Fast contour propagation for MR-guided prostate radiotherapy using convolutional neural networks

K.A.J. Eppenhof
Medical Image Analysis Group, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

M. Maspero, and M.H.F. Savenije
Computational Imaging Group for MR Diagnostics & Therapy, Center for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands

J.C.J. de Boer, J.R.N. van der Voort van Zyp, and B.W. Raaymakers
Department of Radiotherapy, University Medical Center Utrecht, Utrecht, The Netherlands

A.J.E. Raaijmakers
Medical Image Analysis Group, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

M. Veta
Medical Image Analysis Group, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

C.A.T. van den Berg
Computational Imaging Group for MR Diagnostics & Therapy, Center for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands

J.P.W. Pluim
Medical Image Analysis Group, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands

(Received 23 October 2019; revised 9 December 2019; accepted for publication 18 December 2019; published xx xxxx xxxx)

Purpose: To quickly and automatically propagate organ contours from pretreatment to fraction images in magnetic resonance (MR)-guided prostate external-beam radiotherapy.

Methods: Five prostate cancer patients underwent 20 fractions of image-guided external-beam radiotherapy on a 1.5 T MR-Linac system. For each patient, a pretreatment T2-weighted three-dimensional (3D) MR imaging (MRI) scan was used to delineate the clinical target volume (CTV) contours. The same scan was repeated during each fraction, with the CTV contour being manually adapted if necessary. A convolutional neural network (CNN) was trained for combined image registration and contour propagation. The network estimated the propagated contour and a deformation field between the two input images. The training set consisted of a synthetically generated ground truth of randomly deformed images and prostate segmentations. We performed a leave-one-out cross-validation on the five patients and propagated the prostate segmentations from the pretreatment to the fraction scans. Three variants of the CNN, aimed at investigating supervision based on optimizing segmentation overlap, optimizing the registration, and a combination of the two were compared to results of the open-source deformable registration software package Elastix.

Results: The neural networks trained on segmentation overlap or the combined objective achieved significantly better Hausdorff distances between predicted and ground truth contours than Elastix, at the much faster registration speed of 0.5 s. The CNN variant trained to optimize both the prostate overlap and deformation field, and the variant trained to only maximize the prostate overlap, produced the best propagation results.


Key words: contour propagation, deep learning, image registration, MR-guided radiotherapy, prostate
1. INTRODUCTION

External-beam radiotherapy is one of the standard treatments for prostate cancer. Because of the superior soft-tissue contrast, magnetic resonance imaging (MRI) is increasingly used in planning and guiding prostate radiotherapy. Extreme hypofractionation with stereotactic body radiotherapy (SBRT) in prostate cancer leads to low genitourinary (GU) and gastrointestinal (GI) toxicity. Recently, MR-guided radiotherapy (MRgRT) has become viable, resulting in even lower GU and GI toxicity. In MRgRT, a pretreatment MRI is used to delineate the clinical target volume (CTV), prior to the daily fractions of radiotherapy. At the start of each fraction, the pretreatment scan is registered to the daily fraction scan. The CTV contour is propagated by deforming it according to the registration and, if necessary, it is manually adjusted. Registration and the manual adjustment of contours are time-consuming and hinder the effectiveness of the treatment due to potential intra-fraction motion of the prostate.

A contour propagation method that is fast and also requires minimal manual adjustments is therefore desirable. Although the prostate is often considered to move rigidly, analyses of prostate motion have shown that a variable degree of deformation is present. For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostatemotion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of deformation is present.10 For this application, deformable analyses of prostate motion have shown that a variable degree of de

Medical Physics, 0 (0), xxxx
in 20 daily fractions of 3.1 Gy. The images used in this study include a T2-weighted three-dimensional (3D) Cartesian turbo spin-echo sequence acquired on the Elekta Unity 1.5 T system with acquisition parameters as specified in Table I. An MRI was acquired in a pretreatment session in advance of the daily fractions to perform structure delineation and enable treatment planning. Similar MRIs were made in advance of each daily treatment session. On this so-called pre-beam MRI, the pretreatment contours need to be propagated such that a new plan can be generated that fits optimally to the patient’s anatomy of that given fraction. In total, 21 scans per patient were collected: 1 pretreatment scan and 20 daily fraction scans. Examples of a pretreatment scan and a daily fraction scan are shown in Fig. 1.

