



Anatomy-Informed Data Augmentation for Enhanced Prostate Cancer Detection

Balint Kovacs^{1,2,3}(✉) , Nils Netzer^{2,3} , Michael Baumgartner^{1,4,5} ,
 Carolin Eith^{2,3}, Dimitrios Bounias^{1,3}, Clara Meinzer², Paul F. Jäger^{5,6} ,
 Kevin S. Zhang², Ralf Floca¹, Adrian Schrader^{2,3} , Fabian Isensee^{1,5} ,
 Regula Gnirs², Magdalena Görtz^{7,8}, Viktoria Schütz⁷, Albrecht Stenzinger⁹,
 Markus Hohenfellner⁷, Heinz-Peter Schlemmer², Ivo Wolf¹⁰ ,
 David Bonekamp² , and Klaus H. Maier-Hein^{1,5,11}

¹ Division of Medical Image Computing, German Cancer Research Center (DKFZ), Heidelberg, Germany

balint.kovacs@dkfz-heidelberg.de

² Division of Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

³ Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany

⁴ Faculty of Mathematics and Computer Science, Heidelberg University, Heidelberg, Germany

⁵ Helmholtz Imaging, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁶ Interactive Machine Learning Group, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁷ Department of Urology, University of Heidelberg Medical Center, Heidelberg, Germany

⁸ Junior Clinical Cooperation Unit ‘Multiparametric Methods for Early Detection of Prostate Cancer’, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁹ Institute of Pathology, University of Heidelberg Medical Center, Heidelberg, Germany

¹⁰ Mannheim University of Applied Sciences, Mannheim, Germany

¹¹ Pattern Analysis and Learning Group, Department of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany

Abstract. Data augmentation (DA) is a key factor in medical image analysis, such as in prostate cancer (PCa) detection on magnetic resonance images. State-of-the-art computer-aided diagnosis systems still rely on simplistic spatial transformations to preserve the pathological label post transformation. However, such augmentations do not substantially increase the organ as well as tumor shape variability in the training set, limiting the model’s ability to generalize to unseen cases with more diverse localized soft-tissue deformations. We propose a new anatomy-informed transformation that leverages information from adjacent organs to simulate typical physiological deformations of the prostate and generates unique lesion shapes without altering their label. Due to its lightweight

D. Bonekamp and K. H. Maier-Hein—Equal contribution.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43990-2_50.

computational requirements, it can be easily integrated into common DA frameworks. We demonstrate the effectiveness of our augmentation on a dataset of 774 biopsy-confirmed examinations, by evaluating a state-of-the-art method for PCa detection with different augmentation settings.

Keywords: Data augmentation · Soft-tissue deformation · Prostate cancer detection

1 Introduction

Data augmentation (DA) is a key factor in the success of deep neural networks (DNN) as it artificially enlarges the training set to increase their generalization ability as well as robustness [22]. It plays a crucial role in medical image analysis [8] where annotated datasets are only available with limited size. DNNs have already successfully supported radiologists in the interpretation of magnetic resonance images (MRI) for prostate cancer (PCa) diagnosis [3]. However, the DA scheme received less attention, despite its potential to leverage the data characteristic and address overfitting as the root of generalization problems.

State-of-the-art approaches still rely on simplistic spatial transformations, like translation, rotation, cropping, and scaling by globally augmenting the MRI sequences [12, 20]. They exclude random elastic deformations, which can change the lesion outline but might alter the underlying label and thus produce counter-productive examples for training [22]. However, soft tissue deformations, which are currently missing from the DA schemes, are known to significantly affect the image morphology and therefore play a critical role in accurate diagnosis [6].

