



OPEN

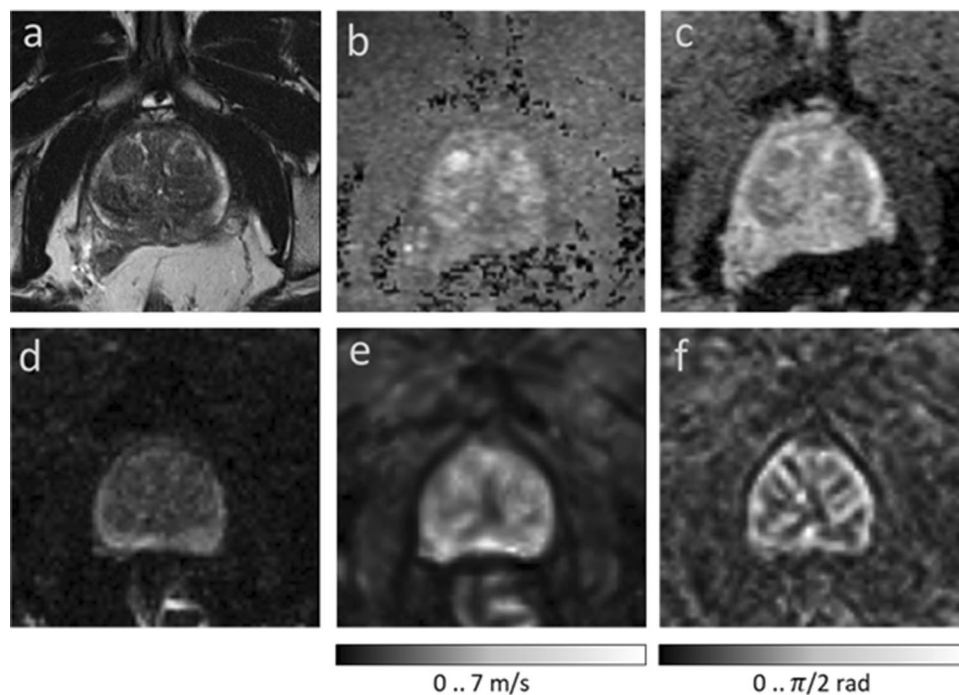
## Fully automated quantification of in vivo viscoelasticity of prostate zones using magnetic resonance elastography with Dense U-net segmentation

Nader Aldoji<sup>1</sup>, Federico Biavati<sup>1</sup>, Marc Dewey<sup>1,2,3</sup>, Anja Hennemuth<sup>4</sup>, Patrick Asbach<sup>1</sup> & Ingolf Sack<sup>1</sup>✉

Magnetic resonance elastography (MRE) for measuring viscoelasticity heavily depends on proper tissue segmentation, especially in heterogeneous organs such as the prostate. Using trained network-based image segmentation, we investigated if MRE data suffice to extract anatomical and viscoelastic information for automatic tabulation of zonal mechanical properties of the prostate. Overall, 40 patients with benign prostatic hyperplasia (BPH) or prostate cancer (PCa) were examined with three magnetic resonance imaging (MRI) sequences: T2-weighted MRI (T2w), diffusion-weighted imaging (DWI), and MRE-based tomoelastography, yielding six independent sets of imaging data per patient (T2w, DWI, apparent diffusion coefficient, MRE magnitude, shear wave speed, and loss angle maps). Combinations of these data were used to train Dense U-nets with manually segmented masks of the entire prostate gland (PG), central zone (CZ), and peripheral zone (PZ) in 30 patients and to validate them in 10 patients. Dice score (DS), sensitivity, specificity, and Hausdorff distance were determined. We found that segmentation based on MRE magnitude maps alone (DS, PG:  $0.93 \pm 0.04$ , CZ:  $0.95 \pm 0.03$ , PZ:  $0.77 \pm 0.05$ ) was more accurate than magnitude maps combined with T2w and DWI\_b (DS, PG:  $0.91 \pm 0.04$ , CZ:  $0.91 \pm 0.06$ , PZ:  $0.63 \pm 0.16$ ) or T2w alone (DS, PG:  $0.92 \pm 0.03$ , CZ:  $0.91 \pm 0.04$ , PZ:  $0.65 \pm 0.08$ ). Automatically tabulated MRE values were not different from ground-truth values ( $P > 0.05$ ). In conclusion, MRE combined with Dense U-net segmentation allows tabulation of quantitative imaging markers without manual analysis and independent of other MRI sequences and can thus contribute to PCa detection and classification.

Prostate cancer (PCa) has the highest incidence of all types of cancer in men and is the second leading cause of cancer deaths in men<sup>1,2</sup>. Radiological imaging modalities such as magnetic resonance imaging (MRI) play a central role in the diagnosis of PCa and therapy planning. In particular, multiparametric MRI as defined in the Prostate Imaging Reporting and Data System (PI-RADS) has contributed to the standardization of prostate MRI worldwide<sup>3</sup>. However, inter- and intrareader agreement is still moderate<sup>4</sup>. This is mainly due to the use of subjective imaging criteria in the assessment of lesion shape and signal intensity as well as manual segmentation procedures of prostate regions, resulting in high variability of reference values. It is noteworthy that even quantitative imaging parameters can vary widely with the size and site of the regions selected for their analysis<sup>5</sup>. Therefore, accurate segmentation of the prostate gland and its zones is critical for the detection and management of PCa<sup>6–9</sup>. Automated prostate segmentation using multiparametric MRI (mpMRI) is often performed based on morphologic images such as T2-weighted MRI and transferred to other images which depict microstructural information such as the apparent diffusion coefficient (ADC) measured by diffusion-weighted imaging (DWI). Complementary to DWI, magnetic resonance elastography (MRE)<sup>10</sup> has been recently introduced for the clinical assessment of PCa<sup>11–22</sup>. MRE provides maps of stiffness and viscosity, which are quantitatively linked with

<sup>1</sup>Department of Radiology, Charité – Universitätsmedizin Berlin, Berlin, Germany. <sup>2</sup>DKTK (German Cancer Consortium), Partner Site Berlin, Berlin, Germany. <sup>3</sup>Berlin Institute of Health at Charité, Universitätsmedizin Berlin, Berlin, Germany. <sup>4</sup>Institute of Computer-assisted Cardiovascular Medicine, Charité – Universitätsmedizin Berlin, Berlin, Germany. ✉email: ingolf.sack@charite.de



**Figure 1.** Images illustrating the appearance of the prostate in mpMRI in a volunteer with benign prostatic hyperplasia (BPH). **(a)** T2-weighted MRI, **(b)** DWI image with a b-value of 1400 s/mm<sup>2</sup>, **(c)** ADC map, **(d)** MRE magnitude signal, **(e)** SWS map as a surrogate marker of tissue stiffness, and **(f)**  $\varphi$  map, which signifies friction or tissue fluidity.

mechanical microstructures in biological soft tissues and their changes due to disease<sup>23,24</sup>. Unlike DWI, where ADC is reconstructed from variations in signal magnitude, MRE values are reconstructed from phase images, while magnitude images have been largely unused except for recent approaches to water diffusion analysis<sup>25,26</sup>. Our hypothesis was that MRE magnitude images provide anatomical information that can be used for automated segmentation of prostate zones. Full exploitation of anatomic and viscoelastic information contained in a single set of MRE data would greatly facilitate quantitative parameter extraction without co-registration artifacts. However, for several reasons, automated prostate segmentation is a challenging task<sup>27</sup>. For example, the prostate is a highly heterogeneous organ with complex 3D geometry giving rise to well-known ambiguities of tissue boundaries in MRI. Furthermore, prostate morphology greatly varies among individuals, especially when benign prostatic hyperplasia (BPH) or advanced PCa is present. Figure 1 illustrates these challenges in a representative case of BPH. Numerous studies have tackled the problem of automatic segmentation of MR images of the prostate using various approaches such as atlas segmentation<sup>28</sup>, deformable and statistical modeling<sup>29</sup> or machine learning<sup>30,31</sup>. For a few years, deep convolutional neural networks (CNNs) have been extensively used for segmentation tasks in various radiological applications, thanks to their suitability for generalization<sup>32</sup>. In the prostate, CNNs were used for prostate gland segmentation using either slice-wise<sup>33</sup> or full 3D approaches<sup>34,35</sup>. However, only a few approaches were actually able to segment the prostate gland (PG) and its zones such as the central zone (CZ) and the peripheral zone (PZ)<sup>36,37</sup>. CNNs, such as<sup>38,39</sup>, are mathematical models that can learn a number of complex features such as intensities, texture or morphology presented in the input images and, thus, are suitable to classify different tissue types to accomplish fully automated segmentation tasks.