Patients were positioned with the aid of a laser positioning system using anatomy-based tattooed skin markers. These markers are meant to be aligned with the prostate’s axial position to enable scans with the prostate consistently centered in the field-of-view (FOV). At the time of the pretreatment scan, no CTV delineation has been performed, and typically an offset can occur between the marker position and the center of the prostate. After CTV delineation on the pretreatment scans, the offset is known and is accounted for in the patient setup in the daily fraction, centering the CTV in the FOV of the fraction scans. The pretreatment CTV delineation is performed shortly after acquisition. This delineation is not time-critical given the fact that the first fraction is at least one day later than the pretreatment image acquisition.

In contrast to the delineation on the pretreatment scan, the adaptive workflow to propagate the contours to the fraction scans is time-critical. As part of the treatment workflow, during the daily fraction, the FOV of the images was reduced in the axial direction to speed up the workflow. Specifically, the prostate contour was propagated from the pretreatment scan to the daily fraction using rigid registration, after which the image was cropped with a margin of 30 mm superior and inferior to the prostate. The rigid registration was only used to estimate the amount of required cropping, which results in a translation of the prostate in the axial direction. To obtain a similar FOV in the pretreatment images, we cropped the pretreatment images around the prostate CTV with a 30 mm margin.

An expert radiation oncologist drew gross tumor volume (GTV) contours on the pretreatment scans. The CTV was taken as a 4-mm margin around the GTV (excluding the bladder and rectum), and further extended to include the prostate, and the seminal vesicles if they touched or had overlap with the GTV. On the daily fraction scans, CTV contours were constructed by propagation of the CTV contours from the pretreatment to the fraction scans using the ADMIRE deformable registration algorithm (Elekta AB). When necessary, the radiation oncologist manually adjusted or redrew the contours using VolumeTool, an in-house software package. The planning target volume (PTV) was defined as the CTV with an isotropic 5-mm margin in all directions. The manually delineated CTV contours of the pretreatment and daily fraction scans are used and considered the ground truth in this study. Each contour was converted to a binary segmentation, by assigning voxels of which the centers lie inside the contour to the segmented area.

![Pre-treatment scan and Fraction scan](image)

**Fig. 1.** Examples of pretreatment (a) and daily fraction scans (c) with ground truth prostate contours from the same patient in yellow. Both images show the center slice of the volume. (b) and (d) show zoomed-in versions of (a) and (c), respectively. The images are from T2-weighted three-dimensional (3D) turbo spin-echo scans.
2.B. Network architecture

A 3D CNN was designed to propagate the CTVs from the pretreatment images to the fraction images. This network consists of two parts. The first part receives the two images as two input channels and has a deformation vector field composed of three components $u_x$, $u_y$, $u_z$ as output. In the second part of the network, the deformation vector field is used by a spatial transformer layer (STL) to transform the prostate segmentation in the pretreatment image as delineated by the contour. The resulting transformed segmentation is the output of the second part of the network. The network can be jointly trained to optimize the transformed segmentation and the deformation vector field through a loss function that consists of two terms. The first term minimizes the $L_2$-norm of the difference between the estimated and true deformation fields. The other term maximizes the Dice coefficient of the fraction segmentation and transformed pretreatment segmentation (Fig. 2).

