HHS Public Access

Author manuscript

Acad Radiol. Author manuscript; available in PMC 2023 September 01.

Published in final edited form as:

Acad Radiol. 2022 September ; 29(9): 1404–1412. doi:10.1016/j.acra.2022.01.009.

Development of a 3D CNN-based AI Model for Automated Segmentation of the Prostatic Urethra

Mason J. Belue, B.S.[1], **Stephanie A. Harmon, Ph.D.**[1], **Krishnan Patel, M.D.**[2], **Asha Daryanani, B.S.**[1], **Enis Cagatay Yilmaz, M.D.**[1], **Peter A. Pinto, M.D.**[3], **Bradford J. Wood, M.D.**[4,5], **Deborah E. Citrin, M.D.**[2], **Peter L. Choyke, M.D.**[1], **Baris Turkbey, M.D.**[1,*]

[1]Molecular Imaging Branch, NCI, NIH, Bethesda, MD

[2]Radiation Oncology Branch, NCI, NIH, Bethesda, MD

[3]Urologic Oncology Branch, NCI, NIH, Bethesda, MD

[4]Center for Interventional Oncology, NCI, NIH, Bethesda, MD

[5]Depart of Radiology, Clinical Center, NIH, Bethesda, MD

Abstract

Rationale-objective: The combined use of prostate cancer radiotherapy and MRI planning is increasingly being used in the treatment of clinically significant prostate cancers. The radiotherapy dosage quantity is limited by toxicity in organs with de-novo genitourinary toxicity occurrence remaining unperturbed. Estimation of the urethral radiation dose via anatomical contouring may improve our understanding of genitourinary toxicity and its related symptoms. Yet, urethral delineation remains an expert-dependent and time-consuming procedure. In this study, we aim to develop a fully automated segmentation tool for the prostatic urethra.

Materials and methods: This study incorporated 939 patients' T2-weighted MRI scans (train/validation/test/excluded: 657/141/140/1 patients), including in-house and public PROSTATE-x datasets, and their corresponding ground truth urethral contours from an expert genitourinary radiologist. The AI model was developed using MONAI framework and was based on a 3D-UNet. AI model performance was determined by Dice score (volume-based) and the Centerline Distance (CLD) between the prediction and ground truth centers (slice-based). All predictions were compared to ground truth in a systematic failure analysis to elucidate the model's strengths and weaknesses. The Wilcoxon-rank sum test was used for pair-wise comparison of group differences.

Results: The overall organ-adjusted Dice score for this model was 0.61 and overall CLD was 2.56mm. When comparing prostates with symmetrical (n=117) and asymmetrical (n=23) benign prostate hyperplasia (BPH), the AI model performed better on symmetrical prostates compared to

[*]**Corresponding author:** Baris Turkbey M.D., Molecular Imaging Branch, National Cancer Institute, 10 Center Dr., MSC 1182, Building 10, Room B3B85, Bethesda, MD, 20892-1088, turkbeyi@mail.nih.gov.

Declearation of Interest

N/A

asymmetrical in both Dice score (0.64 vs. 0.51 respectively, p<0.05) and mean CLD (2.3 mm vs 3.8 mm respectively, p<0.05). When calculating location-specific performance, the performance was highest at the apex and lowest at the base location of the prostate for Dice and CLD. Dice location dependence: symmetrical (Apex, Mid, Base: 0.69 vs 0.67 vs 0.54 respectively, p<0.05) and asymmetrical (Apex, Mid, Base: 0.68 vs 0.52 vs 0.39 respectively, p<0.05). CLD location dependence: symmetrical (Apex, Mid, Base: 1.43 mm vs 2.15 mm vs 3.28 mm, p<0.05) and asymmetrical (Apex, Mid, Base: 1.83 mm vs 3.1 mm vs 6.24 mm, p<0.05).

Conclusion: We developed a fully automated prostatic urethra segmentation AI tool yielding its best performance in prostate glands with symmetric BPH features. This system can potentially be used to assist treatment planning in patients who can undergo whole gland radiation therapy or ablative focal therapy.

Keywords

prostate; urethra; radiation therapy; focal therapy; planning; MRI: magnetic resonance imaging; CLD: Centerline Distance; BPH: Benign prostate hyperplasia; SBRT: stereotactic body radiation therapy

INTRODUCTION

Prostate cancer is the second leading cause of cancer related deaths among American males (1). The treatment of prostate cancer usually involves prostatectomy, radiotherapy, focal therapy or management with active surveillance (2). The choice of treatment can be highly individualized and depends on a combination of clinical factors including the side effects of each therapy, cancer risk factors and preference of the patient. For radiotherapy planning of prostate cancer, prostate MRI is being increasingly used for delineation of the prostate gland, intraprostatic cancer foci and critical structures such as the urethra, bladder, rectum, and neurovascular bundle. Delineation of these areas is also important for focal ablative treatments as well as for focal dose escalation/de-escalation during radiotherapy (3). In prostate cancer radiotherapy, dose is largely limited by toxicity in organs or structures at risk. Although gastrointestinal toxicity has been reduced relative to other organs by the introduction of new radiotherapy techniques, genitourinary (GU) toxicities have remained relatively the same, with a 5-year rate of Grade >= 2 toxicity of approximately 15 to 20%. Additional retrospective analyses evaluating the predictors of GU toxicity for novel radiotherapies such as stereotactic body radiation therapy (SBRT) have suggested a dosimetric relationship between urethral dose and toxicity (4,5). This relationship appears to be generalizable to other techniques of irradiation with different radiobiologic profiles such as LDR (6) and HDR brachytherapy (7). The urethra is a particularly vulnerable structure within the prostate because it lies centrally. Estimating the delivered dose to the urethra, which depends on accurate urethral delineation, may improve our understanding of urinary toxicity, and enable new methods to reduce radiotherapy toxicity (8). However, accurate delineation of the urethra on MRI requires expertise and is time consuming which mitigate against its routine use.

In a limited number of prior studies, CT-based AI models were developed to delineate foley - catheterized urethras (8,9). These studies have limitations in that their ground truth urethral

contours were obtained with a urinary catheter in place. Moreover, these previous studies have been performed on relatively small samples of 55–120 patients. A recently published review covering AI models for organs at risk also reported that the only other MRI-based AI model made for urethra segmentation also had a small sample size of 50 patients (10,11). In this study, our goal is to build off these prior works and develop an end-to-end urethra segmentation model based on non-catheterized prostate MRI.