In this study, we used CNNs for fully automated segmentation of the prostate zones based on T2-weighted (T2w) MRI, DWI, and, the so far unused, MR Elastography images<sup>40</sup>. DWI provides intensity images (based on a specific b-value, DWI<sub>b</sub>) and quantitative ADC maps. MRE provides magnitude images (mag), shear wave speed maps of stiffness (SWS, in m/s), and maps of loss angle ( $\varphi$  in radians, indicating friction or a material's fluidity)<sup>41</sup>. We explored CNNs for prostate segmentation using 14 possible input combinations of MRE maps as input training data. First, we trained and tested individual networks (models), henceforth termed individual models (IMs), using each of the aforementioned sets with their corresponding manually segmented masks of CZ, PZ, and PG as ground truths. Second, the dataset was rearranged so that it contains all 14 possible input combinations at once to be used as input training data for a single model, henceforth termed unified model (UM). The purpose of this study was to test whether MRE-based tomoelastography data are sufficient to extract anatomic and viscoelastic information for automatic tabulation of zonal mechanical properties of the prostate and to compare MRE-based image segmentation with information extracted by CNNs from other MRI pulse sequences typically acquired in clinical practice.

Map	Axial slices (number)	Slice thickness (mm)	Field of view (mm <sup>2</sup> )	Matrix size (pixels)	Voxel size (mm <sup>3</sup> )	Repetition time (TR) (ms)	Echo time (TE) (ms)	Parallel imaging factor (GRAPPA-algorithm)
MRE	25	2	256 × 256	128 × 128	2 × 2 × 2	3240	69	2
T2w	25	3	180 × 180	384 × 384	0.47 × 0.47 × 3	4000	116	2
DWI	25	3	230 × 230	160 × 160	1.44 × 1.44 × 3	4000	56	2

**Table 1.** Imaging parameters of MRE and MRI.

## Methods

**Subjects.** The imaging data used in this study were used retrospectively and were acquired in a previously reported population of BPH and PCa patients who underwent PI-RADS-compatible mpMRI and MRE<sup>12</sup>. Our local ethical review board approved this study, and all patients gave written in-formed consent. Forty patients were included, 26 with a PI-RADS score of 2 consistent with BPH. The remaining fourteen men received biopsy with subsequent Gleason scoring. The decision to perform a biopsy was made by the referring urologist based on the clinical assessment and the PIRADS score. Those fourteen men had a suspicious focal lesion (PI-RADS 4: n = 2, maximum tumor diameters of 10 and 11 mm; PI-RADS 5: n = 12, maximum tumor diameters ranging from 16 to 66 mm). In these 14 men, prostate biopsy revealed PCa Gleason scores of 3 + 3 (ISUP class 1, n = 1), 3 + 4 (ISUP class 2, n = 2), 4 + 3 (ISUP class 3, n = 2), 4 + 4 (ISUP class 4, n = 4), and ≥ 4 + 5 (ISUP class 5, n = 5).

From the 40 volume datasets of these patients, 30 were randomly selected for training. Each volume dataset included 25 slices, yielding a total of 25 (slices) × 30 (volumes) × 1 (input combination) = 750 training images for each of the 14 input combinations of IMs, and 25 (slices) × 30 (volumes) × 14 (all input combinations) = 10,500 images for UM. In the latter case, all 14 input combinations were included in a single large training set. The remaining 10 datasets with a total of 25 (slices) × 10 (volumes) = 250 slices were used for testing.

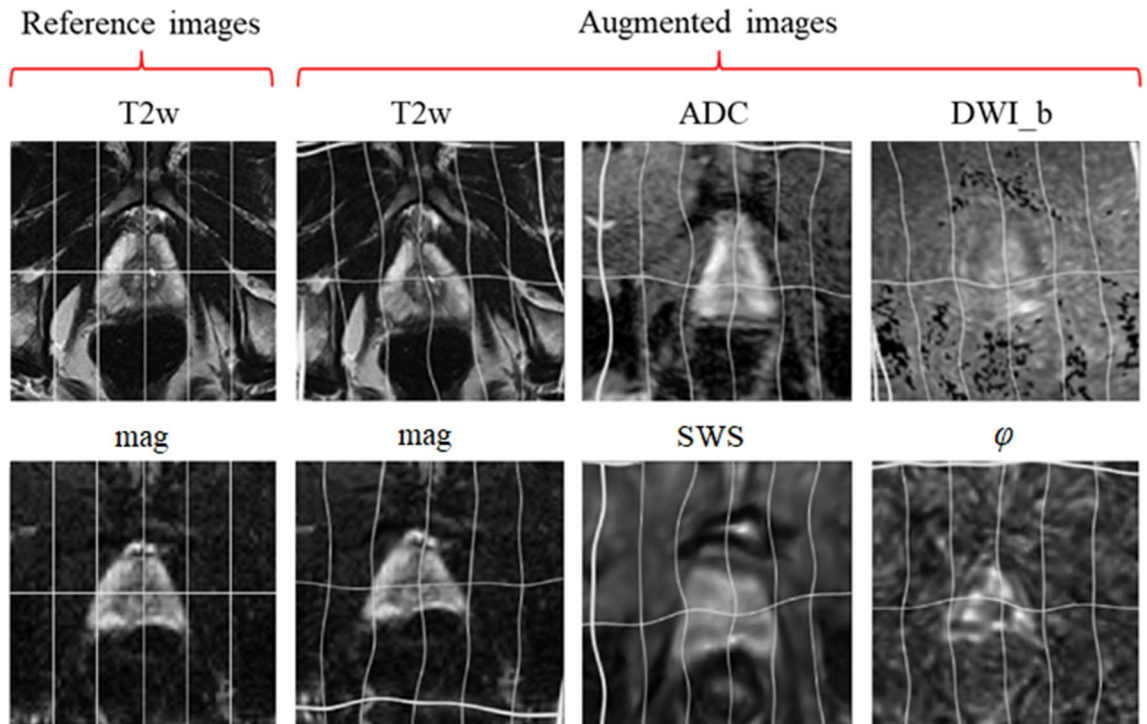
**MR Imaging.** The imaging data were acquired on a 3-Tesla MRI scanner (Magnetom Skyra; Siemens Healthineers, Erlangen, Germany) using both the 18-channel phased-array surface coil and a spine array coil. All patients underwent a clinically indicated mpMRI examination of the prostate in accordance with PI-RADS version 2 (2015)<sup>42</sup> which included T2-weighted sequences (T2w) in the axial and coronal planes and DWI. Here we used b-values of 0, 50, 500, 1000, and 1400 s/mm<sup>2</sup> (b = 1400 s/mm<sup>2</sup> was used as diffusion weighted image, henceforth referred to as DWI\_b). ADC maps were automatically generated by the MRI scanner via mono-exponential fitting of all b-values. After the clinical MRI examination, patients underwent multifrequency MRE-based tomo-elastography using a single-shot spin-echo sequence with three excitation frequencies of 60, 70, 80 Hz<sup>12</sup>. MRE magnitude images (mag) display signal intensities which are T2-weighted with a strong T2\* effect. All imaging parameters are summarized in Table 1.

MRE data processing was based on multifrequency wavefield inversion for generating frequency-compounded SWS and  $\varphi$  maps. SWS was reconstructed by single-derivative finite-difference operators (k-MDEV)<sup>43</sup> while  $\varphi$  was obtained by second-order, Laplacian-based, direct inversion (MDEV)<sup>44</sup>. Although k-MDEV is noise-resistant and well-suited for SWS visualization, it is limited regarding the quantification of viscosity-related parameters such as  $\varphi$ <sup>43</sup>. Hence, MDEV was used for  $\varphi$  recovery. Both k-MDEV and MDEV pipelines are publicly available at bioqic-apps.charite.de<sup>45</sup>. Two experienced radiologists (PA with more than 10 years and FB with 2 years of experience in PIRADS-based PCa detection and classification) segmented and revised the prostate and its zones based on T2w-MRI, DWI, and MRE. The resulting masks of CZ (which includes the transition zone), PZ, and PG were used as ground truth (GT) for training and validation of CNNs. Overall, six independent imaging data were further used for segmentation analysis: T2w, DWI\_b, ADC, mag, SWS, and  $\varphi$ .