The full loss function is defined as

$$L = k_{\text{overlap}}L_{\text{overlap}} + k_{\text{dvf}}L_{\text{dvf}}$$  \hspace{1cm} (1)

where

$$L_{\text{overlap}} = 1 - \frac{2 \sum_{x \in \Omega_F} C_F(x)C_M(\hat{T}(x))}{\sum_{x \in \Omega_F} C_F(x) + \sum_{x \in \Omega_M} C_M(T(x))}$$ \hspace{1cm} (2)

optimizes the overlap of the segmented volumes, and

$$L_{\text{dvf}} = \frac{1}{|\Omega_F|} \sum_{x \in \Omega_F} ||T(x) - \hat{T}(x)||^2_2$$ \hspace{1cm} (3)

optimizes the deformation vector field (DVF). $\hat{T}$ is the network’s estimate of the actual transformation $T$, and $k_{\text{overlap}}, k_{\text{dvf}} \in [0, 1]$ are weighting parameters. The focus of optimization will change based on these parameters. When $k_{\text{dvf}}$ is set to zero, the network will be trained to maximize the Dice coefficient between the segmentations. No explicit deformation field is required in this case, and as a result, the deformation fields estimated by the network cannot be guaranteed to have a real physical interpretation. When we set $k_{\text{overlap}}$ to zero, the network is trained to predict deformation fields similar to those in the training set. In this case, the segmentation overlap is completely ignored. The hybrid case ($k_{\text{overlap}} = k_{\text{dvf}} = 1$) combines both training objectives.

For the estimation of the deformation field, we use a standard three-dimensional U-net architecture and adapt it to accept two-channel inputs and three-channel outputs. The U-net is composed of five resolution levels in which the convolutional layers have outputs of specific dimensions, ranging from $128 \times 128 \times 128$ at the top of the architecture, to $8 \times 8 \times 8$ at the bottom. Each level has a distinct color in Fig. 3. In previous work we have shown that it is beneficial to train the U-net for image registration progressively. To this end, we add input and output layers to each of the resolution levels. Outputs are scaled to $128 \times 128 \times 128$ dimensions and a weighted sum is taken as the final output of the network. The weights of this sum are not learned, but set during the training process, which allows us to control which level contributes to the result. Initially, we only let the lowest dimension ($8 \times 8 \times 8$) contribute to the result. After a fixed number of iterations $N$, we linearly decrease the weight of this level to zero and linearly increase the weight of the level above it to one ($16 \times 16 \times 16$) over $M$ iterations. Then we let the two-level U-net train for another $2N$ iterations followed by another $M$ iterations of transition to a three-level U-net, and so on until we obtain a typical five-level U-net. The five-level U-net is then trained further in the normal way. At test time, the architecture does not differ from the typical U-net architecture. The chosen schedule for this study used $N = 1000$ and $M = 2000$.

The output of the U-net is a set of three maps for the $u_x$, $u_y$, and $u_z$ components of the deformation field. We use these maps as input to a STL that transforms the pretreatment segmentation. The STL uses nearest neighbor interpolation to sample the pretreatment segmentation based on the deformation vector field by the U-net. The STL is differentiable, allowing the gradients of the overlap loss in Eq. (2) to back-propagate through the STL. The U-net can map the input images to a deformation field that leads to a high overlap of the ground truth fraction segmentation and the transformed pretreatment segmentation.

![Fig. 2. General overview of the method. The method consists of two parts. In the first part, a convolutional neural network predicts a deformation field from a pretreatment image and fraction image. In the second part, the predicted deformation field is used by a spatial transformer layer to deform the segmentation as delineated on the pretreatment image. The loss functions $L_{\text{dvf}}$ and $L_{\text{overlap}}$ are only computed during training.](image-url)
2.C. Training set

The training set consists of pairs of synthetically deformed images and delineated prostate segmentations, for which the ground truth deformation field is generated. During each iteration of training, a unique combination of input images, segmentations, and deformation fields is constructed by applying random deformations. This allows for supervised training with a known ground truth. We hypothesize that the network should be able to generalize from these synthetic examples to real registration problems, as we have shown in previous work.36