METHODS

Study Population

For this dataset, patients were scanned between July 2011 and August 2021 for clinical suspicion or follow-up of prostate cancer. This cohort includes patients recruited as part of one or more IRB approved protocols, including radiologic profiling of prostate cancer (ClinicalTrials.gov Identifier: NCT03354416), patients undergoing MRI-fusion-guided prostate biopsy (NCT00102544), and patients undergoing surgical treatment for intermediate or high-risk prostate cancer (NCT02594202). Patients in the in-house institutional cohort underwent were performed using a 3 Tesla magnet (Achieva 3.0-T-TX or Ingenia Elition X, Philips Healthcare, Best, the Netherlands) with a 16-channel surface coil (SENSE, Philips Healthcare, Best, the Netherlands) and an endorectal coil (BPX-30, MEDRAD, Pittsburgh, Pennsylvania) in 402 patients. For patients where an endorectal coil was not able to be utilized, a 32-channel cardiac coil (SENSE, InVivo, Gainesville, Florida) was used instead (n= 447). The endorectal coil (BPX-30, MEDRAD, Pittsburgh, Pennsylvania) was inflated with perfluorocarbon (3mol/l Fluorinert) to a volume of 45 ml, therefore no additional susceptibility artifacts are observed with its use. In both acquisitions, T2-weighted turbo-spin-echo acquisition sequences were obtained in the axial, sagittal, and coronal planes. Patient scans were excluded if T2W MRI quality was non-diagnostic (n=1). All T2W imaging acquisitions were oblique with respect to the longitudinal axis of the prostate. Further details on the acquisition parameters are provided in supplementary table.

An additional data source from the publicly available ProstateX challenge set was utilized. Briefly, the cohort is from 344 patients who underwent T2-weighted, proton density-weighted, dynamic contrast enhanced and diffusion weighted imaging using the 3 Tesla MAGNETOM Trio and Skyra scanner systems (Siemens Healthineers, Erlangen, Germany) without the use of an endorectal coil (12,13). T2-weighted images were obtained using a turbo spin echo sequence. The in-house institutional database (n=844) plus a portion of the public PROSTATEx dataset (n=95), combined to an overall sample size of 939 patients.

MRI Data Preprocessing

Prostatic urethra was contoured manually by one dedicated genitourinary radiologist (cumulative experience of 14 years in reading prostate MRIs, 1000 MRI reads/year) on axial T2W MRI data using an in-house annotation tool (pseg, iCAD) (14). Only the prostatic urethra is segmented in each patient scan, excluding the membranous urethra.

Data was split into training and validation/test groups using the train_test_split function from pythons sklearn module (15). The partitions were 70% training (n=657) and the

remaining 30% was evenly split into validation/testing (n=282) using a random seed of 42. Partitions are disjoint at the patient level, incorporating and randomizing all sources of scans (Institutional and External cases) between training, validation, and testing groups. Every MRI underwent the following preprocessing pipeline: field of view normalization, intensity normalization, intensity scaling from 0–10, image resizing to (256,256,32) by either padding or center cropping which creates data subsets from raw images. Patient data was de-identified by exporting only the image voxel information from patient DICOMs into 3D Nifti files. All model inference/evaluation was performed using these generated de-identified Nifti files. Missing or corrupted data was excluded from the final analysis.

AI Model Development

The proposed prostate urethra segmentation AI model was based on 3D UNet architecture which had residual units implemented with the ResidualUnit class (16,17). Each layer of the network had an encode and decode path with a skip connection between them. This model had down-/up-sampling channels of (16, 32, 64, 128, 256), strides of (2, 2, 2, 2), 4 residual units, and dropout rate of 0.5 (Figure 1). The loss function used was the Dice loss which also excluded the background from calculation. The Adam optimizer was used with a learning rate of 3e-4. The proposed AI was developed using the MONAI framework, which is an open source, freely available, PyTorch-based framework for deep learning in healthcare imaging (18). To speed up testing different hyperparameters, the weights of the previously best trained model were used to initialize training of the new model. This model would take in raw MRI data, perform all preprocessing, and output the predicted urethra segmentation, making this an end-to-end product. Additionally, the following augmentations were applied to the training data: Adding random gaussian noise and performing random image rotations based on the affine matrix. The final model was selected based on evaluating the model outcome metrics such as Dice score on the validation data. The final model had the highest overall/bulk Dice score on the validation data.

AI Performance Evaluation and Statistical Analysis

The metrics of model performance that were gathered included the overall Dice score, organ-adjusted Dice score, organ-adjusted and length-partitioned Dice score, organ-adjusted Centerline Distance (CLD), and organ-adjusted and length-partitioned CLD (figures 2a and 2b).

- **CLD** is defined as the average of the distances between the ground truth centerline and the centerline from the segmentation. The center of mass of each ground truth and prediction slice was found using the center_of_mass function from pythons SciPy module (19). The outputs of this function are the coordinate pairs of the center for each input image. The difference between these two coordinate pairs is computed and is then converted from voxels to millimeters using the defined spacing extracted from each scan.
- **Dice:** The “Overall” metric represents the Dice score on the validation/test cases without any form of post processing which includes predictions of both the prostatic and membranous urethra.

- **Organ-adjusted Dice** refers to a corrected form of the Dice score that doesn't penalize for cases where the prediction extends past the ground truth (ex. AI predicting both prostatic and membranous urethra). This correction only uses prediction slices where there is a corresponding ground truth (within prostate) and excludes the prediction slices where the corresponding ground truth label is absent (outside of prostate).
- **Length partitioning** refers to taking the organ-adjusted Dice score or CLD and stratifying the prostatic urethra into three separate groups corresponding to the upper third (base), mid third, and bottom third (apex).

The Wilcoxon rank sum test was used to determine statistical significance between groups. The output of the model is a 3D segmentation file that is registered to the T2-weighted image allowing for easy visualization of the predicted urethra path.

Failure Analysis

A failure analysis was incorporated to document and elucidate the AI model strengths and weaknesses by looking into the test cases and comparing ground truth to predictions. Examples of information gathered included if the predictions extended past the ground truth, relative size of predictions compared to ground truth, and presence of an asymmetrical prostate or not (assessed blindly by radiologist). An asymmetric prostate was defined as one containing BPH nodule(s) crossing midline of the transition zone and significantly altering the course of the urethra. Failure analysis revealed significant differences in performance on asymmetrical vs symmetrical prostates (figures 3a and 3b). Additionally, the failure analysis led to incorporating an organ-adjusted correction to correct for penalizing correct segmentation of membranous urethra by the AI.

RESULTS

Basic demographics of the study population within each model partition are presented in Table 1. The number of asymmetrical prostates was not assessed for training data but was assessed for the validation and testing data and found to be 20/141 and 23/140, respectively. Any comparisons of symmetric and asymmetric groups constitute the failure analysis are described in the methods.

Evaluation using Dice Coefficient

The non-organ-adjusted Dice score in the overall population was 0.59. The Wilcoxon rank sum test, shown in Figure 4, revealed a significant difference in Dice score between symmetrical and asymmetrical prostates (0.61 vs. 0.52, $p<0.05$).