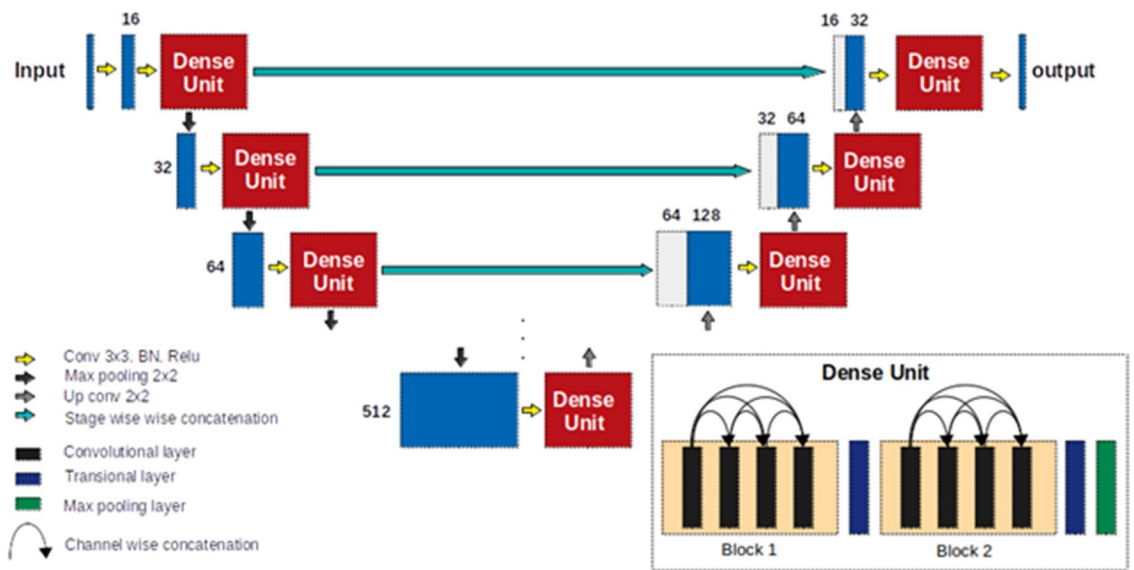
**Image preparation and augmentation.** Since T2w, DWI, and MRE images had different resolutions, all images were resampled to a common resolution of 0.5 mm isotropic edge length. Images of the same size and resolution were obtained by positioning a cropping window with a size of 256 × 256 pixels at the center of each 3D imaging volume. For image augmentation, 9 random elastic deformations were applied to the original images to increase the number of training sets. As described in<sup>46</sup>, elastic deformation can be driven by two main parameters:  $\sigma$ , which represents the elasticity coefficient, and  $\alpha$ , which represents a scaling factor that controls the amplitude of the deformation. The two parameters were set to  $\alpha = 21$  and  $\sigma = 512$ . Examples of augmented images are provided in Fig. 2.

**Dense U-net.** A previously developed Dense U-net<sup>40</sup> based on state-of-the-art U-net convolutional network architecture<sup>38</sup> was used. The network comprised two main parts—an encoder and a decoder—between which skip connections connected feature maps with similar resolutions. Normal stacks of convolutional layers at each stage were substituted with two densely connected blocks. Each dense block consisted of four convolutional layers with 3 × 3 kernel size followed by a transitional layer. The structure of the network is illustrated in Fig. 3.

**Network training.** Stochastic gradient descent with a learning rate of 10<sup>-3</sup>, a momentum of 0.9, and a decay of 10<sup>-6</sup> was used in this study to train and test all proposed models. We used cross-entropy (CE) loss as the main loss function, and performed pixel-wise comparison of ground-truth images and the resulting masks.

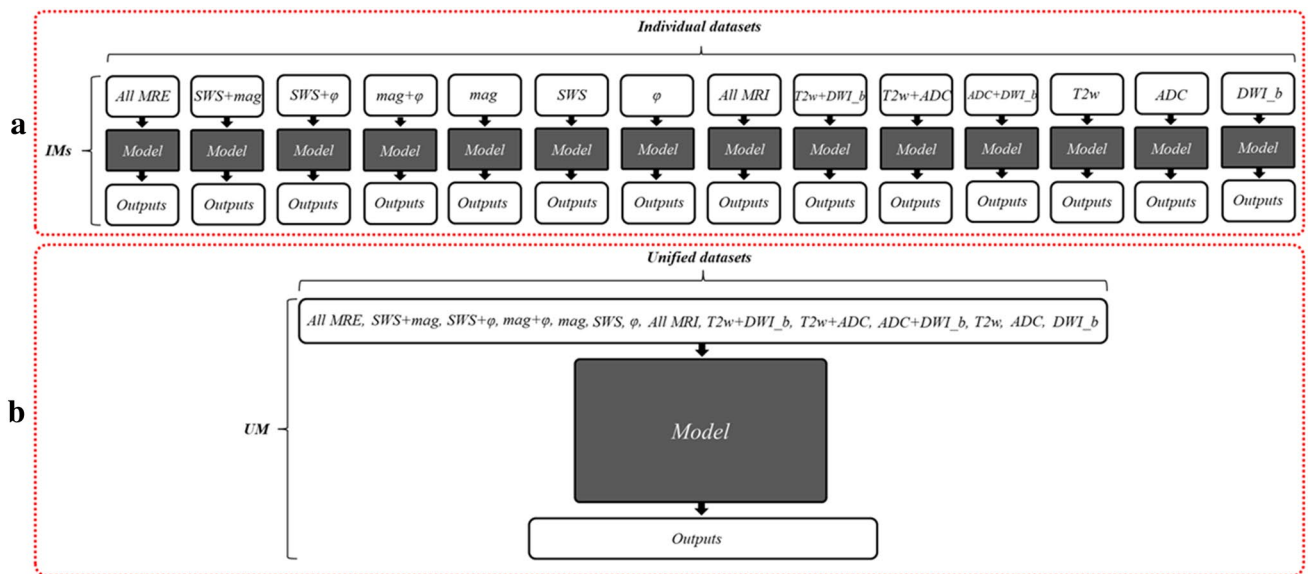


**Figure 2.** Image augmentation using elastic deformation. First column shows the reference non-augmented T2w images in the first row and MRE magnitude maps (mag) in the second row. Augmented images are shown in the remaining columns: first row: T2w, ADC, and DWI\_b. Second row: mag, SWS and  $\varphi$ .



**Figure 3.** Dense U-net architecture. Numbers in the figure indicate the number of feature maps at each stage.

Two main approaches were tested: (1) Individual models (IMs), where we trained and tested a separate model for each combination of sequences/maps or individual sequence/map input. (2) Unified model (UM) without MRE-MRI co-registration (i.e., a single model for T2w, DWI\_b, ADC, mag, SWS, and  $\varphi$  images, trained on masks that were manually and separately segmented from T2w and mag images), yielding a re-arranged dataset, where all image combinations were taken into account during training, validation and testing. Fourteen input combinations were used for training and testing IMs and the UM, more specifically mag+SWS+ $\varphi$ , SWS+mag, SWS+ $\varphi$ , mag+ $\varphi$ , mag,  $\varphi$ , SWS, T2w+ADC+DWI\_b, T2w+ADC, T2w+ADC, ADC+DWI\_b, T2w, ADC and DWI\_b. While each of these input combinations was used to train an individual model in IMs, in UM, each input combination represented a subset and was combined with all other subsets to generate a single, large dataset of all 14 input combinations for training and testing the UM. Figure 4 shows how imaging data were used as inputs for IMs and the UM.



**Figure 4.** Diagrams of the two approaches investigated in this study. **(a)** individual models (IMs), and **(b)** unified model (UM). The figure shows all 14 input combinations used individually as separate datasets in IMs and combined into a single large dataset for training of the UM. The gray box with the ‘Model’ refers to the Dense U-net used in this study, as detailed in Fig. 3.

**Evaluation.** We evaluated all resulting segmentations against manually delineated ground-truth masks using standard evaluation statistics such as mean dice score (DS)  $\pm$  standard deviation (SD), sensitivity (Sen), specificity (Sp), and Hausdorff distance (HD) as a contour consistency measure, and a t-test with  $p < 0.05$  indicating a statistically significant difference.

The dice score is a similarity measure that quantifies the overlap between predicted masks and ground-truth labels, allowing straightforward comparison of segmentation performance<sup>47</sup>.

The Hausdorff distance is the maximum distance between two edge points from two different sets (predicted mask and ground truth). It is expressed in millimeters (mm) and was defined as follows:

$$HD(A, B) = \max\{\max_{i \in A} \min_{j \in B} d(i, j), \max_{i \in B} \min_{j \in A} d(i, j)\}$$

where  $d(i, j)$  is the Euclidean distance between two points from different sets  $A$  and  $B$ .

**Implementation details.** All models were trained on individual or combinations of MRI/MRE images in a slice-wise fashion, where all images had a size of 256x256 and an in-plane resolution of 0.5x0.5 mm. We used the SimpleITK library for image preprocessing<sup>48,49</sup>, and Keras library (version 2.4.3) with Tensorflow library (version 2.3.1 Google Inc.) back-end<sup>50</sup> as the main library for model implementation, training, and testing.

All models were trained on a TitanXP GPU with 2 GB video memory (CUDA version of 10.1) and a batch size of 25 images. Training time of the Dense U-net was around 8.5 hours. The computation time during testing for a single 3D volume (of around 25 slices) was approximately 1.5 s.

**Ethical approval.** All the procedures and experimental protocols were confirmed and approved by Ethics Committee of Charité – Universitätsmedizin. All methods were carried out in accordance with relevant guidelines and regulations.