The process of creating training examples is outlined in Fig. 4, and is similar to the proposed method for training data generation in Ref. [36,37]. Every iteration of training, an image \( I \) and associated CTV segmentation \( S \) from the training set are selected. A random deformable transformation \( T \) is sampled and applied to both, resulting in \( I(T) \) and \( S(T) \). From the images \( I \) and \( I(T) \), the network has to predict the transformation \( T \) that can be applied to \( S \) to obtain an estimate for the ground truth segmentation \( S(T) \). To increase the amount of data, we also apply free-form deformations for data augmentation using an additional random transformation to the images and segmentations. To limit interpolation artifacts, the augmentation transformations consist of a random rigid transformation with rotation angles sampled from a uniform distribution between \(-0.1\) and \(0.1\) radians, and translation vectors sampled from a uniform distribution between \(-12.8\) and \(12.8\) voxels, followed by a random deformable B-spline transformation similar to \( t_1 \) in Table II. Both transformations are concatenated with a scaling transformation that effectively crops the images around the center. We scaled the images by a factor of two in-plane, and a factor of 1.33 in the axial direction. For the images used in this study this means that the effective FOV of the network is \(200 \times 200 \text{ mm}^2\) in the transversal directions and between 100 and 130 mm in the axial direction. Because the prostate appears larger in the images, the network can use a larger part of the estimated deformation vector field for the prostate itself. A requirement is that the prostate is positioned in the center of the FOV, which was the case for all images used in the study. The transformations are sampled on a \(128 \times 128 \times 128\) voxel grid using third-order B-spline interpolation. This results in the input images, the deformation fields, and the output segmentation having \(128 \times 128 \times 128\) voxels, with each voxel having between a \(1.56 \times 1.56 \times 0.86\) and \(1.56 \times 1.56 \times 1.01 \text{ mm}^3\) voxel size. At training, validation, and test time, the pairs of images that are used as input to the network are preprocessed by linearly scaling the image intensities between 0 and 1. The deformation fields are expressed in voxel coordinates.

The network is optimized using stochastic gradient descent with momentum. The momentum parameter was set to

---

Medical Physics, 0 (0), xxxx
images, an additional augmentation transformation is applied to the original data first.

**Transformation Grid size Distribution**

**TABLE II. B-spline transformation parameters for the learned transformation.**

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Grid size</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_1 )</td>
<td>( 2 \times 2 \times 2 )</td>
<td>( U(-6.4, 6.4) )</td>
</tr>
<tr>
<td>( t_2 )</td>
<td>( 4 \times 4 \times 4 )</td>
<td>( U(-3.2, 3.2) )</td>
</tr>
<tr>
<td>( t_3 )</td>
<td>( 8 \times 8 \times 8 )</td>
<td>( U(-1.6, 1.6) )</td>
</tr>
<tr>
<td>( t_4 )</td>
<td>( 16 \times 16 \times 16 )</td>
<td>( U(-0.8, 0.8) )</td>
</tr>
</tbody>
</table>

\( U(a,b) \) is the multivariate uniform distribution that samples vectors with \( n \) components in the \( [a,b] \) interval in voxels. The resulting transformation is defined as \( T_{\text{learned}} = t_1 \circ t_2 \circ t_3 \circ t_4 \).

0.5, and the learning rate was set to 0.01. A batch size of one was used throughout training. All convolutional layers were trained with batch normalization, using the exponential moving average of the batch normalization parameters, as proposed by Ioffe et al.\(^4\)

**2.D. Experiments**

To maximize the number of patients in the training set, we perform a variation of leave-one-out cross validation. We used the fraction scans of three patients for training, the scans of the fourth patient for validation, and the fifth patient’s scans as a test set. We run five permutations, where each patient is used once in the test set, and once in the validation set. The validation set was used to monitor the model’s performance on independent data during training. For each of the five folds, hyperparameters like the learning rate were kept the same. Each of the three networks was trained for 15 000 iterations for each of the five folds.