Likewise, the Dice score for the organ-adjusted urethra was 0.61 in the overall population, and similarly, there was a significant difference in Dice score between asymmetrical prostates and asymmetrical (0.64 vs 0.51, $p<0.05$) as seen in Figure 5.

The organ-adjusted Dice score showed statistically significant differences between all pair-wise comparison for both symmetrical (Apex, Mid, Base: 0.69 vs 0.67 vs 0.54 respectively, $p<0.05$) and asymmetrical (Apex, Mid, Base: 0.68 vs 0.52 vs 0.39 respectively, $p<0.05$)

prostates. Independent of symmetry, the algorithm predicted the location of the apical one-third of the urethra with greatest agreement and the base one-third with the lowest agreement. Box plots showing the distribution of these data are shown in Figure 6.

The performance of the algorithm showed a trend towards greater differences towards the base of the prostate between the overall population and symmetric and asymmetric subpopulations as shown in Figure 7. At the apex, there was no significant difference between each group by pair-wise comparison (Apex, Mid, Base: 0.69, 0.68, 0.69, $p>0.05$). At the mid-gland, the pair-wise comparison between these three groups revealed significant differences between the asymmetric and the symmetric subpopulations (0.52 vs. 0.67, $p < 0.05$) and the asymmetric and overall population (0.52 vs 0.65, $p<0.05$) but no significant difference between the symmetric and overall population (0.67 vs 0.65 respectively, $p>0.05$). Finally, there was a similar pattern of differences observed at the base of the prostate with significant differences seen between the asymmetric subpopulation and the symmetric subpopulation (0.29 vs. 0.54, $p < 0.05$) and the asymmetric and overall population (0.29 vs. 0.50, $p < 0.05$). Again, no significant difference in the organ-adjusted Dice score was observed between the symmetric subpopulation and overall population (0.54 vs. 0.50, $p > 0.05$).

Evaluation using Centerline Distance

The AI algorithm was found to have mean performance dependent on craniocaudal sublocation within the gland as measured by the CLD metric as shown in Figure 8. Statistical testing revealed a significantly higher mean CLD in the asymmetric subpopulation as compared to the symmetric subpopulation (3.8mm vs. 2.3mm, $p < 0.05$) and overall population (3.8mm vs 2.6mm, $p < 0.05$). There was no significant difference between the mean CLD of the symmetric subpopulation and overall population (2.3mm vs. 2.6mm).

Within each subpopulation, the mean performance of the AI model was also dependent on craniocaudal sublocation as measured by CLD and displayed in Figure 9. For the symmetric prostate subpopulation, the mean CLD at the apex, mid-gland, and base were 1.43mm, 2.15mm, and 3.28mm respectively. For the asymmetric subpopulation, the mean CLD were 1.83mm, 3.1mm, and 6.24mm, respectively. In both subpopulations, a pair-wise comparison between the three sublocation showed that these differences were statistically significant.

The CLD of each subcomponent of the urethra (apex, mid-gland, base) was also compared between the symmetric subpopulation, asymmetric subpopulation, and overall population as shown in Figure 10. For the apical urethra, a pair-wise comparison showed no significant differences between the symmetric, asymmetric, and overall subpopulation ($p>0.05$). At the mid-gland, significant differences were seen between the mean CLD of the asymmetric subpopulation when compared to the symmetric subpopulation (3.11mm vs. 2.15mm, $p < 0.05$) and the overall population (3.11mm vs. 2.31mm, $p < 0.05$), but not between the symmetric subpopulation and overall population (2.15mm vs. 2.31mm, $p > 0.05$). Similarly, at the base, significant differences were seen in the mean CLD of the asymmetric subpopulation when compared to the symmetric subpopulation (6.24 mm vs. 3.28mm, $p < 0.05$) and the overall population (6.24mm vs. 3.68mm, $p < 0.05$), but not between the symmetric subpopulation and overall population (3.28mm vs. 3.73mm, $p > 0.05$).

DISCUSSION

In this study, a fully automated AI model was successfully developed for segmenting the prostatic urethra from anatomic MRIs. The general trend for performance was highest at the apex and lowest at the base of the prostate (Figures 6,7,9,10). This urethra location-dependent accuracy was more pronounced in cases that were asymmetrically enlarged with BPH, but only towards the base where the influence of the transition zone BPH is most prominent (Figures 5–10). The performance metrics suggest that this AI model can be accurately used in anatomic prostate MRIs as a tool that supports the clinicians' workflow while they segment the prostatic urethra for treatment planning purposes. In the case of prostates with symmetrical BPH, this model can confidently predict the full urethra path, whereas for asymmetrical prostates, this model can confidently predict the path of the distal two-thirds and would likely need manual correction for the proximal one-third given the tortuosity of the urethra often observed at the base of the gland. In both scenarios, this AI model is likely to assist this labor-intensive segmentation task for clinicians and medical physicists who may be less familiar with detailed radiologic anatomy of the prostate gland.

In comparison with previous studies, our AI model was found to perform similarly. When our AI model is compared to the method found in Acosta et al. (8), the proposed AI model achieves better mean CLD in the mid and apical prostatic urethra positions, comparable mean CLD in the base, with an overall difference of 2.56mm vs 3.25mm. In comparison to the method found in Takagi et al. (9), the proposed AI model achieves worse mean CLD in the base and the mid, better mean CLD in the apex but worse overall mean CLD of 2.56mm vs 2.09mm. It is notable that the magnitude of the mean error in our predictive model as measured by the CLD of the overall urethra is within previously specified safety margins used in prospective clinical trials (20), although, as noted, at this time for clinical practice, manual verification with correction is still warranted especially in cases of asymmetric prostates at the base.

The performance differences between our study and prior works may be attributed to limitations with CT scans and use of urinary catheters for ground truth creation in the prior studies. As mentioned, the previously published work (8,9) trained their models on CT scans with each urethra localized with a urinary catheter in situ. Imaging with a urinary catheter is not commonplace for external beam radiotherapy planning and requires increased resource utilization and may increase the risk of infection (21). The application of a model trained and validated on datasets with a urinary catheter in situ may not predict the natural course of the urethra. One prior study which has investigated the difference in urethral location conditional on the presence or absence of a urinary catheter in situ provides evidence for this, estimating a median in-plane difference between urethral location of 2 mm between the two conditions and a clinically significant difference, defined by ≥ 3mm, in more than 20% of patients studied. Of interest, these investigators found that the differences were the greatest at the base of the prostate gland (22) which is the sublocation that accounted for the greatest performance gap from the previous model proposed by Acosta and Takagi. Another prior study has shown that catheterization may produce rotational changes to both the urethra and prostate leading to subsequent errors in planning and delivered dosimetry (23). In combination, these studies highlight concerns with using models built on catheterized

CT data for standard clinical practice when urinary catheters are not regularly used in radiation planning studies. These models may have worse performance than our model under these conditions, inaccurately predicting non-catheterized urethral location and leading to a clinically significant overdose to the urethral tissue itself.