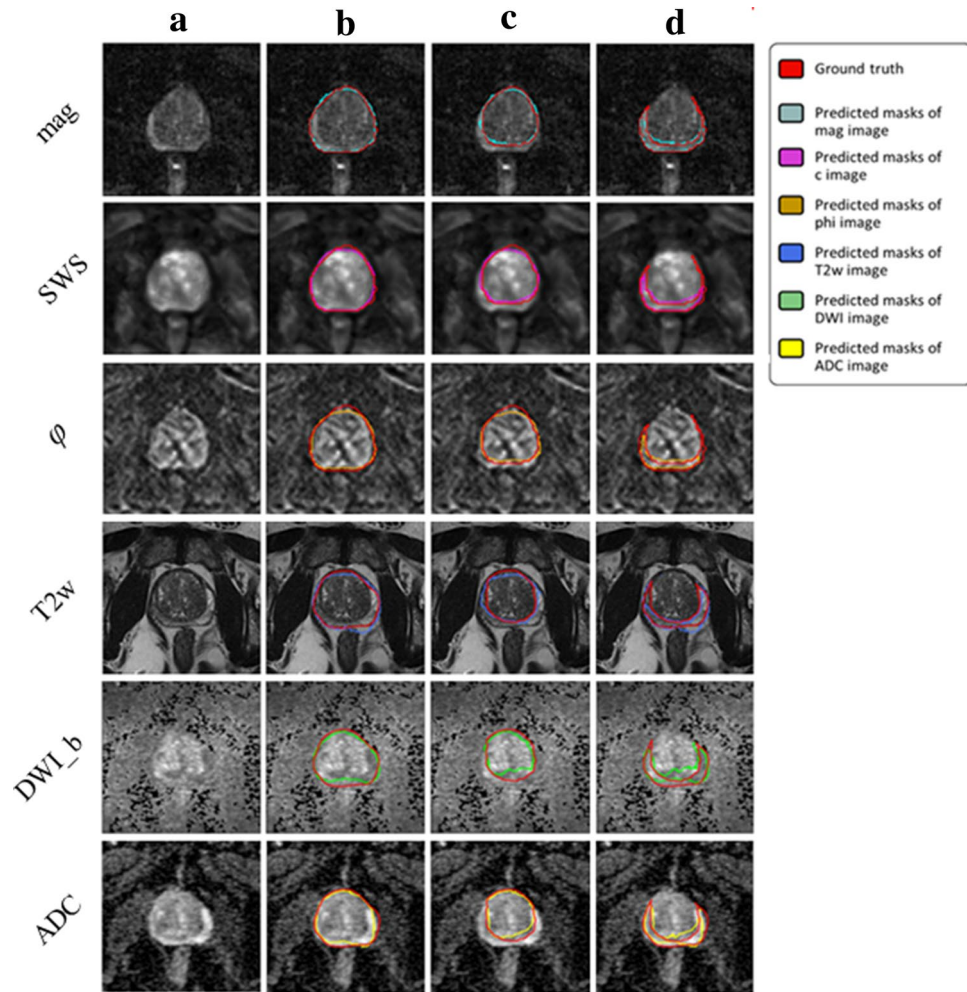
## Results

This section presents the performance results of our Dense U-net that was used on MRI and MRE images. Dense U-net was applied in two forms: UM and IMs evaluated using DS, HD and MRE values of the automatically segmented prostate zones.

Figures 5 and 6 illustrate segmentation results for each of the two training approaches. The boundary of automatically segmented regions closely follows the boundary of manual segmentations.

Our experimental results for IMs trained on different combinations of maps showed that the proposed method was capable of efficiently segmenting PG, CZ, and PZ from the input MRI/MRE images. DS ranged from  $0.87 \pm 0.04$  (ADC) to  $0.93 \pm 0.04$  (mag), from  $0.85 \pm 0.09$  (DWI\_b) to  $0.95 \pm 0.03$  (mag), and from  $0.53 \pm 0.10$  (ADC+DWI\_b) to  $0.77 \pm 0.05$  (mag), for PG, CZ, and PZ, respectively. Also, HD was minimized (i.e., best contour consistency) based on mag with 0.86, 0.76, 0.98 mm (Fig. 5 and Table 2). However, performance differences between MRE (mag, SWS,  $\phi$ ) and MRI maps (T2w, DWI\_b, ADC) were statistically not significant except for CZ, where MRE (mag, SWS,  $\phi$ ) had a significantly higher DS than MRI (T2w, DWI\_b, ADC,  $p < 0.05$ ).

Unlike IMs, the UM can process any map or combination of maps, which facilitates clinical applications. Compared with IMs, the UM had higher DS in PG and CZ and lower DS in PZ. Assessed for different prostate



**Figure 5.** Examples of segmentation results achieved with IMs: first column (a) shows the original image, second (b), third (c) and fourth column (d) show masks of prostate, central and peripheral zones, respectively.

zones, DS of the UM ranged from  $0.77 \pm 0.11$  (DWI\_b) to  $0.92 \pm 0.04$  (mag, SWS,  $\varphi$ ), from  $0.65 \pm 0.08$  (DWI\_b) to  $0.86 \pm 0.06$  (mag, SWS,  $\varphi$ ), and from  $0.28 \pm 0.10$  (DWI\_b) to  $0.57 \pm 0.05$  (mag), for PG, CZ, and PZ, respectively. The smallest HD of 1.15, 1.45, and 1.81 for PG, CZ, and PZ, respectively, was found for mag. Images are presented in Fig. 6, and the results are summarized in Table 3.

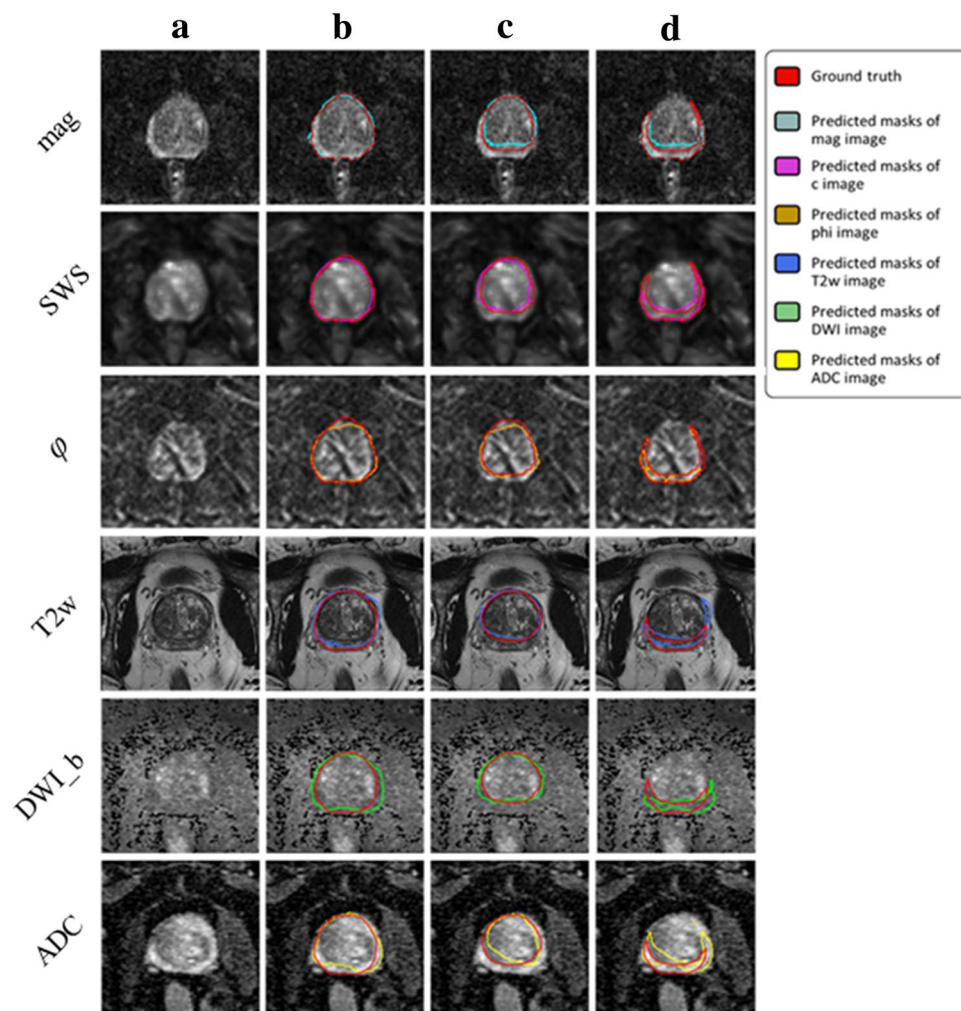
Unlike IMs, the UM showed a significantly higher DS when using MRE data in comparison with MRI data ( $p < 0.001$ , 0.05, and 0.05 for PG, CZ, and PZ, respectively). Overall, IMs had significantly more accurate results compared with the UM in terms of both DS and HD with  $p$ -values  $< 0.01$  for PG, CZ, and PZ. This is shown in Fig. 7, illustrating that DS was higher and HD lower for IMs than for the UM. Figure 8 shows a case where model segmentations were inaccurate compared with the ground-truth masks. Quantitative analysis, for both IMs and the UM, of pixel values in PG, CZ, and PZ showed no significant difference ( $p > 0.05$ ) between ground-truth and automated prostate segmentation. Group mean values are presented in Tables 4 and 5.