Both lesion and prostate shape geometrical appearance influence the clinical assessment of Prostate Imaging-Reporting and Data System (PI-RADS) [24]. The prostate constantly undergoes soft tissue deformation dependent on muscle contractions, respiration, and more importantly variable filling of the adjacent organs, namely the bladder and the rectum. Among these sources, the rectum has the largest influence on the prostate and lesion shape variability due to its large motion [4] and the fact that the majority of the lesions are located in the adjacent peripheral prostate zone [1]. However, only one snapshot of all these functional states is captured within each MRI examination, and almost never will be exactly the same on any repeat or subsequent examination. Ignoring these deformations in the DA scheme can potentially limit model performance.

Model-driven transformations attempting to simulate organ functions - like respiration, urinary excretion, cardiovascular- and digestion mechanics - offer a high degree of diversity while also providing realistic transformations. Currently, the finite element method (FEM) is the standard for modeling biomechanics [13]. However, their computation is overly complex [10] and therefore does not scale to on-the-fly DA [7]. Recent motion models rely on DNNs using either a FEM model [15] or complex training with population-based models [18]. Motion models have not been integrated into any deep learning framework as an online data augmentation yet, thereby leaving the high potential of inducing application-specific knowledge into the training procedure unexploited.

In this work we propose an anatomy-informed spatial augmentation, which leverages information from adjacent organs to mimic typical deformations of the prostate. Due to its lightweight computational requirements, it can be easily integrated into common DA frameworks. This technique allows us to simulate different physiological states during the training and enrich our dataset with a wider range of organ and lesion shapes. Inducing this kind of soft tissue deformation ultimately led to improved model performance in patient- and lesion-level PCa detection on an independent test set.

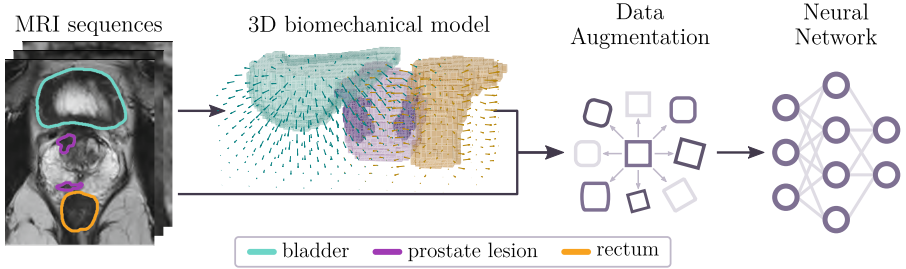


Fig. 1. The proposed anatomy-informed prostate augmentation. Simulating typical physiologic changes in adjacent organs enlarges the training set with realistic soft tissue deformations of the prostate thereby increasing the generalization ability as well as the robustness of the network. Due to its lightweight computational requirements, it can be easily integrated into online network training.

2 Methods

2.1 Mathematical Model of the Anatomy-Informed Deformation

Model-driven spatial transformations simulate realistic soft-tissue deformations, which are part of the physiology, and can highly affect the shape of the prostate as well as the lesions in it. As the computation of state-of-the-art FEM models does not scale to on-the-fly DA, we introduce simplifications to be able to integrate such a biomechanical model as an online DA into the model training:

- soft tissue deformation of the prostate is mostly the result of morphological changes in the bladder and rectal space [4,6],
- due to the isotropic mechanical behavior of the rectum and the bladder [19], we apply isotropic deformation to them,
- we assume similar elastic modulus between the prostate and surrounding muscles [16], allowing us to approximate these tissue classes as homogeneous,
- we introduce a non-linear component into the model by transforming the surrounding tissue proportionally to the distance from the rectum and bladder in order to generate realistic deformations [23].