Our research findings have a few future implications. This end-to-end AI model only requires axial T2W MRI data and operates fully automated, which can assist urethral avoidance during external beam radiotherapy planning. Furthermore, automated segmentation of the urethra could be used to aid in large scale retrospective analyses attempting to find dosimetric correlates of toxicity or can be used for sessions of direct MRI-guided radiotherapy (e.g., MRIdian ViewRay). Finally, this AI model can be utilized to choose ideal candidates for image guided focal therapy applications and can help to avoid iatrogenic urethra injuries.

Our study has five main limitations. First, urethra annotations were prepared by a single highly experienced genitourinary radiologist. For AI model development, ideally, annotations should be made by multiple readers, however, our institutional infrastructure in this field did not permit this and although this can be viewed as a limitation, a uniform method of urethra segmentation is an advantage for obtaining consistent performance for this tedious task. Second, our model was not prospectively or retrospectively used for treatment planning purposes in this study, therefore its actual impact on this task is unknown. However, we aim to test this AI model in our radiation and focal therapy planning in near future. Third, the proportion of asymmetrical vs symmetrical BPH cases was not assessed for the training population since these variations were encountered during retrospective failure analysis phase of our study. Fourth, the results of the model are likely influenced by the heterogeneity of the training data from different MRI machines/protocols and multiple institutions. As a result, this model is not directly comparable to other models built on single institution, homogenous data. However, we view heterogeneity of data as a strength for AI models allowing for greater generalizability. Fifth, we did not split the data into cohorts and conduct sub-analyses comparing AI results on each source of data heterogeneity.

In conclusion, we developed an end-to-end AI approach which can automatically segment prostatic urethra with promising performance metrics. This AI model yielded superior performance in symmetric prostate glands and underperformed in prostate gland with heavy burdens of hyperplastic adenomas. This system can be potentially used to assist treatment planning in patients undergoing radiation therapy or focal therapy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

The full study protocol can be accessed in the online supplementary material. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the

U.S. Government. Additional research support was provided by the NIH Medical Research Scholars Program, a public-private partnership supported jointly by the NIH and contributions to the Foundation for the NIH from the Doris Duke Charitable Foundation, the American Association for Dental Research, and the Colgate-Palmolive Company.

REFERENCES

1. Key Statistics for Prostate Cancer | Prostate Cancer Facts [Internet]. [cited 2021 Sep 7]. Available from: https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html
2. Donovan JL, Hamdy FC, Lane JA, Mason M, Metcalfe C, Walsh E, et al. Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. N Engl J Med. 2016 Oct 13;375(15):1425–37. [PubMed: 27626365]
3. Dinh CV, Steenbergen P, Ghobadi G, Heijmink SWTJP, Pos FJ, Haustermans K, et al. Magnetic resonance imaging for prostate cancer radiotherapy. Phys Med. 2016 Mar 1;32(3):446–51. [PubMed: 26858164]
4. Wang K, Mavroidis P, Royce TJ, Falchook AD, Collins SP, Sapareto S, et al. Prostate Stereotactic Body Radiation Therapy: An Overview of Toxicity and Dose Response. Int J Radiat Oncol Biol Phys. 2021 May 1;110(1):237–48. [PubMed: 33358229]
5. Leeman JE, Chen Y-H, Catalano P, Bredfeldt J, King M, Mouw KW, et al. Radiation Dose to the Intraprostatic Urethra Correlates Strongly With Urinary Toxicity After Prostate Stereotactic Body Radiation Therapy: A Combined Analysis of 23 Prospective Clinical Trials. Int J Radiat Oncol Biol Phys. 2021 Aug 16;S0360–3016(21)00844–0.
6. Merrick GS, Butler WM, Tollenaar BG, Galbreath RW, Lief JH. The dosimetry of prostate brachytherapy-induced urethral strictures. Int J Radiat Oncol Biol Phys. 2002 Feb 1;52(2):461–8. [PubMed: 11872293]
7. Akimoto T, Ito K, Saitoh J, Noda S, Harashima K, Sakurai H, et al. Acute genitourinary toxicity after high-dose-rate (HDR) brachytherapy combined with hypofractionated external-beam radiation therapy for localized prostate cancer: correlation between the urethral dose in HDR brachytherapy and the severity of acute genitourinary toxicity. Int J Radiat Oncol Biol Phys. 2005 Oct 1;63(2):463–71. [PubMed: 16168838]
8. Acosta O, Mylona E, Le Dain M, Voisin C, Lizee T, Rigaud B, et al. Multi-atlas-based segmentation of prostatic urethra from planning CT imaging to quantify dose distribution in prostate cancer radiotherapy. Radiother Oncol. 2017 Dec 1;125(3):492–9. [PubMed: 29031609]
9. Takagi H, Kadoya N, Kajikawa T, Tanaka S, Takayama Y, Chiba T, et al. Multi-atlas–based auto-segmentation for prostatic urethra using novel prediction of deformable image registration accuracy. Med Phys. 2020;47(7):3023–31. [PubMed: 32201958]
10. Kalantar R, Lin G, Winfield JM, Messiou C, Lalondrelle S, Blackledge MD, et al. Automatic Segmentation of Pelvic Cancers Using Deep Learning: State-of-the-Art Approaches and Challenges. Diagnostics. 2021 Oct 22;11(11):1964. [PubMed: 34829310]
11. Elguindi S, Zelefsky MJ, Jiang J, Veeraraghavan H, Deasy JO, Hunt MA, et al. Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy. Phys Imaging Radiat Oncol. 2019 Oct;12:80–6. [PubMed: 32355894]
12. Litjens G, Debats O, Barentsz J, Karssemeijer N, Huisman H. Computer-aided detection of prostate cancer in MRI. IEEE Trans Med Imaging. 2014 May;33(5):1083–92. [PubMed: 24770913]
13. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging. 2013 Dec;26(6):1045–57. [PubMed: 23884657]
14. Sanford TH, Harmon SA, Kesani D, Gurram S, Gupta N, Mehralivand S, et al. Quantitative Characterization of the Prostatic Urethra Using MRI: Implications for Lower Urinary Tract Symptoms in Patients with Benign Prostatic Hyperplasia. Acad Radiol. 2021 May;28(5):664–70. [PubMed: 32307270]
15. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Mach Learn PYTHON. :6.