## Discussion

Our study shows that Dense U-net segmentation of prostate zones based on MRE data allows automated tabulation of quantitative imaging markers for the total prostate and for the central and peripheral zones.

Our results show that IMs performed excellent across all maps and sequences with high values for DS and low values for HD. In our experiments, segmentation was most reliable when we used T2w and MRE magnitude images, which provide sufficiently rich anatomical details for automated prostate segmentation while quantitative parameter maps such as SWS,  $\varphi$  and ADC lack those details. Figure 9 depicts a variety of maps in a patient demonstrating that anatomy of prostate boundaries is well preserved on T2w and MRE magnitude images while it is less clearly visible on ADC, DWI\_b, SWS and  $\varphi$  maps.

For IMs, we found no significant difference between DS of MRE and DS of MRI in PG and PZ while, in CZ, DS of MRE was higher than that of MRI. This may be explained in part by blurring due to larger slice thickness in MRI (3 mm) than MRE (2 mm). Furthermore, MRI slice volumes covered the entire prostate gland, the seminal vesicles and the periprostatic tissues while MRE volumes were solely focused on the prostate gland given the smaller slice thickness.



**Figure 6.** Examples of segmentation results achieved with the UM: first column (a) shows the original image, second (b), third (c) and forth column (d) show masks of prostate, central and peripheral zones, respectively.

Model	Prostate gland						Central zone						Peripheral zone					
	Dice ↑	Std ↓	Median ↑	Sen ↑	Spc ↑	HD ↓	Dice ↑	Std ↓	Median ↑	Sen ↑	Spc ↑	HD ↓	Dice ↑	Std ↓	Median ↑	Sen ↑	Spc ↑	HD ↓
All MRE	0.92	0.04	0.91	0.92	0.99	1.04	0.94	0.03	0.93	0.93	0.99	1.08	0.73	0.05	0.74	0.72	1.00	1.35
SWS+mag	0.91	0.04	0.90	0.90	0.99	1.22	0.93	0.04	0.92	0.91	0.99	1.41	0.69	0.04	0.69	0.65	1.00	1.72
SWS+φ	0.91	0.05	0.90	0.89	0.99	1.12	0.92	0.05	0.93	0.92	0.99	1.10	0.64	0.10	0.62	0.68	1.00	1.43
Mag+φ	0.91	0.04	0.91	0.94	0.99	1.95	0.93	0.04	0.93	0.95	0.99	1.92	0.69	0.06	0.68	0.68	1.00	1.97
Mag	<b>0.93</b>	0.04	0.93	0.93	0.99	<b>0.86</b>	<b>0.95</b>	0.03	0.95	0.96	0.99	<b>0.76</b>	<b>0.77</b>	0.05	0.79	0.79	1.00	<b>0.98</b>
φ	0.90	0.05	0.88	0.87	0.99	1.18	0.91	0.05	0.92	0.89	0.99	1.19	0.52	0.11	0.54	0.57	1.00	1.66
SWS	0.82	0.08	0.80	0.78	0.99	2.41	0.88	0.06	0.87	0.84	0.99	2.56	0.55	0.07	0.53	0.59	1.00	2.64
All MRI	0.90	0.02	0.88	0.83	1.00	1.06	0.88	0.05	0.86	0.82	1.00	1.16	0.66	0.11	0.67	0.61	1.00	1.81
T2w+ADC	0.90	0.04	0.89	0.88	0.99	1.30	0.87	0.06	0.88	0.84	1.00	1.35	0.62	0.13	0.66	0.65	0.99	2.82
T2w+DWI_b	0.91	0.04	0.90	0.90	1.00	0.98	0.91	0.06	0.90	0.86	1.00	1.00	0.63	0.16	0.66	0.68	0.99	2.11
ADC+DWI_b	0.88	0.05	0.87	0.84	0.99	1.44	0.87	0.06	0.86	0.83	0.99	1.46	0.53	0.10	0.55	0.52	0.99	2.84
T2w	0.92	0.03	0.92	0.93	0.99	1.05	0.91	0.04	0.91	0.89	1.00	0.87	0.65	0.08	0.65	0.72	0.99	2.90
ADC	0.87	0.04	0.86	0.86	0.99	1.41	0.86	0.06	0.84	0.84	0.99	1.41	0.55	0.08	0.57	0.55	0.99	2.38
DWI_b	0.88	0.06	0.86	0.87	0.99	1.38	0.85	0.09	0.83	0.82	1.00	1.48	0.54	0.17	0.51	0.61	0.99	2.07
All MRE ave.	0.91	0.04	0.90	0.91	0.99	1.23	0.93	0.04	0.93	0.93	0.99	1.24	0.67	0.07	0.67	0.68	1.01	1.52
All MRI ave.	0.89	0.04	0.88	0.87	0.99	1.23	0.88	0.06	0.87	0.84	1.00	1.25	0.60	0.12	0.61	0.62	0.99	2.42

**Table 2.** Statistical analysis of the segmentation results of IMs. Bold indicates the highest values, while *Italic* indicates the lowest values. ↑ (↓): the higher (the lower) the better.