Based on them, we define the vector field V for the transformation as the gradient of the convolution between the Gaussian kernel G_σ and the indicator function S_{organ} , multiplied by a scalar C to control deformation amplitude and direction:

$$V = \nabla(G_\sigma * S_{organ}(x, y, z)) \cdot C. \quad (1)$$

The resulting V serves as the deformation field for an MRI sequence $I(x, y, z)$:

$$I_{deformed}(x, y, z) = I(x + V_x(x, y, z), y + V_y(x, y, z), z + V_z(x, y, z)). \quad (2)$$

It allows us to simulate the distension or evacuation of the bladder or rectal space. We refer to this transformation as anatomy-informed deformation. We make it publicly available in Batchgenerators [9] and integrate it into a nnU-Net trainer https://github.com/MIC-DKFZ/anatomy_informed_DA.

2.2 Experimental Setting

We evaluate our anatomy-informed DA qualitatively as well as quantitatively.

First, we visually inspect whether our assumptions in Sect. 2.1 regarding pelvic biomechanics resulted in realistic transformations. We apply either our proposed transformation to the rectum or the bladder, random deformable or no transformation in randomly selected exams and conduct a strict Turing test with clinicians having different levels of radiology expertise (a freshly graduated clinician (C.E.) and resident radiologists (C.M., K.S.Z.), 1.5 - 3 years of experience in prostate MRI) to determine if they can notice the artificial deformation.

Finally, we quantify the effect of our proposed transformation on the clinical task of patient-level PCa diagnosis and lesion-level PCa detection. We derive the diagnosis through semantic segmentation of the malignant lesions following previous studies [5, 11, 12, 20, 21]. Semantic segmentation provides interpretable predictions that are sensitive to spatial transformations, making it appropriate for testing spatial DAs. To compare the performance of the trained models to radiologists, we calculate their performance using the clinical PI-RADS scores and histopathological ground truths. To consider clinically informative results, we use the partial area under the Receiver Operating Characteristic (pAUROC) for patient-level evaluation with the sensitivity threshold of 78.75%, which is 90% of the sensitivity of radiologists for PI-RADS ≥ 4 . Additionally, we calculate the F_1 -score at the sensitivity of PI-RADS ≥ 4 . Afterward, we evaluate model performances on object-level using the Free-Response Receiver Operating Characteristic (FROC) and the number of detections at the radiologists' lesion level performance for PI-RADS ≥ 4 , at 0.32 average number of False Positives per scan. Objects were derived by applying a threshold of 0.5 to the softmax outputs followed by connected component analysis to identify connected regions in the segmentation maps. Predictions with an Intersection over Union of 0.1 with a ground truth object were considered True Positives. To systematically compare the effect of our proposed anatomy-informed DA with the commonly used settings, we create three main DA schemes:

1. **Basic** DA setting of nnU-Net [8], which is an extensive augmentation pipeline containing simple spatial transformations, namely translation, rotation and scaling. This setting is our reference DA scheme.
2. **Random deformable** transformations as implemented in the nnU-Net [8] DA pipeline extending the basic DA scheme (1) to test its presence in the medical domain. Our hypothesis is that it will produce counterproductive examples, resulting in inferior performance compared to our proposed DA.

3. Proposed **anatomy-informed** transformation in addition to the simple DA scheme (1). We define two variants of it:
 - (a) Deforming only the rectum, as rectal distension has the highest influence among the organs on the shapes of the prostate lesions [4].
 - (b) Deforming the bladder in addition to the rectum, as bladder deformations also have an influence on lesions, although smaller.

2.3 Prostate MRI Data

774 consecutive bi-parametric prostate MRI examinations are included in this study, which were acquired in-house during the clinical routine. The ethics committee of the Medical Faculty Heidelberg approved the study (S-164/2019) and waived informed consent to enable analysis of a consecutive cohort. All experiments were performed in accordance with the declaration of Helsinki [2] and relevant data privacy regulations. For every exam, PI-RADS v2 [24] interpretation was performed by a board-certified radiologist. Every patient underwent extended systematic and targeted MRI trans-rectal ultrasound-fusion transperineal biopsy. Malignancy of the segmented lesions was determined from a systematic-enhanced lesion ground-truth histopathological assessment, which has demonstrated reliable ground-truth assessment with sensitivity comparable to radical prostatectomy [17]. The samples were evaluated according to the International Society of Urological Pathology (ISUP) standards under the supervision of a dedicated uropathologist. Clinically significant prostate cancer (csPCa) was defined as ISUP grade 2 or higher. Based on the biopsy results, every csPCa lesion was segmented on the T2-weighted sequences retrospectively by multiple in-house investigators under the supervision of a board-certified radiologist. In addition to the lesions, the rectum and the bladder segmentations were automatically predicted by a model built upon nnU-Net [8] trained iteratively on an in-house cohort initially containing a small portion of our cohort. Multiple radiologists confirmed the quality of the predicted segmentations.