For each patient, we train three variants of the network: one trained on segmentation overlap \( (k_{\text{overlap}} = 1, k_{\text{def}} = 0) \), one trained on deformation field estimation \( (k_{\text{overlap}} = 0, k_{\text{def}} = 1) \), and a hybrid of these two \( (k_{\text{overlap}} = 1, k_{\text{def}} = 1) \). We do this for all instances in the cross validation. To distinguish the variants of the network, we will call them “overlap loss network,” “deformation loss network,” and “hybrid loss network” for the remainder of this paper. We compare the presented method to a more conventional deformable image registration method in Elastix, an open-source image registration software package.\(^4\) We use a deformable image registration algorithm implemented in Elastix published by Klein et al.\(^1\) The original purpose of this algorithm was to perform automatic atlas-based segmentation of the prostate in 3D MR images. The algorithm by Klein et al. performs a rigid registration first, followed by a deformable registration based on B-spline transformations, by optimizing the localized mutual information.\(^4\)

For all network variants and Elastix, we measure the Dice coefficient of the fraction segmentation and propagated pretreatment segmentation. As a measure of registration error for the prostate, we calculate the difference between the centroid of the prostate in the propagated pretreatment segmentation and the fraction segmentation. As a metric for contour distance, we calculate the 95th percentile of the Hausdorff distance between the segmentations. We check for folding by inspecting the determinant of the Jacobian of the deformation field. As a measure of folding, we count the percentage of voxels with a negative Jacobian determinant, both for the full image and for the voxels inside the prostate. To assess the robustness of the methods, we assess how they handle additional shifts of the prostate that may occur during a fraction, by testing the performance on a range of superior–inferior shifts. These shifts are likely to occur in clinical practice due to the filling of the bladder, which is right above the prostate. We apply shifts along the axial direction between \( -5 \) and \( +5 \) mm with \( 1 \) mm steps to each of the pretreatment images and test the propagation to all fractions for each patient. The shifts are applied by changing the cropping window that is described in Section 1.

To show the difference in speed of each of the algorithms, we record the time necessary to complete each registration.
problem including the transformation of the segmentation for the Elastix method and the proposed networks.

3. RESULTS

In Fig. 5, we show box plots for the Dice coefficient, 95th percentile of the Hausdorff distance, and the registration error for the prostate’s centroid for the propagation from the pre-treatment scan to each of the 20 fraction scans. For each patient, we show five boxes: red boxes represent the metrics without propagation, gray boxes represent Elastix’ propagation results, and green, blue, and purple boxes represent the results of the overlap, deformation and hybrid networks, respectively. On the x-axis, boxes are grouped by patient.

In general, the results obtained by the overlap loss network and the hybrid loss network are superior to the results obtained with Elastix, with the deformation loss network having a somewhat lower Dice coefficient, and larger Hausdorff registration and centroid distances compared to the other methods (Table IV). The Wilcoxon signed-rank test was used to analyze whether the Dice score, Hausdorff distances, and prostate centroid distances obtained by the CNNs were significantly different from those obtained by Elastix. Full results are shown in Table III. At a significance level $\alpha = 0.01$, we can conclude significant improvement of the Hausdorff distances and Dice coefficients over Elastix for the overlap loss and hybrid loss networks, whereas Elastix is significantly better than the deformation loss network. In Fig. 6,
examples of propagated contours are shown for each of the
methods, showing that generally the contours have better
overlap for the overlap loss and hybrid loss networks than for
Elastix and the deformation loss network.

In Table IV, we show the percentage of voxels with nega-
tive Jacobian. Elastix and the deformation loss network suffer
from little to no folding, whereas the networks that were also
trained using the Dice coefficient show substantially more
folding. Fig. 7, shows the results of the axial shifting experi-
ment, in which the prostate is axially shifted between −5 and
+5 mm. For each of the shifts, we show the distribution of
the 95th percentile of the Hausdorff distance for all possible
pretreatment-to-fraction registration problems across the five
patients. The shifts induce little variation in the Hausdorff
distance for any of the methods except for the deformation
loss network. This indicates that the overlap loss and hybrid
loss networks and Elastix are insensitive to the shifts. How-
ever, the overlap loss and hybrid loss network outperform
Elastix over the whole range of applied shifts.

For each of the methods, we recorded the time required to
complete the propagation of the segmentation on a system
with an Intel Xeon E5-2640 v4 CPU, 512 GB of memory
and an Nvidia Titan XP graphics card with 12 GB of GPU
memory. For each of the network variants, the timing is the
same, amounting to 0.49 ± 0.10 s (μ ± σ) when running on
the GPU. For Elastix, executing the propagation problem
required on average 43.2 ± 0.29 s, using both the CPU and
GPU for the registration.