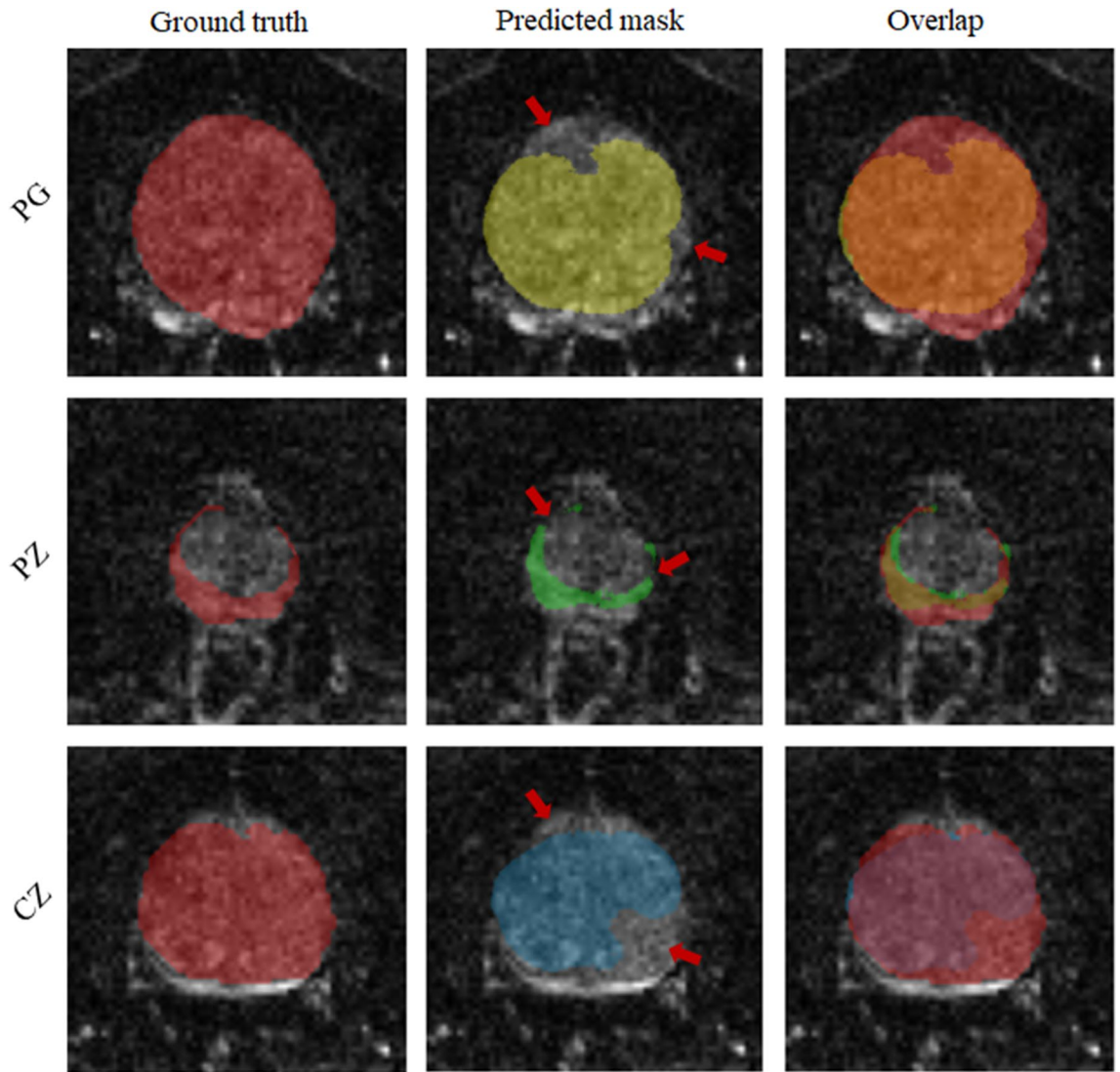
Model	Prostate gland						Central zone						Peripheral zone					
	Dice ↑	Std ↓	Median ↑	Sen ↑	Spc ↑	HD ↓	Dice ↑	Std ↓	Median ↑	Sen ↑	Spc ↑	HD ↓	Dice ↑	Std ↓	Median ↑	Sen ↑	Spc ↑	HD ↓
All MRE	<b>0.92</b>	0.04	0.92	0.87	1.00	1.17	0.86	0.06	0.82	0.80	0.99	1.61	<b>0.56</b>	0.05	0.56	0.54	0.99	2.00
SWS+mag	0.91	0.04	0.91	0.87	1.00	1.19	0.85	0.06	0.81	0.78	0.99	1.64	0.55	0.05	0.56	0.54	0.99	2.01
SWS+φ	0.90	0.04	0.89	0.87	0.99	1.41	0.85	0.05	0.83	0.79	0.99	1.93	0.49	0.06	0.52	0.53	0.99	2.16
Mag+φ	0.91	0.04	0.91	0.88	0.99	1.23	0.87	0.05	0.85	0.84	0.99	1.66	0.57	0.03	0.57	0.55	0.99	1.86
Mag	0.91	0.05	0.91	0.87	0.99	<b>1.15</b>	<b>0.87</b>	0.05	0.84	0.84	0.99	<b>1.45</b>	0.57	0.05	0.57	0.52	0.99	<b>1.81</b>
φ	0.88	0.04	0.88	0.85	0.99	1.54	0.83	0.04	0.82	0.76	0.99	2.03	0.44	0.07	0.47	0.49	0.98	2.43
SWS	0.89	0.05	0.89	0.87	0.99	1.63	0.82	0.08	0.80	0.76	0.99	2.18	0.46	0.06	0.47	0.48	0.99	2.19
All MRI	0.80	0.09	0.78	0.91	0.98	2.37	0.73	0.06	0.71	0.56	1.00	2.80	0.51	0.09	0.53	0.56	0.99	3.42
T2w+ADC	0.80	0.09	0.78	0.91	0.98	2.34	0.72	0.05	0.70	0.55	1.00	2.86	0.49	0.09	0.52	0.55	0.99	3.56
T2w+DWI_b	0.81	0.09	0.80	0.90	0.98	2.04	0.74	0.06	0.72	0.58	1.00	2.63	0.40	0.08	0.41	0.41	0.99	3.82
ADC+DWI_b	0.80	0.09	0.75	0.88	0.98	2.25	0.670	<i>.07</i>	0.64	0.48	1.00	3.45	0.45	0.11	0.44	0.53	0.99	4.10
T2w	0.78	0.09	0.78	0.87	0.98	2.43	0.75	0.07	0.72	0.60	1.00	2.35	0.42	0.12	0.43	0.38	0.99	4.12
ADC	0.80	0.08	0.76	0.88	0.98	2.21	0.65	0.07	0.62	0.46	1.00	3.54	0.43	0.11	0.43	0.52	0.98	4.19
DWI_b	0.77	0.11	0.76	0.88	0.98	2.42	0.65	0.08	0.64	0.49	1.00	3.38	0.28	0.10	0.30	0.33	0.98	5.29
All MRE ave.	0.89	0.05	0.89	0.87	0.99	1.46	0.83	0.06	0.81	0.77	0.99	1.91	0.52	0.06	0.53	0.53	0.99	2.23
All MRI ave.	0.79	0.09	0.77	0.89	0.98	2.29	0.70	0.07	0.68	0.53	1.00	3.00	0.43	0.10	0.44	0.47	0.99	4.07

**Table 3.** Statistical analysis of the segmentation results of the UM. Bold indicates the highest values, while italic indicates the lowest values. ↑(↓): the higher (the lower) the better.



**Figure 7.** Comparison of the performance of the two approaches: IMs with orange and UM with blue color. The left and right figures show the values of dice score and Hausdorff distance, respectively.





**Figure 8.** Illustration of inaccurate segmentations: left column is the ground truth; middle is predicted masks and right is the overlap between ground truth and predicted masks. Rows from top to bottom, images of PG, PZ and CZ respectively.

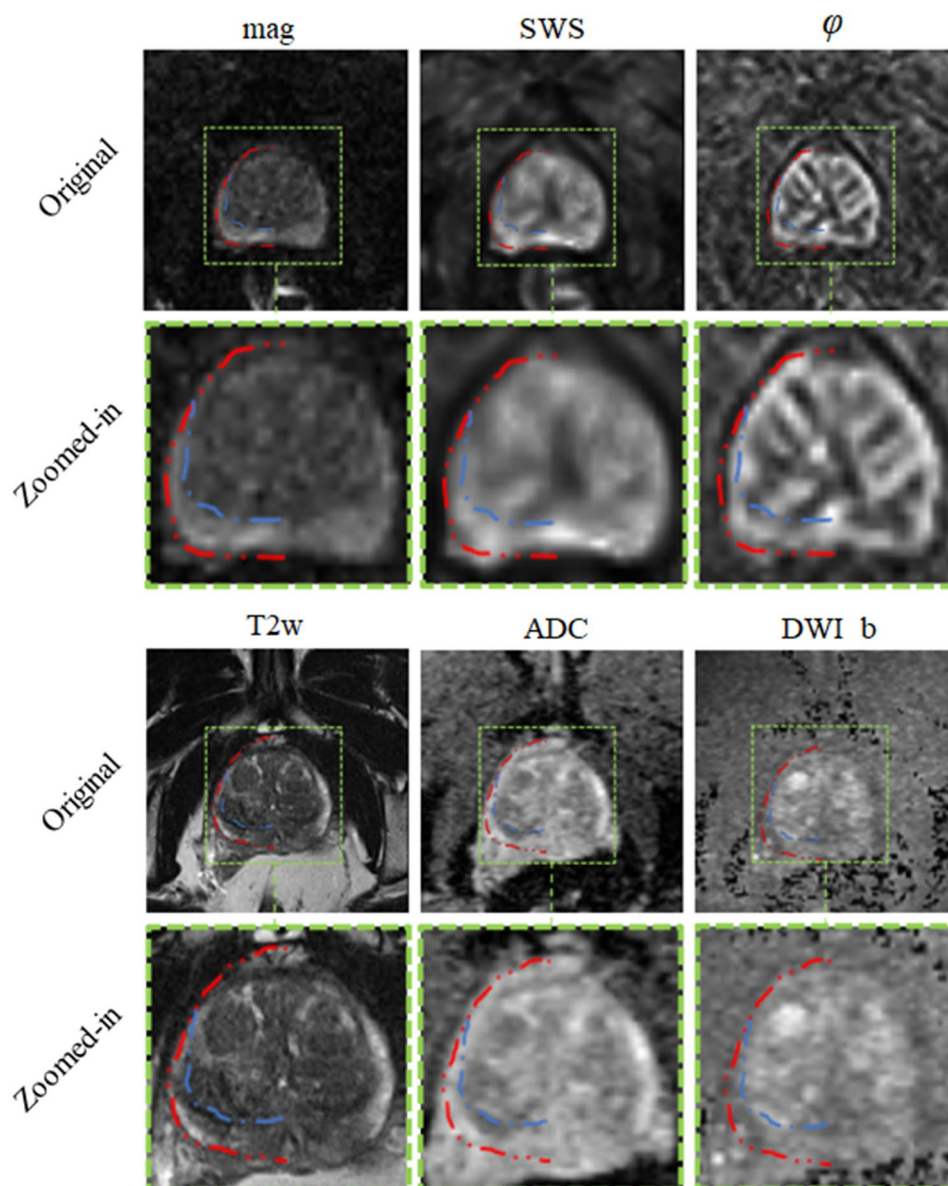
Map	Pixel values (average $\pm$ standard deviation)					T-test (significance at $P < 0.05$ )			
	PG mask	PG predicted	CZ mask	CZ predicted	PZ mask	PZ predicted	PG P-value	CZ P-value	PZ P-value
SWS	1.28 $\pm$ 0.39	1.35 $\pm$ 0.4	1.25 $\pm$ 0.32	1.36 $\pm$ 0.33	1.39 $\pm$ 0.22	1.28 $\pm$ 0.22	0.87	0.81	0.83
mag	33.1 $\pm$ 10.46	36.56 $\pm$ 11.05	29.6 $\pm$ 7.85	34.03 $\pm$ 8.31	43.21 $\pm$ 7.2	39.7 $\pm$ 7.29	0.83	0.78	0.85
$\varphi$	0.61 $\pm$ 0.19	0.63 $\pm$ 0.19	0.64 $\pm$ 0.17	0.66 $\pm$ 0.17	0.58 $\pm$ 0.09	0.56 $\pm$ 0.09	0.91	0.88	0.87

**Table 4.** Values obtained in PG, CZ and PZ regions of interest using ground truth and predicted masks. “mask” and “predicted” in the column headers refer to the ground truth and model output, respectively. T-test results between ground truth and predicted masks for PG, CZ, and PZ are reported in the right part of the table (based on segmentation results for the IMs using magnitude images as input).