2.4 Training Protocol

774 exams were split into 80% training set (619 exams) and 20% test set (155 exams) by stratifying them based on the prevalence of csPCa (36.3%). The MRI sequences were registered using B-spline transformation based on mutual information to match the ground-truth segmentations across all modalities [12, 14]. As the limited number of exams with csPCa and the small lesion size compared to the whole image can cause instability during training, we adapted the cropping strategy from [21] by keeping the organ segmentations to use the anatomy-informed DA (offsets of ± 9 mm axial to the prostate and ± 11.25 mm in the axial plane to the rectum and the bladder). The images are preprocessed by the automated algorithm of nnU-Net [8]. We trained 3D nnU-Net models in 5-fold cross-validation with different spatial DA schemes, see Sect. 2.2. The hyperparameter C of the anatomy-informed DA was optimized using validation

results, sampled during training with uniform distribution constrained by amplitude values in positive and negative directions of $C = \{300, 600, 900, 1200, 1500\}$. $C_{rectum} = 1200$ and $C_{bladder} = 600$ were selected for the final models. Compared to the standard nnU-Net settings, we implemented balanced sampling regarding the prevalence of csPCa and reduced the number of epochs to 350 to avoid overfitting. We used Mish activation function, Ranger optimizer, cosine anneal learning rate scheduler, and initial learning rate of 0.001 following [12]. The final models are ensembled and evaluated on the independent test set using bootstrapping with 1000 replications to provide standard deviation and to calculate p-values for the F_1 -score and for the number of detected lesions using two-sided t-test to determine statistical significance.

3 Results

The anatomy-informed transformation produced highly realistic soft tissue deformations. Figure 2 shows an example of the transformation simulating rectum distensions with prostate lesions at different distances from the rectum. 92% of the rectum and 93% of the bladder deformation from the randomly picked exams became so realistic that our freshly graduated clinician did not detect them, but our residents noticed 87.5% of the rectum and 25% of the bladder deformations based on small transformation artifacts and their expert intuition. Irregularities resulted from the random elastic deformations can be easily detected, in contrast to our method being challenging to detect its artificial nature.

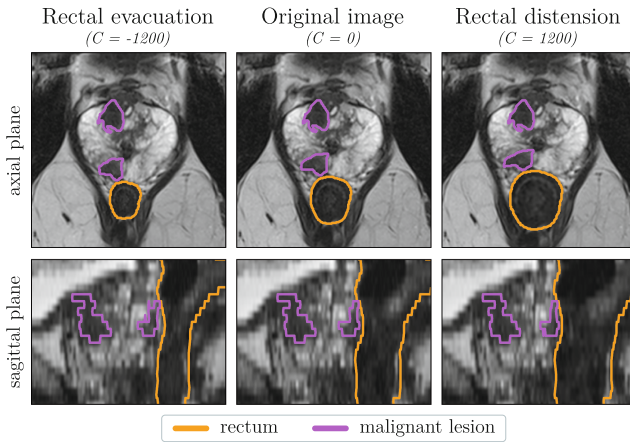


Fig. 2. Results of the proposed anatomy-informed deformation on the rectum, showing the outlines of the rectum and malignant lesions. The middle images show the original MRI sequence, the left images simulate rectal space evacuation, while the right images rectal distension. The transformation induces localized soft tissue deformations, resulting in changes in the shape of lesions only in the adjacent peripheral prostate zone.

In Table 1 we summarize the patient-level pAUROC and F_1 -scores; and lesion-level FROC results on the independent test set showing the advantage of using anatomy-informed DA. To further highlight the practical advantage of the proposed augmentation, we compare the performance of the trained models to the radiologists' diagnostic performance for PI-RADS ≥ 4 , which locate the most informative performance point clinically on the ROC diagram, see Fig. 3.

Table 1. Prostate cancer detection results on our independent test set

DA scheme	pAUROC	F_1 -score	FROC
1. basic (reference)	$44.33 \pm 11.65\%$	$57.31 \pm 3.14\%$	$58.14 \pm 5.79\%$
2. random elastic	$38.94 \pm 14.38\%$	$56.98 \pm 3.08\%$	$58.63 \pm 5.42\%$
3.a) proposed (rectum)	$59.92 \pm 13.27\%$	$61.64 \pm 3.61\%$	$59.55 \pm 5.97\%$
3.b) proposed (rectum + bladder)	$53.27 \pm 13.42\%$	$62.42 \pm 3.84\%$	$59.93 \pm 5.53\%$

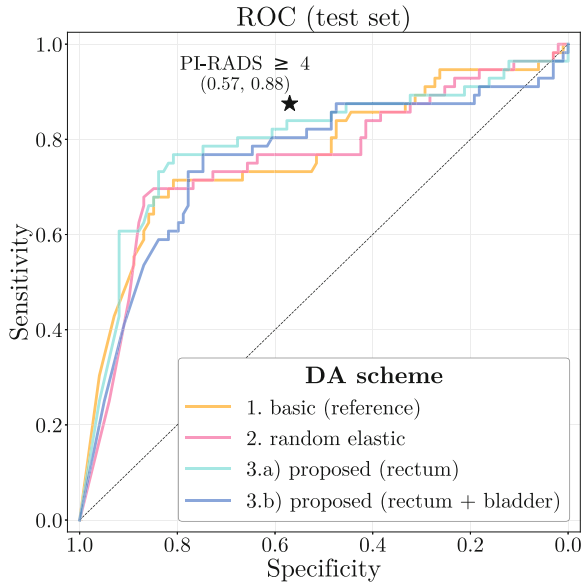


Fig. 3. The effect of different spatial DA schemes on the ROC. The radiologists' performance with PI-RADS ≥ 4 is marked to locate the most informative performance point clinically. Both variants of the proposed anatomy-informed DA (3.a and 3.b) increased the sensitivity value around the clinical PI-RADS ≥ 4 performance point compared to the simple (1) and random elastic (2) DA schemes, approaching it closely.

Extending the basic DA scheme with the proposed anatomy-informed deformation not only increased the sensitivity closely matching the radiologists' patient-level diagnostic performance but also improved the detection of PCa

on a lesion level. Interestingly, while the use of random deformable transformation also improved lesion-level performance, it did not approach the diagnostic performance of the radiologists, unlike the anatomy-informed DA.

At the selected patient- and object-level working points, the model with the proposed rectum- and bladder-informed DA scheme reached the best results with significant improvements ($p < 0.05$) compared to the model with the basic DA setting by increasing the F_1 -score with 5.11% and identifying 4 more lesions (5.3%) from the 76 lesions in our test set.

The time overhead introduced by anatomy-informed augmentation caused no increase in the training time, the GPU remained the main bottleneck.

4 Discussion

This paper addresses the utilization of anatomy-informed spatial transformations in the training procedure to increase lesion, prostate, and adjacent organ shape variability for the task of PCa diagnosis. For this purpose, a lightweight mathematical model is built for simulating organ-specific soft tissue deformations. The model is integrated into a well-known DA framework and used in model training for enhanced PCa detection.