4. DISCUSSION

In this paper, we have proposed a deep learning-based
method for fast deformable propagation of clinical target
volume contours from pretreatment to fraction scans. Three
variants of a CNN trained with different loss functions have
been tested, based on contour overlap, prediction of the defor-
mation field, and a hybrid of the two.

The results show that the propagation accuracy of the net-
works trained with the overlap loss and hybrid loss perform
significantly better (P < 0.01) at Dice coefficients and
Hausdorff distances compared to the open source registration
package Elastix. Hausdorff distances measured between the
predicted contours and the ground truth indicate that the reg-
istration error at the contours is on average 5.7 mm for the
hybrid loss network compared to 7.5 mm for Elastix. The
deformation loss network on average performs worse than the
other two network variants and Elastix. A crucial aspect for
the application is timing: the networks can propagate the

Table III. P-values for the Wilcoxon signed-rank test between Elastix and
each of the three network variants.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Overlap loss</th>
<th>Deformation loss</th>
<th>Hybrid loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dice coefficient</td>
<td>P &lt; 10⁻⁵★</td>
<td>P &gt; 0.99†</td>
<td>P &lt; 10⁻⁵★</td>
</tr>
<tr>
<td>95th perc. Hausdorff</td>
<td>P &lt; 10⁻⁵★</td>
<td>P &gt; 0.99†</td>
<td>P &lt; 10⁻⁵★</td>
</tr>
<tr>
<td>Prostate centroid</td>
<td>P &gt; 0.11○</td>
<td>P &gt; 0.99†</td>
<td>P &lt; 0.05★</td>
</tr>
</tbody>
</table>

Significant improvements (P < 0.01) over Elastix are indicated with ★, no signifi-
cant differences with ○, and cases where Elastix is superior with †.
prostate segmentations within 0.5 s, much faster than the 43 s it takes to perform the propagation in Elastix.

Inspection of folding in the predicted deformation fields shows that the deformation loss network and Elastix perform much better in this respect, with less than two percent of the FOV showing any folding. The deformation loss network will cause the network to mimic the transformations in the training set, which do not fold by construction. Because the overlap loss and hybrid loss networks maximize prostate overlap, the learned deformation fields can contain unrealistic deformations. It should be noted that for the current application the true deformation field can show folding as well, as sliding motion between the prostate and surrounding organs is possible.

A shift experiment shows that the overlap and hybrid loss networks are not very sensitive to larger displacements of up to an additional 5 mm between the prostate in the pretreatment and fraction scan. The fraction scans used in this study show that it is possible to perform the procedure accurately based on the tattooed markers, and that expected displacements between fraction scans falls within this range. The network trained on deformation loss was shown to be more sensitive to additional shifts, which can be attributed to limitations in the distribution of deformations in the training set.

Our results show that it is possible for the networks to generalize from a set of four patients (three in the training set, one in the validation set) to a fifth patient. This also suggests that with additional data becoming available, it will be possible to train a single neural network that may generalize to new patients. To cope with the currently limited number of patients, we have adopted a leave-one-out strategy. With more data becoming available, it will ultimately be possible to train a single model that can be used for multiple new patients.

The proposed method introduces two hyperparameters that need to be optimized, $k_{\text{overlap}}$ and $k_{\text{dvf}}$, that weight high prostate overlap and correct estimation of the deformation vector field respectively. In the experiments we have tried binary settings of these parameters, and there is potentially a better, non-binary setting for the current application. With the limited data available, we have not been able to test this, but future work could focus on the optimization of the hyperparameters.

Delineation methods using deep learning based prostate segmentation can also be employed to delineate CTVs on fraction scans, provided the CTV always covers the same type of tissue, as discussed in the introduction section. Segmentation methods perform well when the CTV covers only the whole prostate, but fail when other surrounding tissue should be included in the CTV. A CNN trained to propagate contours can account for these variations in the pretreatment delineations, by using information from previous contours. Although a quantitative comparison to prostate segmentation is difficult because of these differences, the Dice coefficients and Hausdorff distances for the two best performing networks

### Table IV. Results of the evaluated methods in terms of Dice coefficient, 95th percentile of the Hausdorff distance, prostate centroid distance, amount of folding (Jacobian determinant $< 0$), and duration of each algorithm.