16. Kerfoot E, Clough J, Oksuz I, Lee J, King AP, Schnabel JA. Left-Ventricle Quantification Using Residual U-Net. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young A, et al., editors. Statistical Atlases and Computational Models of the Heart Atrial Segmentation and LV Quantification Challenges. Cham: Springer International Publishing; 2019. p. 371–80. (Lecture Notes in Computer Science).
17. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Cham: Springer International Publishing; 2015. p. 234–41. (Lecture Notes in Computer Science).
18. Consortium M. MONAI: Medical Open Network for AI [Internet]. Zenodo; 2020 [cited 2021 Nov 30]. Available from: https://zenodo.org/record/5728262
19. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020 Mar 2;17(3):261–72. [PubMed: 32015543]
20. Jaccard M, Zilli T, Dubouloz A, Escude L, Jorcano S, Linthout N, et al. Urethra-Sparing Stereotactic Body Radiation Therapy for Prostate Cancer: Quality Assurance of a Randomized Phase 2 Trial. Int J Radiat Oncol Biol Phys. 2020 Nov 15;108(4):1047–54. [PubMed: 32535161]
21. Kranz J, Schmidt S, Wagenlehner F, Schneidewind L Catheter-Associated Urinary Tract Infections in Adult Patients. Dtsch Ärztebl Int. 2020 Feb;117(6):83–8. [PubMed: 32102727]
22. Dekura Y, Nishioka K, Hashimoto T, Miyamoto N, Suzuki R, Yoshimura T, et al. The urethral position may shift due to urethral catheter placement in the treatment planning for prostate radiation therapy. Radiat Oncol Lond Engl. 2019 Dec 12;14(1):226.
23. Litzenberg DW, Muenz DG, Archer PG, Jackson WC, Hamstra DA, Hearn JW, et al. Changes in prostate orientation due to removal of a Foley catheter. Med Phys. 2018 Apr;45(4):1369–78. [PubMed: 29474748]

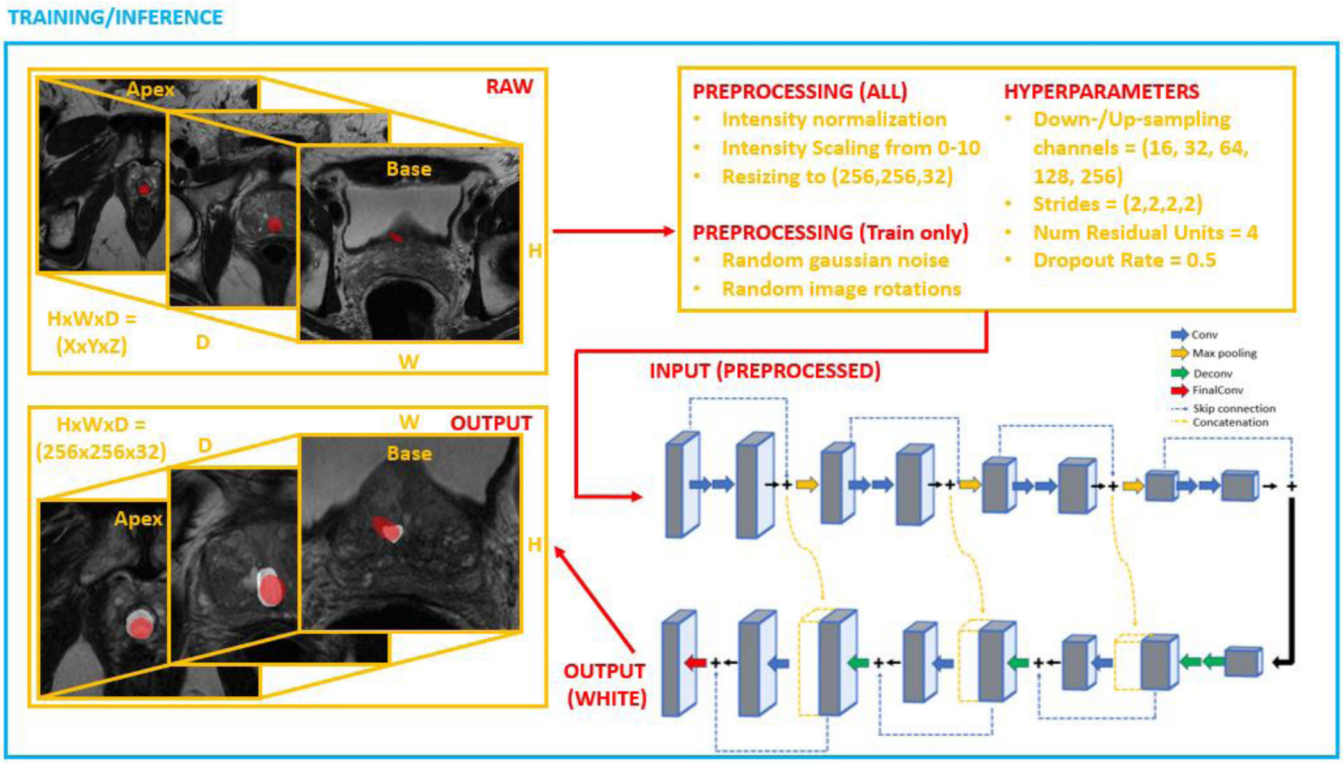

Figure 1.
UNet Training/Inference Workflow from raw input (MRI) to the output (white binary segmentation)

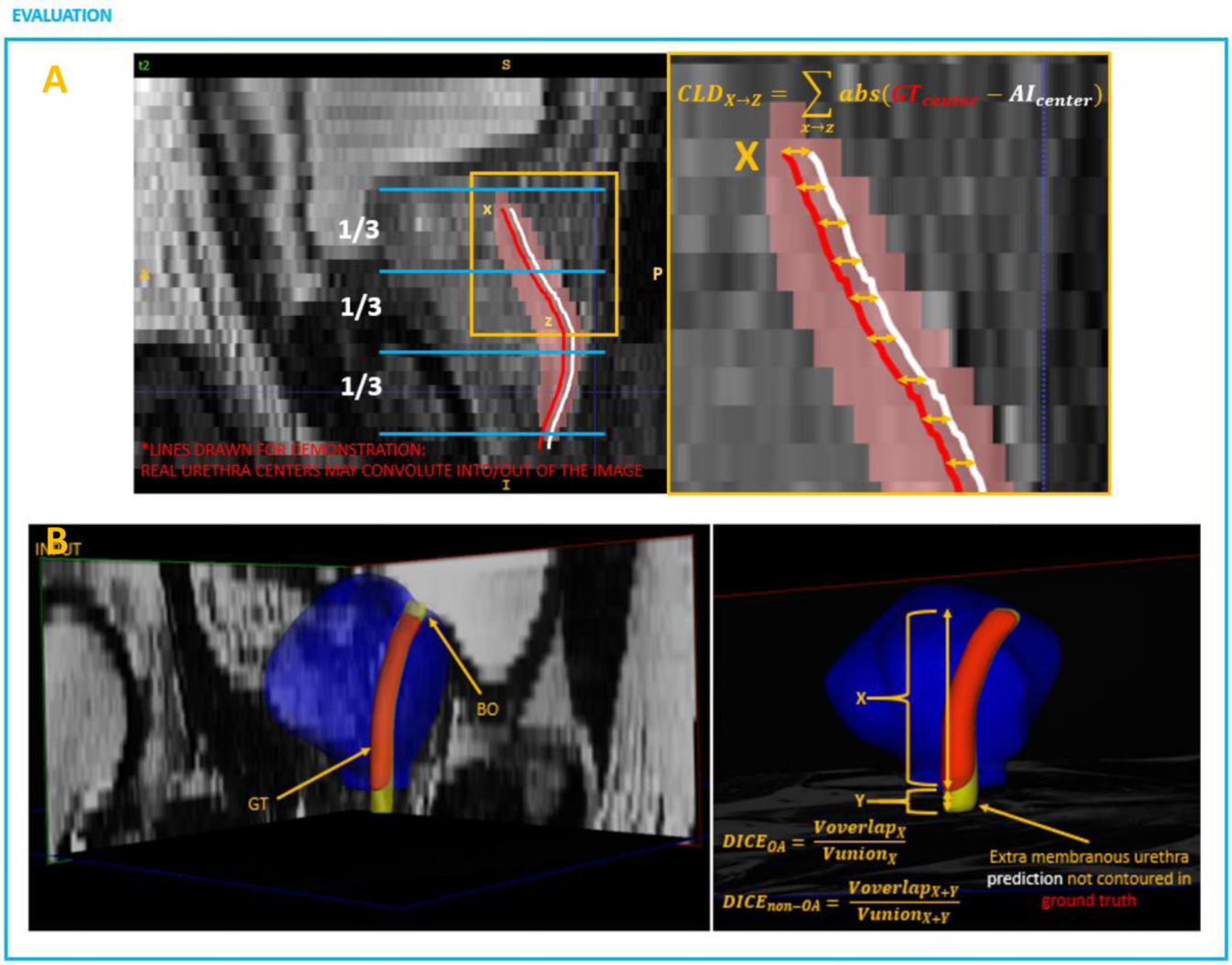

Figure 2:

(A) Centerline distance (CLD) calculation between ground truth and AI prediction. For length-partitioning, the urethra is split into 1/3s based on the z-dimension. (B) Organ-adjusted (OA) calculations are based on preventing the penalization of the AI correctly segmenting the membranous urethra which is not captured by the ground truth.

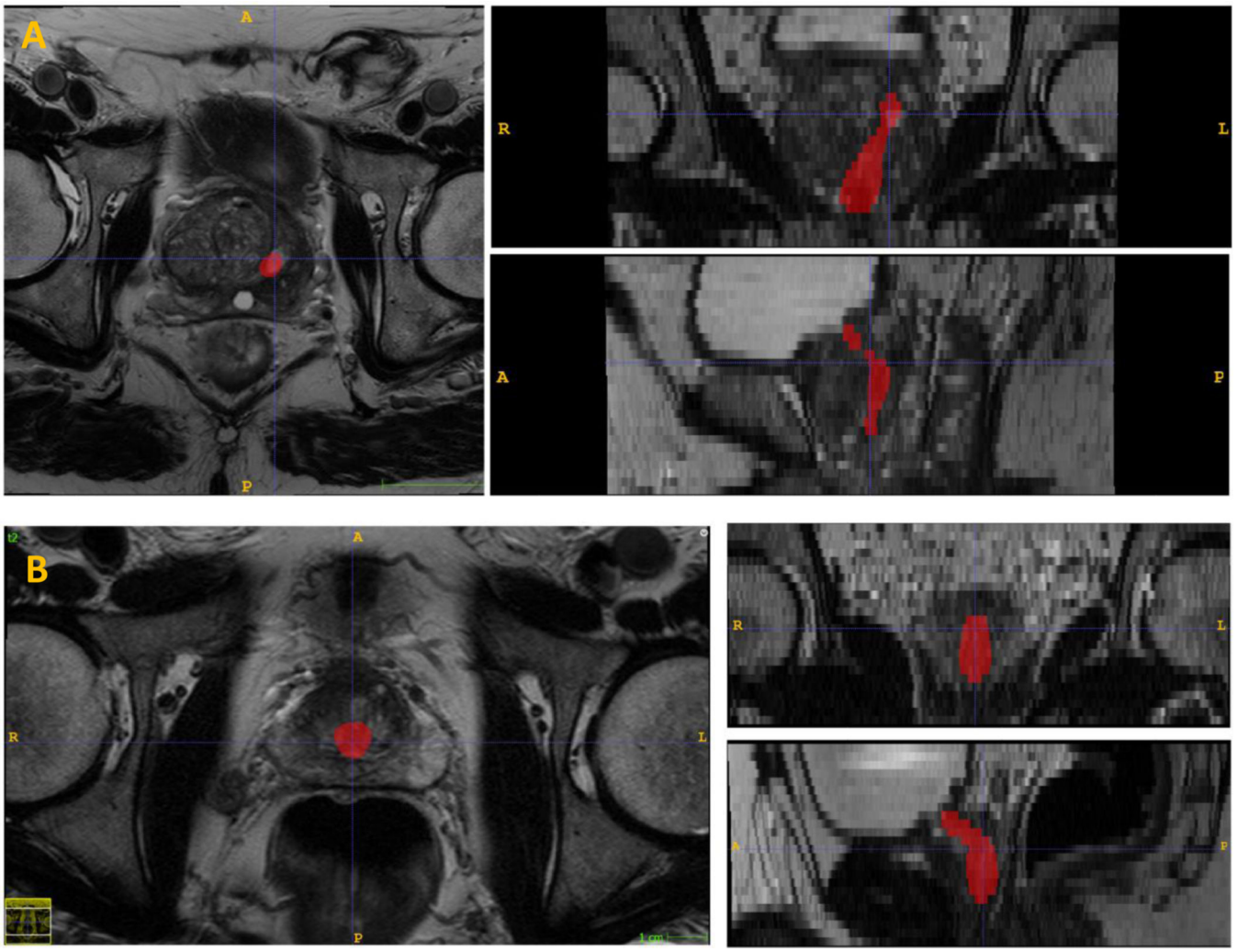

Figure 3.

(A) Example of asymmetrical prostate gland with convoluted urethra path and (B) example of a symmetrical prostate gland with non-convoluted urethra path

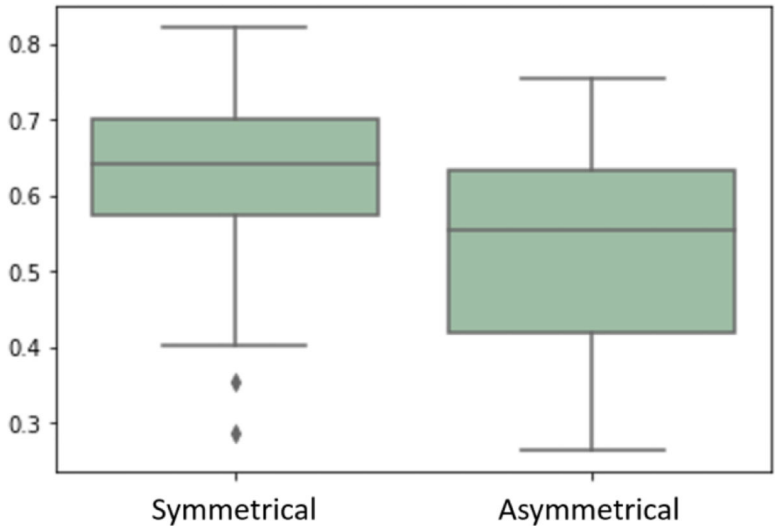

Figure 4.
Non-organ Adjusted Dice Score: Asymmetrical vs Symmetrical

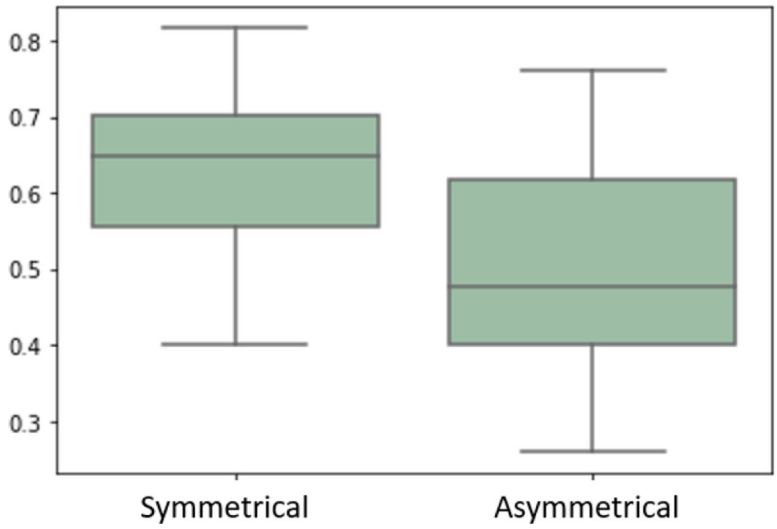

Figure 5.
Organ-adjusted Dice Score: Asymmetrical vs Symmetrical

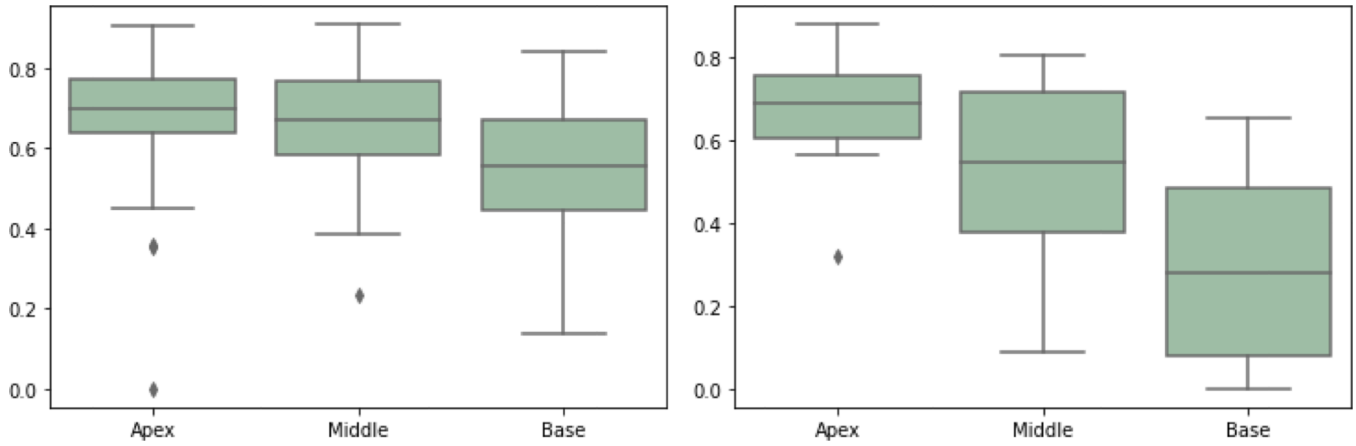

Figure 6.
Organ-adjusted Dice Score vs Urethra Position for Symmetrical and Asymmetrical Prostates

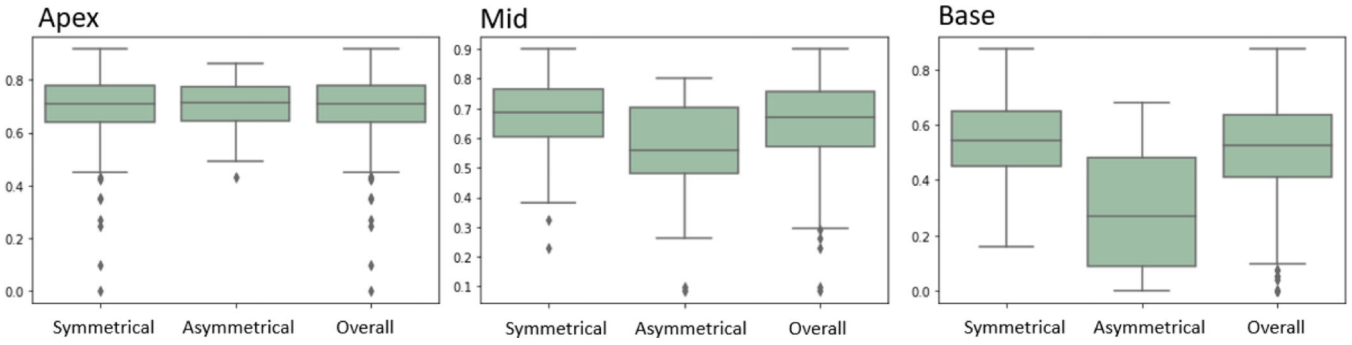

Figure 7.