Towards Radiologists’ Performance. Inducing lesion shape variability via anatomy-informed augmentation to the training process improved the lesion detection performance and increased the sensitivity value towards radiologist-level performance in PCa diagnosis in contrast to the training with the basic DA setting. These soft tissue deformations are part of physiology, but only one snapshot is captured from the many possible functional states within each individual MR examination. Our proposed DA simulates examples of physiologic anatomical changes that may have occurred in each of the MRI training examples at the same exam time points, thereby aiding the generalization ability as well as the robustness of the network. We got additional, but slight improvements by extending the DA scheme with bladder distensions. A possible explanation for this result is that less than 30% of the lesions are located close to the bladder, and our dataset did not contain enough training examples for more improvements.

Realistic Modeling of Organ Deformation. Our proposed anatomy-informed transformation was designed to mimic real-world deformations in order to preserve essential image features. Most of the transformed sequences successfully passed the Turing test against a freshly graduated clinician with prostate MRI expertise, and some were even able to pass against radiology residents with more expertise. To support the importance of realism in DA quantitatively, we compared the performance of the basic and our anatomy-informed DA scheme with that of the random deformable transformation. The random deformable DA scheme generated high lesion shape variability, but it resulted in lower performance values. This could be due to the fact that it can also cause implausible or

even harmful image warping, distorting important features, and producing counterproductive training examples. In comparison, our proposed anatomy-informed DA outperformed the basic and random deformable DA, demonstrating the significance of realistic transformations for achieving superior model performance.

High Applicability with Limitations. The easy integration into DA frameworks and no increase in the training time make our proposed anatomy-informed DA highly applicable. Its limitation is the need for additional organ segmentations, which requires additional effort from the annotator. However, pre-trained networks for segmenting anatomical structures like nnU-Net [8] have been introduced recently, which can help to overcome this limitation. Additionally, our transformation computation allows certain errors in the organ segmentations compared to applications where fully accurate segmentations are needed. The success of anatomy-informed DA opens the research question of whether it enhances performance across diverse datasets and model backbones.

5 Conclusion

In this work, we presented a realistic anatomy-informed augmentation, which mimics typical organ deformations in the pelvis. Inducing realistic soft-tissue deformations in the model training via this kind of organ-dependent transformation increased the diagnostic accuracy for PCa, closely approaching radiologist-level performance. Due to its simple and fast calculation, it can be easily integrated into DA frameworks and can be applied to any organ with similar distension properties. Due to these advantages, the shown improvements in the downstream task strongly motivate to utilize this model as a blueprint for other applications.

References

1. Ali, A., Du Feu, A., Oliveira, P., Choudhury, A., Bristow, R.G., Baena, E.: Prostate zones and cancer: lost in transition? *Nat. Rev. Urol.* **19**(2), 101–115 (2022)
2. Association, W.M., et al.: 64th WMA general assembly Fortaleza Brazil (2013). WMA Declaration of Helsinki-Ethical Principles for Medical Research Involving Human Subjects (2018)
3. Bhattacharya, I., et al.: A review of artificial intelligence in prostate cancer detection on imaging. *Ther. Adv. Urol.* **14**, 17562872221128792 (2022)
4. Boubaker, M.B., Ganghoffer, J.F.: Bladder/prostate/rectum: biomechanical models of the mobility of pelvic organs in the context of prostate radiotherapy. In: *Biomechanics of Living Organs*, pp. 307–324. Elsevier (2017)
5. Duran, A., Dussert, G., Rouvière, O., Jaouen, T., Jodoin, P.M., Lartizien, C.: ProstAttention-Net: a deep attention model for prostate cancer segmentation by aggressiveness in MRI scans. *Med. Image Anal.* **77**, 102347 (2022)
6. Engels, R.R., Israël, B., Padhani, A.R., Barentsz, J.O.: Multiparametric magnetic resonance imaging for the detection of clinically significant prostate cancer: what urologists need to know. Part 1: acquisition. *Eur. Urol.* **77**(4), 457–468 (2020)

7. Hu, Y., et al.: Adversarial deformation regularization for training image registration neural networks. In: Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G. (eds.) MICCAI 2018. LNCS, vol. 11070, pp. 774–782. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00928-1_87
8. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**(2), 203–211 (2021)
9. Isensee, F., et al.: Batchgenerators - a python framework for data augmentation (2020). <https://github.com/MIC-DKFZ/batchgenerators>
10. Khallaghi, S., et al.: Statistical biomechanical surface registration: application to MR-TRUS fusion for prostate interventions. *IEEE Trans. Med. Imaging* **34**(12), 2535–2549 (2015)
11. Kohl, S., et al.: Adversarial networks for the detection of aggressive prostate cancer. In: Workshop on Machine Learning for Health (NIPS ML4H 2017) (2017)
12. Netzer, N., et al.: Fully automatic deep learning in bi-institutional prostate magnetic resonance imaging: effects of cohort size and heterogeneity. *Invest. Radiol.* **56**(12), 799–808 (2021)
13. Payan, Y., Ohayon, J.: Biomechanics of living organs: hyperelastic constitutive laws for finite element modeling. World Bank Publications (2017)
14. Pellicer-Valero, O.J., et al.: Deep learning for fully automatic detection, segmentation, and Gleason grade estimation of prostate cancer in multiparametric magnetic resonance images. *Sci. Rep.* **12**(1), 1–13 (2022)
15. Pfeiffer, M., Riediger, C., Weitz, J., Speidel, S.: Learning soft tissue behavior of organs for surgical navigation with convolutional neural networks. *Int. J. Comput. Assist. Radiol. Surg.* **14**(7), 1147–1155 (2019). <https://doi.org/10.1007/s11548-019-01965-7>
16. Qasim, M., et al.: Biomechanical modelling of the pelvic system: improving the accuracy of the location of neoplasms in MRI-TRUS fusion prostate biopsy. *BMC Cancer* **22**(1), 1–10 (2022)
17. Radtke, J.P., et al.: Multiparametric magnetic resonance imaging (MRI) and MRI-transrectal ultrasound fusion biopsy for index tumor detection: correlation with radical prostatectomy specimen. *Eur. Urol.* **70**(5), 846–853 (2016)
18. Romaguera, L.V., Mezheritsky, T., Mansour, R., Carrier, J.F., Kadoury, S.: Probabilistic 4D predictive model from in-room surrogates using conditional generative networks for image-guided radiotherapy. *Med. Image Anal.* **74**, 102250 (2021)
19. Rubod, C., et al.: Biomechanical properties of human pelvic organs. *Urology* **79**(4), e17–e22 (2012)
20. Saha, A., Hosseinzadeh, M., Huisman, H.: End-to-end prostate cancer detection in bpMRI via 3d CNNs: effects of attention mechanisms, clinical priori and decoupled false positive reduction. *Med. Image Anal.* **73**, 102155 (2021)
21. Sanyal, J., Banerjee, I., Hahn, L., Rubin, D.: An automated two-step pipeline for aggressive prostate lesion detection from multi-parametric MR sequence. *AMIA Summits Transl. Sci. Proceed.* **2020**, 552 (2020)
22. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *J. Big Data* **6**(1), 1–48 (2019)
23. Wang, Y., Ni, D., Qin, J., Xu, M., Xie, X., Heng, P.A.: Patient-specific deformation modelling via elastography: application to image-guided prostate interventions. *Sci. Rep.* **6**(1), 1–10 (2016)
24. Weinreb, J.C., et al.: PI-RADS prostate imaging-reporting and data system: 2015, version 2. *Eur. Urol.* **69**(1), 16–40 (2016)