<table>
<thead>
<tr>
<th>Metric</th>
<th>None</th>
<th>Elastix</th>
<th>Overlap loss CNN</th>
<th>Deformation loss CNN</th>
<th>Hybrid loss CNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dice coefficient</td>
<td>0.62 ± 0.14</td>
<td>0.78 ± 0.12</td>
<td>0.86 ± 0.05</td>
<td>0.75 ± 0.07</td>
<td>0.86 ± 0.05</td>
</tr>
<tr>
<td>95th perc. Hausdorff (mm)</td>
<td>13.15 ± 6.37</td>
<td>7.47 ± 4.72</td>
<td>5.82 ± 3.63</td>
<td>8.16 ± 4.11</td>
<td>5.66 ± 3.56</td>
</tr>
<tr>
<td>Prostate centroid (mm)</td>
<td>11.38 ± 4.78</td>
<td>3.29 ± 1.95</td>
<td>2.99 ± 1.57</td>
<td>4.10 ± 2.14</td>
<td>2.85 ± 2.04</td>
</tr>
<tr>
<td>Folding in FOV (%)</td>
<td>–</td>
<td>0.00 ± 0.00</td>
<td>33.97 ± 4.20</td>
<td>1.78 ± 0.79</td>
<td>7.32 ± 2.19</td>
</tr>
<tr>
<td>Duration (s)</td>
<td>–</td>
<td>43.2 ± 0.29</td>
<td>0.49 ± 0.10</td>
<td>0.49 ± 0.10</td>
<td>0.49 ± 0.10</td>
</tr>
</tbody>
</table>

Each figure is the mean ± standard deviation over the 20 pretreatment-to-fraction propagations averaged over all patients.
are in the same range as deep learning-based prostate segmentation methods in literature.\textsuperscript{15–19} However, it should again be noted that prostate segmentation is not the same as CTV segmentation in this case.

Further applications of the proposed method include the propagation of multiple contours at once. The proposed networks can be adapted to transform multiple segmentations using the STL. In that case, the setup stays the same, except for the fact that multiple masks are fed to the STL, and that the overlap loss will need to be replaced with the multi-class generalized overlap loss.\textsuperscript{45} The transformation of multiple segmentations can happen in parallel on the GPU, which means that there will be no additional computation time for multiple contours. The proposed method may also be valuable in \textit{intra}-fraction propagation of contours, by registering the pretreatment or a previous fraction scan to cine-MR images that are continuously acquired during MR-guided treatment.\textsuperscript{36} This is particularly of interest for hypofractionated schemes, in which the individual fractions may take more time, which can potentially increase the impact of registration errors during the treatment. To measure the efficacy of the propagated contours, future work will include a more extensive validation investigating the time required for a specialist to adapt the contours.

It is important to note that the adaptive MRgRT workflow is performed under continuous human supervision and that this would also be the setting in which the current method would potentially be used. The radiation oncologist will always check the contours and improve them manually where necessary. A further speed improvement can be obtained by reducing the manual corrections required from the radiation oncologist.

5. CONCLUSIONS

We have developed a fast contour propagation method for prostate radiotherapy, that can propagate contours from pre-treatment to fraction scans in under half a second. The method achieves superior propagation accuracy to an existing and more time consuming deformable registration method. The proposed method can aid in shortening treatment time by reducing time spent on propagation and manual contour adaptation.

CONFLICT OF INTEREST

The authors have no conflict to disclose.

\textsuperscript{a}Author to whom correspondence should be addressed. Electronic mail: k.a.j.eppenhof@tue.nl.

REFERENCES

24. Dalca AV, Balakrishnan G, Guttaj JV, Sabuncu MR. Unsupervised Learning for Fast Probabilistic Diffeomorphic Registration. in:


