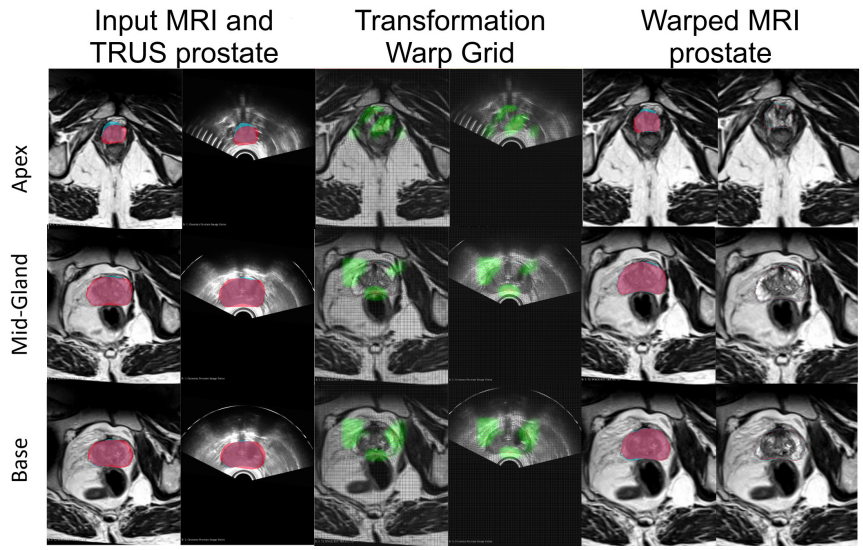

Fig. 6. Sample MRI-TRUS registration result from a 65yr-old male, using 3D segmented surfaces used as inputs to the model, and a B-spline grid (interpolated from the NICP deformation field) to warp the MRI onto the targeted TRUS. The green areas represent the registered prostate surface on the TRUS image and the red areas represent the prostate gland on the MRI.

[7] Grant Haskins, Jochen Kruecker, Uwe Kruger, Sheng Xu, Peter A Pinto, Brad J Wood, and Pingkun Yan, "Learning deep similarity metric for 3d mr–trus image registration," *International journal of computer assisted radiology and surgery*, vol. 14, no. 3, pp. 417–425, 2019.

[8] Qiulan Zeng, Yabo Fu, Jiwoong Jeong, Tian Zhen, Tonghe Wang, Yang Lei, Hui Mao, Ashesh B Jani, Pretesh Patel, Walter J Curran, et al., "Weekly supervised convolutional long short-term memory neural networks for mr-trus registration," in *Medical Imaging 2020: Ultrasonic Imaging and Tomography*. International Society for Optics and Photonics, 2020, vol. 11319, p. 1131910.

[9] Mahsa Shakeri, Herve Lombaert, Shashank Tripathi, Samuel Kadoury, Alzheimer's Disease Neuroimaging Initiative, et al., "Deep spectral-based shape features for alzheimer's disease classification," in *International Workshop on Spectral and Shape Analysis in Medical Imaging*. Springer, 2016, pp. 15–24.

[10] Michal Drozdzal, Gabriel Chartrand, Eugene Vorontsov, Mahsa Shakeri, Lisa Di Jorio, An Tang, Adriana Romero, Yoshua Bengio, Chris Pal, and Samuel Kadoury, "Learning normalized inputs for iterative estimation in medical image segmentation," *Medical image analysis*, vol. 44, pp. 1–13, 2018.

[11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[13] Shiyang Cheng, Ioannis Marras, Stefanos Zafeiriou, and Maja Pantic, "Active nonrigid icp algorithm," in *2015 11th IEEE international conference and workshops on automatic face and gesture recognition (FG)*. IEEE, 2015, vol. 1, pp. 1–8.

[14] Nobuyuki Umetani, "Exploring generative 3d shapes using autoencoder networks," in *SIGGRAPH Asia 2017 Technical Briefs*, New York, NY, USA, 2017, SA 17, Association for Computing Machinery.