In contrast to IMs, the UM could process any input combination of MRI/MRE maps without a need for retraining or fine-tuning the network. Similar to IMs, the UM also favored input combinations with rich anatomical detail, such as T2w and magnitude images. In all experiments, HD of the UM ranged from 1.15 to 5.29 mm, which is significantly inferior to IMs. Moreover, by showing the changes in DS and HD due to all possible input combinations for IMs and UM, Fig. 7 illustrates that both IMs and the UM had decent performance while IMs were slightly better. Nevertheless, given the robustness of the UM combined with decently good segmentation

Map	Pixel values (average $\pm$ standard deviation)					T-test (significance at $P < 0.05$ )			
	PG mask	PG predicted	CZ mask	CZ predicted	PZ mask	PZ predicted	PG P-value	CZ P-value	PZ P-value
SWS	1.28 $\pm$ 0.39	1.29 $\pm$ 0.39	1.25 $\pm$ 0.32	1.32 $\pm$ 0.3	1.39 $\pm$ 0.22	1.27 $\pm$ 0.24	0.9	0.89	0.81
mag	33.1 $\pm$ 10.46	34.1 $\pm$ 10.6	29.6 $\pm$ 7.85	30.54 $\pm$ 7.0	43.21 $\pm$ 7.2	39.6 $\pm$ 8.05	0.95	0.95	0.85
$\varphi$	0.61 $\pm$ 0.19	.62 $\pm$ 0.19	0.64 $\pm$ 0.17	0.66 $\pm$ 0.16	0.58 $\pm$ 0.09	0.57 $\pm$ 0.11	0.99	0.92	0.95

**Table 5.** Values obtained in PG, CZ and PZ regions of interest using ground truth original and predicted masks. “mask” and “predicted” in the column headers refer to the ground truth and model output, respectively. T-test results between ground truth and predicted masks for PG, CZ and PZ are also reported in the right part of the table (based on segmentation results from UM and input images of mag).



**Figure 9.** Different maps of the same patient. Second and fourth rows show zoomed-in images of the first and third rows, respectively. Dotted red and blue lines highlight part of the segmentation for prostate and PZ, respectively and the rest of the delineation was omitted to show the actual pixel intensities.

results, it is recommended to primarily apply IMs and to use the UM as a second opinion for automated prostate segmentation.

Exploiting MRE magnitude images for automatic segmentation instead of high-resolution in-plane T2-weighted MRI had the benefit of not requiring image registration and making full use of the anatomic and viscoelastic information contained in a single MRE dataset. This potentially stabilizes quantitative parameter extraction as any co-registration artifacts are avoided. As the ultimate proof of valid segmentations, viscoelasticity values averaged within volumes of PG, CZ, and PZ obtained from CNNs were not different from those obtained with manually segmented masks (see supplemental information). Thus, we here for the first time used the information of the magnitude signal in prostate MRE, showing that it was fully sufficient for accurate segmentation, which may greatly enhance MRE of the prostate in the future.

Figure 8 shows a case where the model failed to achieve accurate segmentation. Inaccuracies appear to be attributable to under-segmentation and discontinuity. Under-segmentation is visible in both the entire prostate gland (first row) and the CZ (last row), where the model did not properly locate and delineate boundaries. Discontinuity can be seen in the PZ (middle row), where the model resulted in a mask with several unconnected neighboring areas. Many factors can contribute to inaccurate segmentation, including boundary ambiguity, partial volume effects, and tissue heterogeneity. Therefore, radiologists typically use 3D information, which is subjectively interpolated by eye to the ambiguous image slice. However, even including adjacent slices for training in a 2.5-D approach<sup>51</sup> or use of full 3D models does not necessarily lead to better segmentation performance due to partial volume effects<sup>52</sup>.

We will implement the proposed CNN-based segmentation on a server, which is currently used for MRE data processing (<https://bioqic-apps.charite.de>), in order to make it available to the research community. Our ultimate aim is to accomplish fully automated tabulation of prostate SWS and  $\varphi$  values based on multifrequency MRE data. Once installed, the Dense U-net will be trained using other sets of MRE data acquired with other scanners and MRE sequences in order to generalize its applicability for prostate segmentation. In a next step, we aim at automated segmentation of other organs including the liver, kidneys, and pancreas based on MRE magnitude images.

Although MRE of the prostate is not yet part of standard care imaging, the proposed segmentation tool may help to further integrate this innovative imaging marker into routine clinical practice. We are confident that MRE will add diagnostic value to the PIRADS system in the future as shown in<sup>53</sup>. In addition, when enhanced with the segmentation option presented here, MRE may provide quantitative values on the mechanical consistency of prostate subzones as a reference for fully automated tumor classification.

Our study has limitations, including the small number of patients, which is attributable to fact that we performed a proof-of-concept exploration of CNNs for MRE-based prostate segmentation. For further improvement of segmentation quality and generalization to other domains, future studies should use multicenter MRI/MRE data acquired with different imaging protocols. Finally, we assembled all data for training in a way that avoids image alignment and registration procedures. While this approach ensured robust results based on single sets of data, we cannot rule out that combinations of MRI and MRE images (e.g., magnitude and T2w) would have resulted in slightly better DS and HD scores.

In summary, Magnitude images of prostate MRE were used for automated segmentation of prostate subzones based on trained CNNs. As such, MRE data provide all information needed for extraction of viscoelasticity parameters and for delineation of the prostate regions for which those values are of interest for tabulation and automated classification of suspicious prostate lesions. Dense U-net achieved excellent segmentation results using both IMs and the UM and yielded MRE parameters that were not different from ground truth. Compared with standard image segmentation based on T2w images, MRE magnitude images proved to suffice, as demonstrated by excellent Dice scores and Hausdorff dimension results. Quantitative maps of multiparametric MRI including those of DWI and viscoelasticity did not provide adequate anatomic information for learning-based prostate segmentation. Prostate MRE combined with Dense U-nets allows tabulating quantitative imaging markers without manual analysis and independent of other MRI sequences and thus has the potential to contribute to imaging-based PCa detection and classification.

## Data availability

All images in this study were used retrospectively and were acquired in a previously reported population of BPH and PCa patients who underwent PI-RADS-compatible mpMRI and MRE. Our local ethical review board approved this study (IRB approved study), and all patients gave written informed consent. The segmented masks were done in house.

Received: 26 July 2021; Accepted: 5 January 2022

Published online: 07 February 2022

## References

1. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics. *CA: Cancer J. Clin.* **66**, 7–30 (2016).
2. Aldoj, N., Lukas, S., Dewey, M. & Penzkofer, T. Semi-automatic classification of prostate cancer on multi-parametric MR imaging using a multi-channel 3D convolutional neural network. *Eur. Radiol.* **30**, 1243 (2019).
3. Turkbey, B. *et al.* Prostate imaging reporting and data system version 2.1: 2019 Update of prostate imaging reporting and data system version 2. *Eur. Urol.* **76**, 340–351 (2019).
4. Becker, A. S. *et al.* Direct comparison of PI-RADS version 2 and version 1 regarding interreader agreement and diagnostic accuracy for the detection of clinically significant prostate cancer. *Eur. J. Radiol.* **94**, 58–63 (2017).
5. Sack, I. & Schaeffter, T. *Quantification of Biophysical Parameters in Medical Imaging* (Springer, 2018).
6. Wang, Y. *et al.* Towards personalized statistical deformable model and hybrid point matching for robust MR-TRUS registration. *IEEE Trans. Med. Imaging* **35**, 589–604 (2016).