Organ-adjusted Dice Score vs Symmetrical and Asymmetrical Prostates Stratified by Urethra Position

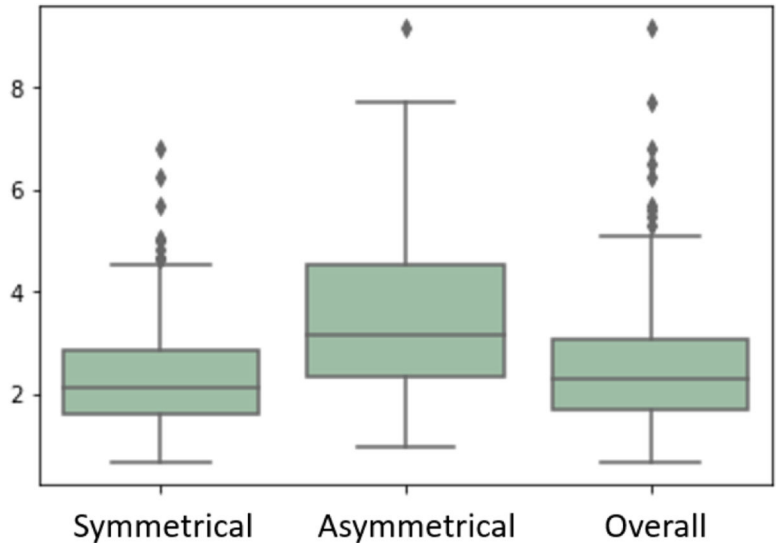

Figure 8.
Centerline Difference of Ground Truth and Prediction stratified by prostate asymmetry status

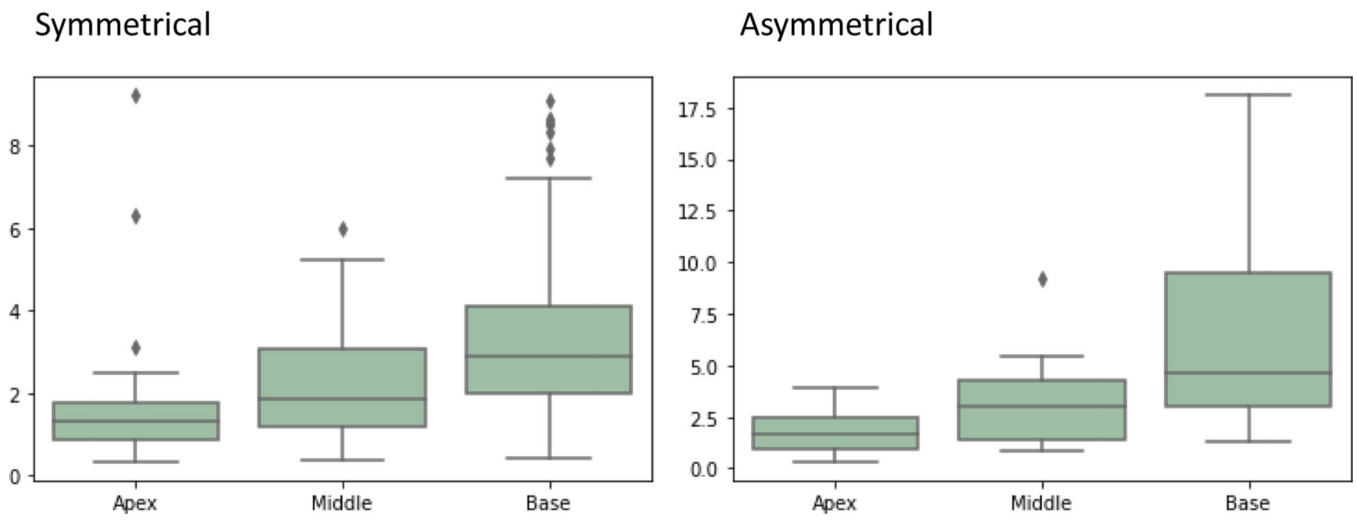

Figure 9.

Centerline Difference of Ground Truth and Prediction Grouped by Symmetrical and Asymmetrical prostates and stratified by prostatic urethra location.

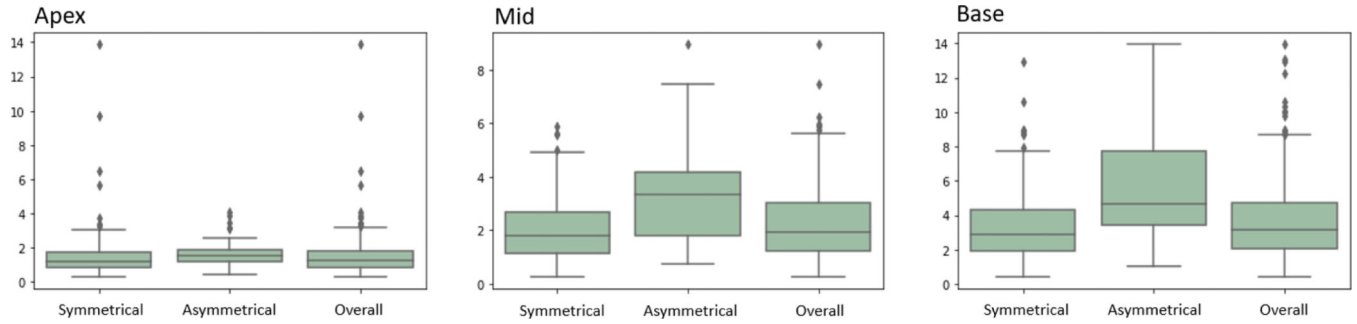

Figure 10.

Centerline Difference of Ground Truth and Prediction: Grouped by prostatic urethra location and stratified by presence of asymmetry.

Table 1

Patient demographics per partition

Group	Age (median)	PSA (median)	Prostate Volume (median)	Asymmetrical Prostates (n=)	PROSTAT Ex Distribution (n=95)	ERC Distribution (n=402)
Train (n=657)	**65.95**	**6.76**	**56**	**---**	**64**	**284**
Validation (n=141)	**65.65**	**8.07**	**51**	**20**	**18**	**63**
Test (n=140)	**66.11**	**6.25**	**56**	**23**	**13**	**55**

Table 2

Model Centerline Distance Results.

	Acosta et al. (2017) Overall (*n* = 1)	**Takagi et al. (2020) Overall (*n* = 20)**	**Belue et al. Non-Asymmetrical (*n* = 117)**	**Belue et al. Asymmetrical (*n* = 23)**	**Belue et al. Overall (*n* = 140)**
Top (Base) 1/3	3.7±1.7 mm	2.49±1.78 mm	3.28±1.91 mm	6.24±4.33 mm	3.73±2.64 mm
Middle 1/3	2.52±1.5 mm	1.86±0.89 mm	2.15±1.18 mm	3.11±1.92 mm	2.31±1.37 mm
Bottom (Apex) 1/3	3.01±1.7 mm	1.92±0.94 mm	1.43±1.03 mm	1.83±1.04 mm	1.49±1.04 mm
Overall	3.25±1.2 mm	2.09±0.89 mm	2.32±0.94 mm	3.79±2.19 mm	2.56±1.34 mm