7. Terris, M. K. & Stamey, T. A. Determination of prostate volume by transrectal ultrasound. *J. Urol.* **145**, 984–987 (1991).
8. Zettinig, O. *et al.* Multimodal image-guided prostate fusion biopsy based on automatic deformable registration. *Int. J. Comput. Assist. Radiol. Surg.* **10**, 1997–2007 (2015).
9. Sabouri, S. *et al.* MR measurement of luminal water in prostate gland: Quantitative correlation between MRI and histology. *J. Magn. Resonance Imaging* **46**, 861–869 (2017).
10. Muthupillai, R. *et al.* Magnetic resonance elastography by direct visualization of propagating acoustic strain waves. *Science* **269**, 1854–1857 (1995).
11. Hu, B. *et al.* Evaluation of MR elastography for prediction of lymph node metastasis in prostate cancer. *Abdom. Radiol.* **46**, 3387 (2021).
12. Asbach, P. *et al.* In vivo quantification of water diffusion, stiffness, and tissue fluidity in benign prostatic hyperplasia and prostate cancer. *Investig. Radiol.* **55**, 524–530 (2020).
13. Li, M. *et al.* Tomoelastography based on multifrequency MR elastography for prostate cancer detection: Comparison with multiparametric MRI. *Radiology* **299**, 362 (2021).
14. Li, S. *et al.* A feasibility study of MR elastography in the diagnosis of prostate cancer at 3.0 T. *Acta Radiol.* **52**, 354–358 (2011).
15. Chopra, R. *et al.* In vivo MR elastography of the prostate gland using a transurethral actuator. *Magn. Reson. Med.* **62**, 665–671 (2009).
16. Dittmann, F. *et al.* Tomoelastography of the prostate using multifrequency MR elastography and externally placed pressurized-air drivers. *Magn. Reson. Med.* **79**, 1325–1333 (2018).
17. Arani, A., Plewes, D., Krieger, A. & Chopra, R. The feasibility of endorectal MR elastography for prostate cancer localization. *Magn. Reson. Med.* **66**, 1649–1657 (2011).
18. McGrath, D. M. *et al.* MR elastography to measure the effects of cancer and pathology fixation on prostate biomechanics, and comparison with T1, T2 and ADC. *Phys. Med. Biol.* **62**, 1126 (2017).
19. Reiter, R. *et al.* Prostate cancer assessment using MR elastography of fresh prostatectomy specimens at 9.4 T. *Magn. Reson. Med.* **84**, 396–404 (2020).
20. Sahebjavaher, R. S., Baghani, A., Honarvar, M., Sinkus, R. & Salcudean, S. E. Transperineal prostate MR elastography: Initial in vivo results. *Magn. Reson. Med.* **69**, 411–420 (2013).
21. Sinkus, R., Nisius, T., Lorenzen, J., Kemper, J. & Dargatz, M. In-vivo prostate MR-elastography. *Proc. Intl. Soc. Mag. Reson. Med.* **11**, 586 (2003).
22. Thörmer, G. *et al.* Novel technique for MR elastography of the prostate using a modified standard endorectal coil as actuator. *J. Magn. Reson. Imaging* **37**, 1480–1485 (2013).
23. Sack, I., Jöhrens, K., Würfel, J. & Braun, J. Structure-sensitive elastography: On the viscoelastic powerlaw behavior of in vivo human tissue in health and disease. *Soft Matter* **9**, 5672–5680 (2013).
24. Hudert, C. A. *et al.* How histopathologic changes in pediatric nonalcoholic fatty liver disease influence in vivo liver stiffness. *Acta Biomater.* **123**, 178–186 (2021).
25. Yin, Z., Kearney, S. P., Magin, R. L. & Klatt, D. Concurrent 3D acquisition of diffusion tensor imaging and magnetic resonance elastography displacement data (DTI-MRE): Theory and in vivo application. *Magn. Reson. Med.* **77**, 273–284 (2017).
26. Yin, Z., Magin, R. L. & Klatt, D. Simultaneous MR elastography and diffusion acquisitions: Diffusion-MRE (dMRE). *Magn. Reson. Med.* **71**, 1682–1688 (2014).
27. Mahapatra, D. & Buhmann, J. M. Prostate MRI segmentation using learned semantic knowledge and graph cuts. *IEEE Trans. Bio-med. Eng.* **61**, 756–764 (2014).
28. Klein, S. *et al.* Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med. Phys.* **35**, 1407–1417 (2017).
29. Toth, R. & Madabhushi, A. Multifeature landmark-free active appearance models: Application to prostate MRI segmentation. *IEEE Trans. Med. Imaging* **31**, 1638–1650 (2012).
30. Litjens, G., Debats, O., van de Ven, W., Karssemeijer, N. & Huisman, H. A pattern recognition approach to zonal segmentation of the prostate on MRI, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* 413–420.
31. Zheng, Y. & Comaniciu, D. Marginal space learning for medical image analysis. *Springer*. **2** (2014).
32. Bengio, Y. *et al.* Scaling learning algorithms towards AI. *Large-Scale Kernel Mach.* **34**, 1–41 (2007).
33. Zhu, Q., Du, B., Turkbey, B., Choyke, P. L. & Yan, P. Deeply-supervised CNN for prostate segmentation, 178–184.
34. Milletari, F., Navab, N. & Ahmadi, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation in, 565–571.
35. Yu, L., Yang, X., Chen, H., Qin, J. & Heng, P.-A. Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images.
36. Zabiollahy, F., Schieda, N., Krishna Jayaraj, S. & Ukwatta, E. Automated segmentation of prostate zonal anatomy on T2-weighted (T2W) and apparent diffusion coefficient (ADC) map MR images using U-Nets. *Med. Phys.* **46**, 3078 (2019).
37. Clark, T., Wong, A., A. Haider, M. & Khalvati, F. Fully deep convolutional neural networks for segmentation of the prostate gland in diffusion-weighted MR images. [https://doi.org/10.1007/978-3-319-59876-5\\_12](https://doi.org/10.1007/978-3-319-59876-5_12).
38. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation, in *International Conference on Medical Image Computing And Computer-assisted Intervention*. Springer. 234–241 (2015).
39. Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning, in *Thirty-first AAAI Conference on Artificial Intelligence* **4**, 12–12.
40. Aldoj, N., Biavati, F., Michallek, F., Stober, S. & Dewey, M. Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-Net. *Sci. Rep.* **10**, 1–17 (2020).
41. Streitberger, K.-J. *et al.* How tissue fluidity influences brain tumor progression. *Proc. Natl. Acad. Sci.* **117**, 128–134 (2020).
42. Barentsz, J. O. *et al.* Synopsis of the PI-RADS v2 guidelines for multiparametric prostate magnetic resonance imaging and recommendations for use. *Eur. Urol.* **69**, 41 (2016).
43. Tzschätzsch, H. *et al.* Tomoelastography by multifrequency wave number recovery from time-harmonic propagating shear waves. *Med. Image Anal.* **30**, 1–10 (2016).
44. Streitberger, K.-J. *et al.* High-resolution mechanical imaging of glioblastoma by multifrequency magnetic resonance elastography. *PLoS One* **9**, e110588 (2014).
45. Meyer, T. *et al.* Online platform for extendable server-based processing of magnetic resonance elastography data in Montreal, Quebec, Canada, in *Proc 23rd Annual Meeting ISMRM* 3966.
46. Simard, P. Y., Steinkraus, D. & Platt, J. C. Best practices for convolutional neural networks applied to visual document analysis in, 958–958.
47. Dice, L. R. Measures of the amount of ecologic association between species. *Ecology* **26**, 297–302 (1945).
48. Lowekamp, B. C., Chen, D. T., Ibáñez, L. & Blezek, D. The design of SimpleITK. *Front. Neuroinf.* **7**, 45 (2013).
49. Yaniv, Z., Lowekamp, B. C., Johnson, H. J. & Beare, R. SimpleITK image-analysis notebooks: A collaborative environment for education and reproducible research. *J. Digital Imaging* **31**, 290–303 (2018).
50. Abadi, M. *et al.* Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint [arXiv:1603.04467](https://arxiv.org/abs/1603.04467) (2016).
51. Soerensen, S. J. C. *et al.* ProGNet: Prostate gland segmentation on MRI with deep learning in medical imaging 2021: Image Processing. *Int. Soc. Opt. Photon.* **11596**, 115962R (2021).

52. Rifai, H., Bloch, I., Hutchinson, S., Wiart, J. & Garnero, L. Segmentation of the skull in MRI volumes using deformable model and taking the partial volume effect into account. *Medical Image Analysis*. **4**, 219–233 (2000).
53. Li, M. *et al.* Tomoelastography based on multifrequency mr elastography for prostate cancer detection: Comparison with multiparametric MRI. *Radiology* **299**, 362–370 (2021).

### Acknowledgements

This work was funded by the German Research Foundation (GRK2260, BIOQIC; SFB1340, CRC Matrix in Vision).

### Author contributions

N.A. performed all implementation, evaluation, and scriptwriting, while manual segmentation and results evaluation were done by P.A. and F.B. A.H., M.D. and I.S. did the project supervision and medical and technical guidance.

### Funding

Open Access funding enabled and organized by Projekt DEAL.

### Competing interests

The authors declare no competing interests. MD is European Society of Radiology (ESR) Research Chair (2019–2022), and the opinions expressed in this article are the author's own and do not represent the view of ESR. Per ESR guiding principles, the work as Research Chair is on a voluntary basis, and only travel expenses are remunerated.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-05878-5>.

**Correspondence** and requests for materials should be addressed to I.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022