Article:
Lorenzo Polo, A., Nix, M. orcid.org/0000-0001-7228-7344, Thompson, C. et al. (6 more authors) (2024) Improving hybrid image and structure-based deformable image registration for large internal deformations. Physics in Medicine & Biology, 69 (9). 095011. ISSN 0031-9155

https://doi.org/10.1088/1361-6560/ad3723

Improving hybrid image and structure-based deformable image registration for large internal deformations

A. Lorenzo Polo[1], M. Nix[2], C. Thompson[2], C. O'Hara[2], J. Entwisle[2], L. Murray [2,3], A. Appelt [3,4], O.Weistrand[1], S. Svensson[1]

[1] RaySearch Laboratories, SE-104 30 Stockholm, Sweden
[2] Leeds Cancer Centre, Department of Medical Physics, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[3] Leeds Cancer Centre, Department of Clinical Oncology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[4] Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, UK

E-mail: andrea.lorenzopolo@raysearchlabs.com

Keywords: Deformable Image Registration, Re-Irradiation, Large Deformations

Abstract.

Objective. Deformable image registration (DIR) is a widely used technique in radiotherapy. Complex deformations, resulting from large anatomical changes, are a regular challenge. DIR algorithms generally seek a balance between capturing large deformations and preserving a smooth deformation vector field (DVF). We propose a novel structure-based term that can enhance the registration efficacy while ensuring a smooth DVF. *Approach.* The proposed novel similarity metric for controlling structures was introduced as a new term into a commercially available algorithm. Its performance was compared to the original algorithm using a dataset of 46 patients who received pelvic re-irradiation (reRT), many of which exhibited complex deformations. *Main results.* The mean Dice Similarity Coefficient (DSC) under the improved algorithm was 0.96, 0.94, 0.76, and 0.91 for bladder, rectum, colon, and bone respectively, compared to 0.69, 0.89, 0.62, and 0.88 for the original algorithm. The improvement was more pronounced for complex deformations. *Significance.* With this work, we have demonstrated that the proposed term is able to improve registration accuracy for complex cases while maintaining realistic deformations.

1. Introduction

Deformable image registration (DIR) algorithms play an important role in medical image analysis and radiotherapy applications. These algorithms aim to align a pair of images by computing a deformation vector field (DVF) that maps each point in one image (the

reference image) to the corresponding anatomical point in another image (the target image).

In the context of radiotherapy, DIR is frequently employed as a tool for propagating organ delineations between images. This is, for instance, used in adaptive radiotherapy (ART) for monitoring interfraction tumor motion, allowing for an accurate consideration of changes in the size and position of organs at risk (OARs) and gross tumor volumes (GTVs) (White et al. 2019). DIR is also used in ART for mapping image densities from planning CTs to daily cone beam CTs (CBCTs) or MRs (O'Hara et al. 2022; Mannerberg et al. 2020). This enables treatment planning on treatment images that do not present electron density information. The DVFs can also be utilized for mapping dose distributions defined on one image set to another (Murr et al. 2023). Dose mapping is valuable for dose summation, such as in cases of localized recurrence requiring re-irradiation (reRT) (Nix et al. 2022) or in the estimation of a dose of the day during ART (Veiga et al. 2014). Other applications of DIR include response assessment to radiotherapy treatments, enabling monitoring of the evolution of treated regions over time (Sen et al. 2020; Bertelsen et al. 2011).

Despite the widespread usage of DIR, deformation algorithms encounter challenges when dealing with complex deformations that involve significant anatomical changes and organ displacements, leading to DIR uncertainties (Nenoff et al. 2023). Such scenarios commonly arise in reRT cases, where a considerable time gap may exist between image acquisitions (Vásquez Osorio, Mayo, et al. 2023; Nix et al. 2022); also in pelvic cases, characterized by difficulties in achieving consistent bladder and rectum filling (Thörnqvist et al. 2010) and internal risk organ movements (Ghose et al. 2015); and, finally, in brachytherapy procedures, where an applicator is present in only one of the images (Vásquez Osorio, Kolkman-Deurloo, et al. 2015). The limitation of currently available DIR algorithms in handling complex deformations is highlighted by Sen et al. (2020). Many efforts in the DIR community are directed at addressing the problem of large deformations while maintaining a reliable deformation field that is smooth and invertible (Kang et al. 2022).

DIR algorithms are characterized by three main components: a similarity measure, a transformation model, and an optimization process, as described in detail by, e.g., Oh and Kim (2017), and briefly presented below. The similarity measure techniques evaluate the correspondence between two images, applying either intensity-based approaches, structure-based approaches, or a combination of both. Intensity-based measures use the intensity values of each image pair to measure how well they match. Structure-based methods rely on landmarks defined in the image pairs as surfaces, points, or curves, and measure how well they align based on the distance between these landmarks. This can be referred to as contour guidance or, which will be used in the following, as registration of controlling structures. While this approach enhances the model's robustness to noisy images, it requires the delineation of anatomical structures and its performance relies on the quality of these delineations. The transformation model specifies the way the reference image is mapped to the target image, and the optimization

process is chosen based on the selected model. The optimization process is usually done in several resolutions, starting with low-resolution images and propagating the results to the next resolution level. Regularization techniques are employed in the cost function to enforce spatial transformation properties. These techniques mitigate issues such as folding, inverted elements, and lack of smoothness in the resultant deformation field.

With the emergence of deep neural networks, deep learning (DL) registration techniques have gained significant attention in medical image analysis (De Vos et al. 2019). Unlike traditional methods that require optimization for each image pair, DL methods offer rapid outputs. However, these DL methods rely on extensive and representative datasets to effectively train the models, which may pose challenges if such datasets are not readily accessible. Balakrishnan et al. (2019) proposed VoxelMorph, a U-Net based model that demonstrated a strong performance. Several modifications of this model have been proposed. Kang et al. (2022) implemented a pyramid architecture to predict multi-scale deformations, which improved the ability to handle large deformations. While this approach outperformed VoxelMorph, it yielded a significant fraction of the DVF which was folded, hence related to a non-realistic deformation. More recently, the use of transformers for registration problems, as TransMorph (Chen et al. 2022), has shown promising results, including a reduced number of folded elements and improved similarity metrics compared to traditional and DL models. Nonetheless, it is crucial to assess whether the larger receptive field advantage offered by transformers translates effectively to complex scenarios involving large deformations, such as the scenario studied in Thörnqvist et al. (2010) for variations in bladder volumes.

This study introduces a novel structure-based term, $Z(v)$, designed to enhance the optimization process by providing guidance for anatomical structures, particularly in scenarios involving large deformations and organ displacements. To illustrate its efficacy, we integrate this term into ANACONDA (Weistrand and Svensson 2015), a commercially available hybrid DIR algorithm integrated into the treatment planning system RayStation (RaySearch Laboratories), and conduct experiments on a dataset comprising pelvic cases treated with reRT. Validation is done based on structure propagation accuracy and the performance is compared to an algorithm tailored for large deformations in reRT (Nix et al. 2022).

1.1. ANACONDA Overview

ANACONDA stands for ANAtomically CONstrained Deformation Algorithm and has an optimization function that combines an image similarity term, a regularization term (in two components), and a structure-based term (Weistrand and Svensson 2015). Given a reference and a target image that are rigidly registered by a matrix M, the DIR algorithm generates a DVF, v, that maps the voxels of the reference image to its corresponding location within the target image. The deformation vector at gridpoint $x_i \in R^3$ is denoted as $v_i \in R^3$, and the registration is expressed as a non-linear optimization problem:

$$f(v) = \alpha C(v) + \beta H(v) + \gamma S(v) + \delta D(v) \quad (1)$$

The image similarity term $C(v)$ measures the correlation coefficient or mutual information between the target and reference images. To enhance the registration accuracy over a specific region, the image similarity can be computed on a certain area defined as the focus region. The default focus region is set to the patient outline, in the following referred to as external. The regularization term consists of $\beta H(v) + \gamma S(v)$. The first component, $\beta H(v)$, enforces smoothness within the DVF and prevents the presence of inverted elements. It operates by minimizing the Dirichlet energy for the coordinate functions of the DVF, quantifying how much it deviates from a harmonic map. The second component, $\gamma S(v)$, is called shape-based regularization and measures the deformation of the ROIs, penalizing large shape deformations. The structure-based term can be divided into two independent components, $\delta_R D_R(v) + \delta_P D_P(v)$, operating only in the presence of controlling ROIs or POIs, respectively. $D_R(v)$ computes the summations of the approximate Euclidean distances between each vertex of the target ROI mesh to the closest voxel of the reference ROI. This minimizes the distance among the surfaces of the controlling ROIs. The point-based term, D_P, penalizes the distance between each point in the target image and the mapped reference image. A more comprehensive description can be found in the original ANACONDA publication (Weistrand and Svensson 2015).

The motivation for using ANACONDA stems from a known limitation: it has been observed to be unable to handle large deformations. This was illustrated by Nix et al. (2022) in a series of pelvic patients, where ANACONDA was only used when bladder volume changes were below 80%, pointing out its limitation in handling large deformations. Similarly, Wang et al. (2021) acknowledged that ANACONDA struggles to achieve accurate registrations of the upper boundary of the bladder. ANACONDA has also been reported to perform poorly in brachytherapy scenarios characterized by complex deformations due to the applicator insertion (Rigaud, Klopp, et al. 2019; Rigaud, Cazoulat, et al. 2020).

These failure cases can be explained by the behavior of the term D_R. As detailed by Weistrand and Svensson (2015), in D_R, the controlling ROIs of the target image are used to obtain the absolute values of its signed distance map. The objective function pushes the surfaces of the reference structure to the areas of the signed distance map with the lowest values, that is, where the surface of the target structure is located. While this strategy generally yields satisfactory results, it may lead to erroneous surface localization of the reference structure in cases of large deformations or displacements.

Figures 1 (a) and (b) illustrate the problem when large deformations are present. The resulting DVF will follow the path of the pink arrow in solid, such that the reference ROI will move towards an incorrect surface of the target ROI. In Figures 1 (a2) and (b2), it can be seen that this occurs since the surface of the reference ROI falls on the side of the distance map closer to the wrong surface of the target ROI.

In addition, the original algorithm presents a lack of robustness in organ

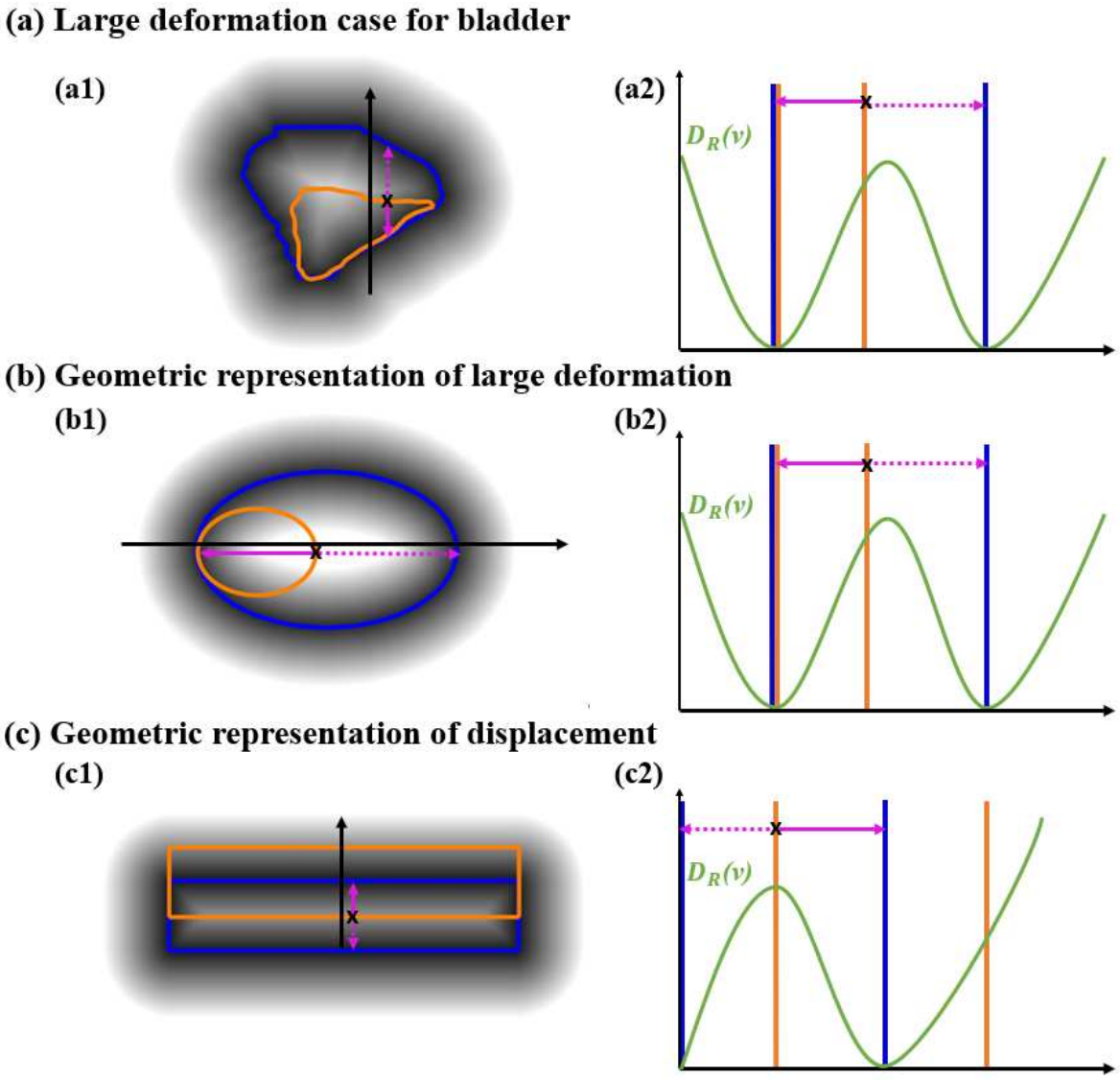

Figure 1: Deformations that lead to erroneous registrations. In (a1), (b1), and (c1), the surface of the ROI is shown in blue for the target and orange for the reference. The absolute values of the distance map of the target ROI are shown as background. The dotted pink arrow corresponds to the desired direction for the surface to move towards, and the solid to the actual direction. The black arrow is a profile used as the x-axis in (a2), (b2), and (c2) to sketch the behavior of $D_R(v)$. The vertical lines represent the surfaces on that profile, and the arrows show the directions that the reference surface can take. The reference surface will move towards the closest minima, following the solid pink arrow.

displacement scenarios. This issue arises under specific circumstances with flat-shaped ROIs like chest walls, when a surface of the structure in the reference image is located in the middle of two surfaces of the structure in the target image. In Figure 1 (c), it can be seen that the algorithm could provide different outputs for very similar scenarios

when the surface is located in between two surfaces.

These two problems have the same root cause: the algorithm fails because a reference surface is closer to a different surface than to its corresponding one in the target structure. We aim to solve this issue while not affecting the performance of success cases. To improve the structure-based guidance, a new term $Z(v)$ is introduced to ensure that the deformed reference structure matches its corresponding target structure. The original term for controlling structures $D_R(v)$ is maintained, but a new weight ϵ models its behavior along the different resolution levels.

2. Methods

The proposed structure-based term $Z(v)$ is implemented as an additional term in ANACONDA's optimization function. The new objective function is defined as:

$$f(v) = \alpha C(v) + \beta H(v) + \delta_P D_P(v) + \delta_R(Z(v) + \epsilon D_R(v)). \quad (2)$$

The original terms for image similarity, $C(v)$, regularization, $H(v)$, and point-based guidance, D_P, are preserved in this study since they are not directly tied to the limitations discussed above. The shape-based regularization term, $S(v)$, was removed from the commercial implementation of ANACONDA. An internal investigation carried out with additional data showed that the term did not influence the end result enough to motivate its usage. The proposed term, $Z(v)$, and an additional weighting term for $D_R(v)$, ϵ, are described below.

2.1. New Similarity Metric for Controlling Structures

The $Z(v)$ term exploits that the ROIs can be defined as volumetric images, so it evaluates the mapping of the volumes as opposite of $D_R(v)$ which evaluates the mapping of the surfaces. $Z(v)$ can be presented as a volume term that ensures a gross volume matching, while $D_R(v)$ provides a precise surface matching.

The controlling structures are represented as volumetric images by generating soft edge volumes, with each structure defined as a separate volume. Figure 2 illustrates how a soft-edge volume is created. From a contour defined on a slice (red line on Figure 2 (b)), each voxel receives a value from 0 to 255, depending on how much of the voxel is included within the contour. This method is beneficial for small structures, like optic nerves, since it is independent of the ROI's size - it only depends on the resolution. By representing the controlling structures as soft edge volumes rather than binary volumes, a more precise definition of the edges is achieved, aiding the algorithm in the registration of the surfaces.

The term $Z(v)$ is computed as the average of the individual correlation coefficients per ROI of the target and deformed reference soft edge volumes. The correlation coefficient measure employed is the same as that used in the similarity term $C(v)$ (Cohen et al. 2009; Weistrand and Svensson 2015):

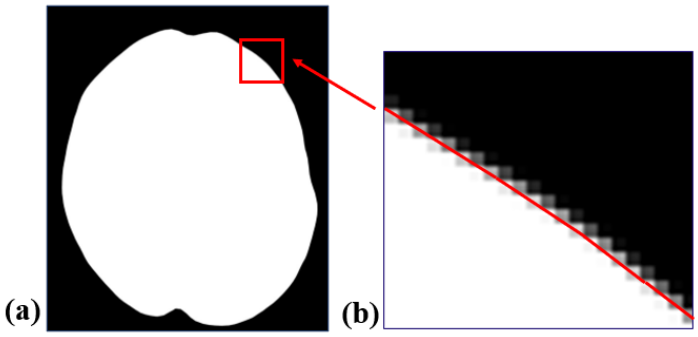

Figure 2: Illustration of the soft-edge concept. In (a) a slice of a 3D volume is shown. In (b) a close-up of the soft edge is shown, with the original contour overlayed in red. Each voxel has a value from 0 to 255, based on how much of the voxel is included within the contour.

$$Z(v) = \frac{1}{k}\sum_{k} \frac{\sum_i (R(x_i) - \bar{R})(T(M(x_i) + v_i) - \bar{T})}{\sqrt{\sum_i (R(x_i) - \bar{R})^2}\sqrt{\sum_i (T(M(x_i) + v_i) - \bar{R})^2}}, \tag{3}$$

where k is the number of controlling structures, R and T are the reference and target images, $\bar{R}$ and $\bar{T}$ are the mean intensities over the set of points contained in the focus region, and $\sum_i$ is a sum over every point x_i contained in the focus region.

While the focus region for computing the correlation coefficient in $C(v)$ can be modified to emphasize the regions where the intensity mapping is more crucial, the focus region in $Z(v)$ is not modifiable and is the same for every structure. In $Z(v)$, the focus region is defined as the union between the external and all the controlling structures. Selecting the external provides more background for the calculation of the correlation coefficient (e.g., the structure in Figure 2 (a) will be contained in a background with the size of the external). Including the controlling structures ensures that structures outside the external are also considered.

2.2. New Weight Term ϵ

The term $Z(v)$ guides the deformation of the reference soft edge volume so that it aligns with the target soft edge volume, while the original structure term $D_R(v)$ provides a more precise surface matching outcome. As mentioned before, in scenarios presenting large deformations or displacements, the $D_R(v)$ term may lead to deforming the reference structure towards the wrong surface of the target structure. In such cases, a conflict arises during optimization, wherein the increase of the $D_R(v)$ term leads to a decrease in $Z(v)$, and vice versa. To address this problem, we include a weight ϵ that updates gradually.

The optimization process typically involves three resolution levels, with 10, 5, and 2.5 mm values. This approach enables the introduction of a coarse-to-fine mapping method, gradually increasing ϵ as the resolution level increases. The term $Z(v)$ is the

average of the correlation coefficient of the controlling structures, ranging between 0 and 1, modified such that the minimum value 0 indicates maximum similarity. The value of $D_R(v)$ measures the distance between surfaces and can take any positive value starting from 0, with zero indicating that the surfaces match. To prevent large deformations to provide a dominant $D_R(v)$ term that could shadow $Z(v)$ during optimization, ϵ is set to zero at the initial iteration. As subsequent optimizations benefit from a more accurate starting deformation field, the maximum value $D_R(v)$ will present significantly lower values than in the first resolution level. Hence, ϵ is incrementally increased across the resolution levels until reaching a value of one in the last level. Figure 3 illustrates the behavior of the new term $Z(v)$ and the new weight ϵ, and how they deal with large deformations and displacements.

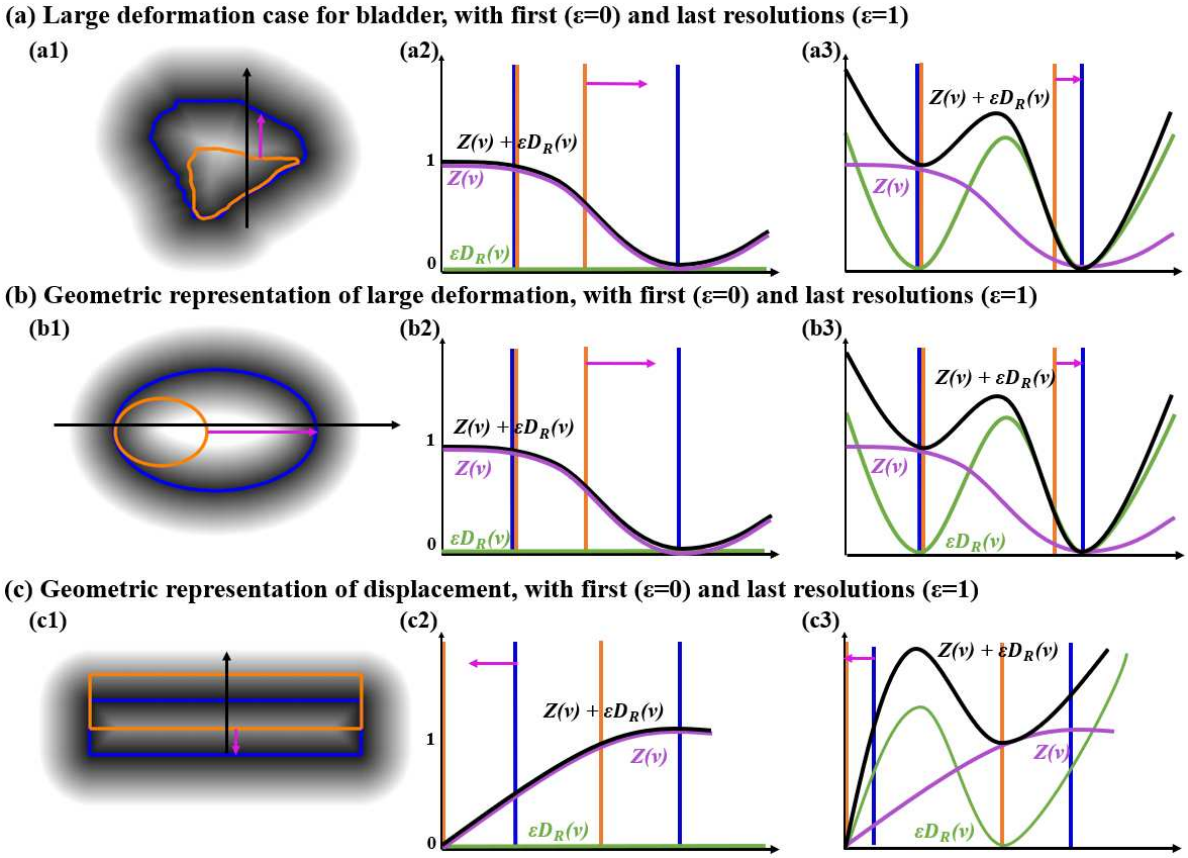

Figure 3: Improved registrations with ANACONDA+. Images in (a1), (b1), and (c1) are the same as in Figure 1, with the pink arrow showing the new path. In (a2), (b2), and (c2), the behavior is sketched for the first resolution level ($\epsilon = 0$). $Z(v)$ pushes the reference structure towards the correct minima, and $D_R(v)$ is always zero. In (a3), (b3), and (c3), the behavior is shown for the last resolution level (ϵ=1). The reference surface has been moved to the previous resolution level, close to its matching target surface, and increasing the weight of $D_R(v)$ provides a finer surface matching.

The behavior of the weight is described in Equation 4, where $R \in \mathbf{N}$ is the number

of resolution levels and $r \in [1, R]$ the current resolution level.

$$\epsilon_r = \begin{cases} 0.2 & \text{if } R = 1 \\ \frac{r-1}{R-1} & \text{if } R > 1 \end{cases} \tag{4}$$

If the resolution of the reference image is lower than the resolution of the deformation field, the optimization can just be conducted over a single resolution level. This is an outlier situation, where the coarse-to-fine behavior of the algorithm is lost and, therefore, is not recommended. Setting ϵ to 0.2 was experimentally found to provide a good compromise, preventing $D_R(v)$ from overtaking the optimization when there is a significant distance between the surfaces while it still aids in achieving improved surface matching once a satisfactory alignment of the controlling structures has been obtained.

2.3. Validation

The $Z(v)$ term was added in a pre-release version of RayStation 2023B which was then used for validation. Validation was made in a retrospective dataset of clinical pelvis reRT patients (n=46) [Research Ethics Committee (REC) approval was granted for used of patient data for research (REC reference 22/YH/0065). No patient consent was needed for retrospective use of deidentified data.]. This dataset combines data used in Nix et al. (2022) with new supplementary data. For a reRT patient, DIR was created with the reRT scan as the reference image as the intent is to use the DIR for mapping previously delivered dose to the reRT scan. In previous work to optimize reRT (Nix et al. 2022), it was found that a bespoke (non-clinical) DIR approach (b-DIR), involving biomechanical and structure-driven hybrid registration was necessary due to the type of large-deformation driven DIR failures in the original ANACONDA algorithm described above. In cases where the bladder volume in the reRT scan was much larger than of one in the original scan (full-to-empty), the reRT scan was used as target image, rather than the reference image, in ANACONDA DIR with controlling structures. The resulting DVF was then inverted and used as a starting point in a biomechanical DIR (Velec et al. 2017). The latter was found needed to correct for residual errors. We will in the following use b-DIR pathway to indicate that b-DIR is used if bladder volume difference is $> 80\%$ and otherwise ANACONDA.

Comparisons were made between the performance of the original and improved ANACONDA algorithms, as well as the b-DIR pathway, using Dice Similarity Coefficient (DSC) and Mean Distance to Agreement (MDA), on a per-ROI basis (Brock et al. 2017). Max distance to agreement (maxDTA) values were also generated but correlated with both DICE and MDA and are therefore not presented. The cohort was filtered by the performance of the original ANACONDA algorithm, to examine the impact of adding the $Z(v)$ term on the failure cases separately. Two subsets of failure cases were studied: previous bladder failures, where ANACONDA provided a DSC < 0.7 for the bladder, and previous per ROI failures, where the compared algorithm (ANACONDA or b-DIR) presented an ROI with a DSC < 0.7 or MDA $> 0.5cm$. For simplicity, we will in the

remaining refer to ANACONDA with $Z(v)$ added as ANACONDA+.

3. Results

Of the n=46 cases used for validation of ANACONDA+, 16 represented complete failure cases under the original algorithm, where the reduction in the volume of the bladder at reRT was so extreme that there were regions of gross failure on the surface of the bladder. Of the remaining 30, many exhibited compromises in other ROIs, in order to accommodate the large geometric change.

Considering the entire cohort (Figure 4 (a)), ANACONDA+ shows clear improvements in DSC for all ROIs in almost all cases when compared to ANACONDA (blue violins). When compared to the previously described b-DIR pathway, hence using b-DIR for cases when the difference in bladder volume (full-to-empty) was $> 80\%$ and otherwise ANACONDA, ANACONDA+ was comparable, with mean delta DSC above zero for all ROIs. DSC for individual organs for ANACONDA, b-DIR pathway, and ANACONDA+ are listed in Table 1. DSC is improved for ANACONDA+ (as well as for b-DIR pathway) with respect to ANACONDA for all organs ($p < 0.01$).

Table 1: DSC measures for the entire cohort presented as average and standard deviation. The best results are in bold.

Dice Similarity Coefficient			
Structure	ANACONDA	b-DIR pathway	ANACONDA+
Bladder (n=46)	0.69 ± 0.40	0.95 ± 0.02	$\mathbf{0.96 \pm 0.06}$
Rectum (n=43)	0.89 ± 0.14	0.91 ± 0.03	$\mathbf{0.94 \pm 0.07}$
Colon (n=36)	0.62 ± 0.19	0.68 ± 0.19	$\mathbf{0.76 \pm 0.19}$
Bone (n=46)	0.88 ± 0.06	0.90 ± 0.03	$\mathbf{0.91 \pm 0.04}$

There was a single case in which ANACONDA+ exhibited reduced performance for colon, compared to ANACONDA. The bladder and bones were grossly misregistered under ANACONDA. b-DIR registered the bladder well at the cost of unphysical distortions in bone, whereas ANACONDA+ was able to achieve bladder registration with minimal effect on bone. Colon was excluded from the registration controlling structures in all cases for this patient, as the images contained anatomically different sections of the organ. Hence, the small reduction of 0.19 in colon DSC was not clinically meaningful.

Considering only those 16 cases which represented failures for the original ANACONDA, shown in Figure 4 (b), the same colon failure case is present. The mean change in bladder DSC relative to original ANACONDA is +0.75, leading to a mean absolute bladder DSC of 0.93 for these difficult cases under the new algorithm, compared to just 0.18 previously. Due to the small number of failure cases, statistical

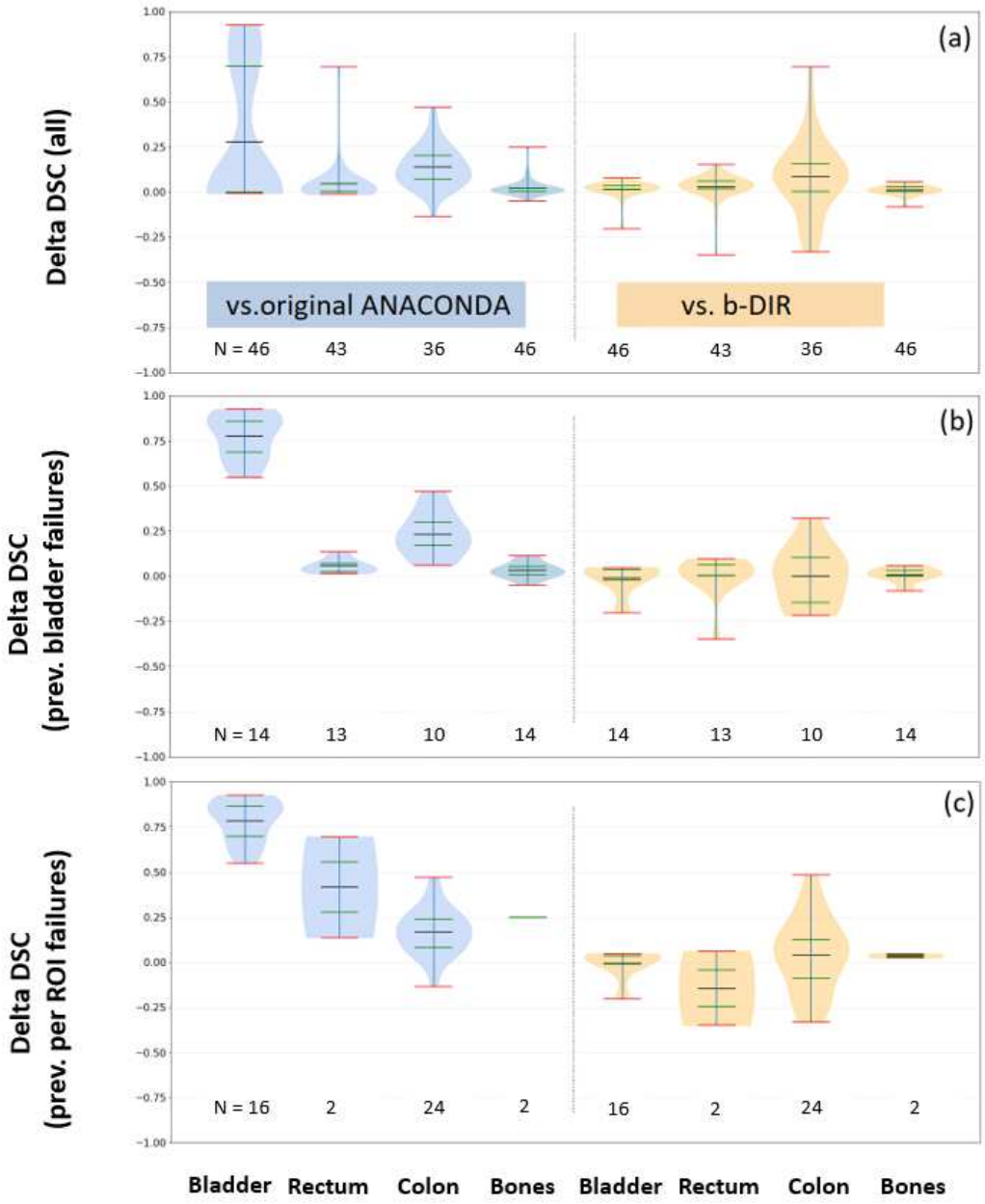

Figure 4: Changes in DSC with ANACONDA+, for four structures. Differences are shown relative to both the original ANACONDA (blue fill) and b-DIR (orange fill). Mean, inter-quartile range and extrema are shown, with kernel estimated density as a violin plot, and the number of samples, N. In (a), the full cohort is represented, in (b) only cases where bladder DIR failed with original ANACONDA (bladder DSC $<$ 0.7), and in (c), only cases where the ROI of interest had DSC $<$ 0.7 with original ANACONDA.

testing was not performed. However, the gross improvements in registration quality were clear numerically and visually.

Considering instead only individual ROIs for which DSC was less than 0.7 under

the comparison algorithm (Figure 4 (c)), DSC was almost universally improved, again with the exception of the colon in the case described above.

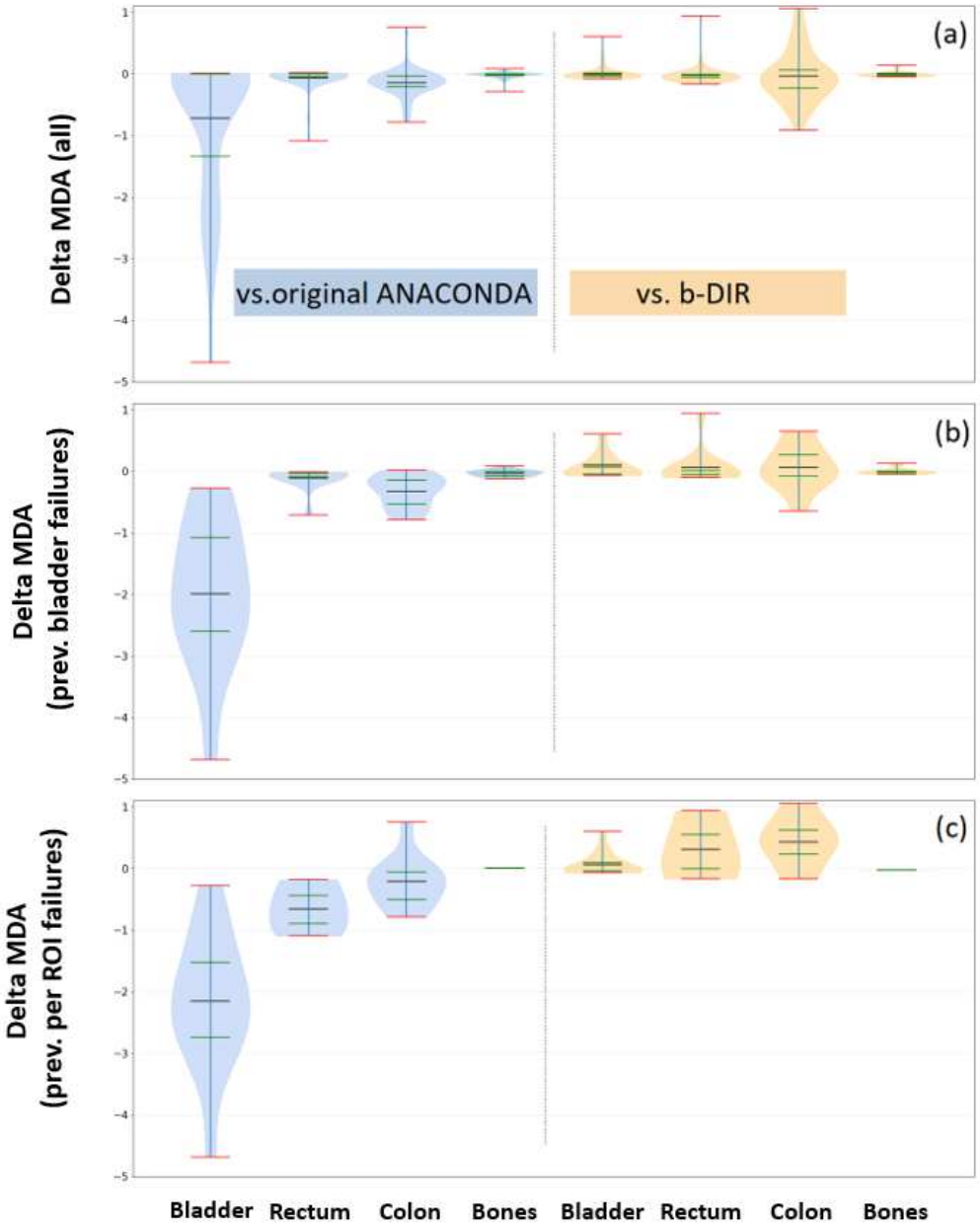

Figure 5: Changes in MDA with ANACONDA+, for four structures. Differences are shown relative to both the original ANACONDA (blue fill) and b-DIR (orange fill). Mean, inter-quartile range, and extrema are shown, with kernel estimated density as a violin plot. In (a), the full cohort is represented, in (b) only cases where bladder DIR failed with original ANACONDA (bladder DSC < 0.7), and in (c), only cases where the ROI of interest had MDA > 0.5 cm with original ANACONDA. The number of samples, N, matches Figure 5 expect (c) where N=15, 3, 11, and 1 for bladder, rectum, colon, and bones, respectively.

Similar improvements were observed for MDA (Figure 5) with universal improvements against the original algorithm and broadly similar performance to the b-DIR method, except for the aforementioned colon failure case. MDA for individual organs for ANACONDA, b-DIR pathway, and ANACONDA+ are listed in Table 2. MDA is improved for ANACONDA+ (and b-DIR pathway) with respect to ANACONDA for all organs ($p < 0.02$). Analysis of the original algorithm failure cohort showed the largest improvements with a mean reduction in MDA of 2.0 cm for the bladder and 0.44 cm for the colon.

Table 2: MDA measures for the entire cohort presented as average and standard deviation. The best results are in bold.

Mean Distance to Agreement (in cm)

Structure	ANACONDA	b-DIR pathway	ANACONDA+
Bladder (n=46)	0.80 ± 1.23	$\mathbf{0.08 \pm 0.02}$	0.08 ± 0.12
Rectum (n=43)	0.16 ± 0.31	0.10 ± 0.04	$\mathbf{0.08 \pm 0.17}$
Colon (n=36)	0.46 ± 0.39	0.36 ± 0.26	$\mathbf{0.32 \pm 0.41}$
Bone (n=46)	0.13 ± 0.07	$\mathbf{0.11 \pm 0.03}$	0.11 ± 0.05

Visual assessment of the DIRs produced was performed in the fusion view of RayStation, revealing excellent registrations for bladder and colon in previous failure cases (Figure 6 - Case A). The rigid registration is shown for reference, to illustrate the bladder volume change between CT scans. The original ANACONDA method failed to accurately register the bladder of the pelvic girdle, whilst also compromising anterior soft tissue and colon registration. ANACONDA+ achieved similar bladder registration to b-DIR, but without the spinal and anterior soft tissue distortions observed with the b-DIR method. This example was representative of the performance for previous large deformation failure cases, with the exception of the outlier case described above, shown in Figure 6 - Case B. This was the most extreme case in the cohort, exhibiting a bladder volume reduction of 85%, leading to large sections of the colon being excluded from the original scan volume, and precluding the use of colon as a controlling structure. For this example, even after excluding colon, ANACONDA was unable to capture the bladder volume change. b-DIR was only successful at the cost of severe and unphysical deformation in the spine, with large folding regions observed in the DVF (inverted elements). ANACONDA+ was successful in registering bladder, rectum, and bone, once colon was excluded.

4. Discussion

The proposed structure-based term which was integrated in the ANACONDA algorithm, resulting in ANACONDA+, seeks to overcome one of the known limitations of the original version of the ANACONDA algorithm. For large deformations, particularly in

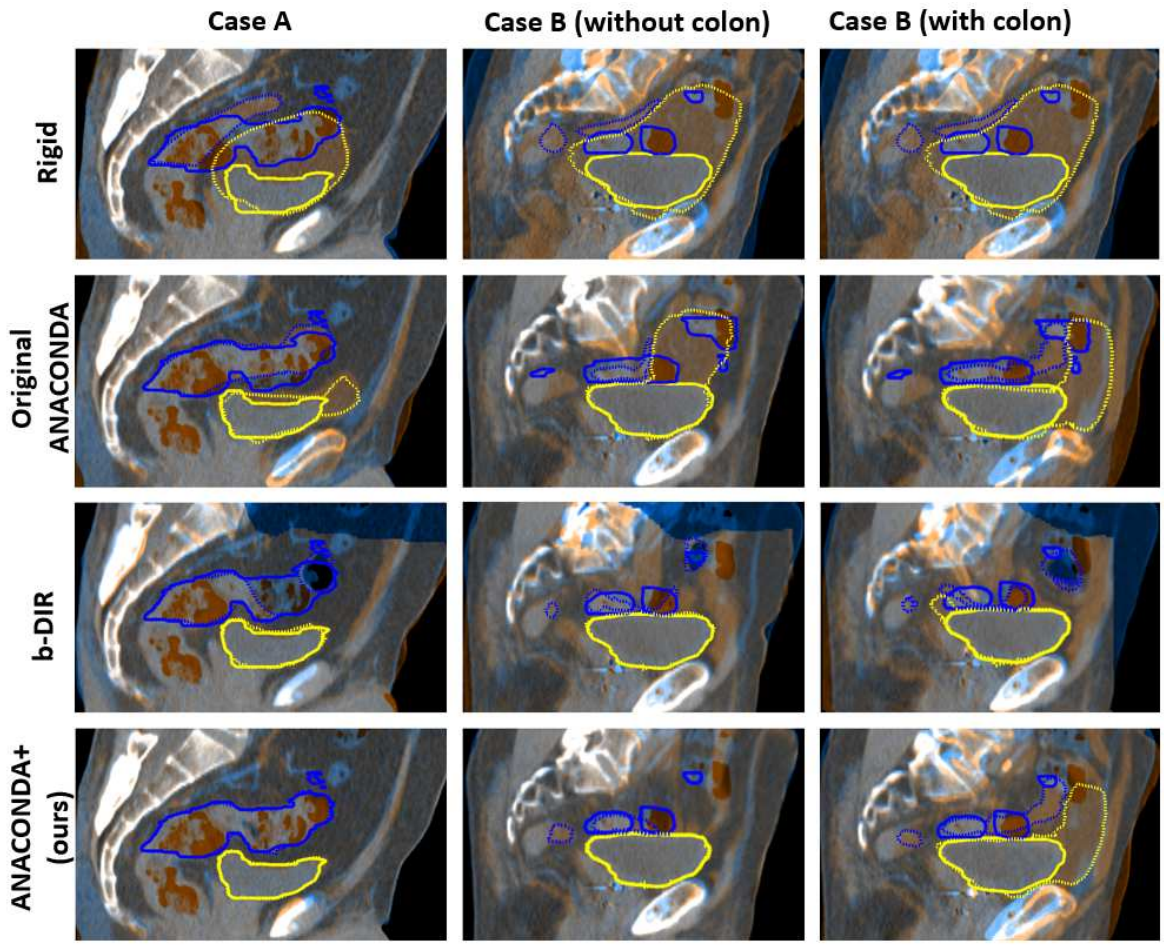

Figure 6: Examples of challenging registrations. Case A (left) shows a failure case for the original ANACONDA algorithm, which is correctly handled by both the b-DIR and proposed algorithms. The proposed algorithm provides better registration of bones. In Case B (middle), without colon as a controlling structure, the original algorithm fails, b-DIR can register the bladder, but only at the cost of severe and unphysical bone deformation. The new ANACONDA algorithm performs well, with minimal bone deformation. Case B (right) with colon, is a failure case for all tested algorithms, as there is a mismatch between the loops included in each image. Bladder (yellow) and colon (blue) ROIs are shown as solid lines (reRT) and dotted lines (deformed original).

compression, the ability of the algorithm to capture the correct surface and use it to guide the DIR is poor. By including a new term $Z(v)$, which maximizes the overlap of the soft edge volumes, we ensure that the deformed reference structure matches its corresponding target structure, avoiding the possibility that the algorithm captures the wrong surface of the ROI. By weighting the original structure-driven term $D_R(v)$ with a parameter ϵ, and optimising this parameter per-resolution level in the coarse-to-fine DIR approach, we retain the accuracy of the original approach, whilst improving robustness to gross deformations.

Validation of ANACONDA+ on a cohort of challenging reirradiation cases, where the bladder was full at original radiotherapy and empty at reRT, demonstrated the benefit of the new technique. 15 of 16 previous failure cases were successfully registered, with no noticeable loss of DIR performance for rectum, colon, or bone, based on changes in DSC and MDA. Visual assessment also showed improvements in soft tissue registration in all cases.

This work was in part motivated by previous work, where a bespoke non-clinical DIR approach was developed to overcome these limitations specifically. Whilst this approach successfully achieved registrations for these full-to-empty bladder cases, this was often at the cost of accuracy in the bones and soft tissue, limiting the usability of these DIRs. An analysis of the determinant of the inverted elements in the DVFs created using b-DIR showed that foldings occurred but they, for the specific cohort, were considered clinically insignificant based on their location. Furthermore, the technique relied on the operator to adjust parameters per-patient and the use of non-clinical steps such as inversion and interpolation of the vector field, leading to inverted elements in the DVF.

ANACONDA+ achieves comparable performance to b-DIR in terms of bladder registration, without severe registration artefacts in neighboring structures, whilst also being available as a clinically usable algorithm, which does not rely on inversion and interpolation methods. The regularisation incorporated in ANACONDA+ guarantees that there are no foldings in the DVFs.

The validation presented here is based on structure propagation overlap to show the effect of the new structure-based term. The DVF will likely give slightly distorted tissue close to the surfaces of the structures used as controlling ROIs if there are large differences between the corresponding structures in the reference and the target image. This is difficult to evaluate without the presence of reliable landmarks, which, in turn, in these regions are difficult to detect in a way that is consistent between the images. The region between, for instance, the bladder and the neighboring colon is to be considered as a sliding interface. Several different approaches on how to model sliding interfaces have been proposed, e.g., in Al-Mayah, Moseley, and Brock (2007), Vandemeulebroucke et al. (2012), and Pace, Aylward, and Niethammer (2013). Introducing sliding interfaces means that the DIR regularization needs to be relaxed and could thereby make the DVF more complex. Depending on anatomical region and application, this can be considered acceptable. Our aim is to explore this option further, for instance, if applying ANACONDA+ in an adaptive cervical cancer treatment planning setting. Within that work, it would also be of interest to do a thorough validation study, both including other metrics, inspired by Brock et al. (2017) and Nenoff et al. (2023), as well as a more diverse patient set to show under which circumstances ANACONDA+ is robust and accurate for dose accumulation. The recent review article by Murr et al. (2023) points out the importance of such studies. As we in this work show that ANACONDA+ performs equal or better than the b-DIR pathway, which has been validated for dose accumulation on a similar patient set, we already have good indications that also ANACONDA+ can be

used for dose accumulation.

In extreme deformation cases, where the anatomy included in the two scans may in fact differ, care is needed to ensure controlling ROIs used for DIR guidance are anatomically similar. ANACONDA+ has shown to be able to register these extremely different geometries, when appropriate controlling structures are chosen. It is also important to ensure that, in scenarios where the scan limits differ, the ROIs must cover only the areas of the organs included in both scans.

Recent registration techniques based on DL, such as VoxelMorph and TransMorph (Kang et al. 2022; Chen et al. 2022), have concentrated on extending the capture range to enhance the model's performance in scenarios characterized by large deformations. However, a trade-off persists between achieving a good similarity metric and eliminating folded elements. Our approach successfully achieves registration for significant deformations while adhering to the regularization prerequisite of not allowing inverted elements.

As a limitation, in contrast to DL techniques, the speed of the proposed algorithm is influenced by the number of controlling ROIs used and the complexity of the deformation. The registration durations span from seconds for straightforward cases to approximately two minutes for very challenging scenarios.

The findings of this work can be applied in DL scenarios. By adding $Z(v)$ as a term in the loss function of a DL DIR method, the registration would benefit from the structure-based guidance term while not increasing the inference time of the model.

Within traditional DIR approaches, combined brachytherapy and external beam radiotherapy where an applicator is present in only one image has gained interest. The original ANACONDA method has demonstrated suboptimal performance in comparison to alternative approaches tailored for these contexts (Rigaud, Klopp, et al. 2019; Rigaud, Cazoulat, et al. 2020). A biomechanical model-based DIR approach, described by Rigaud, Cazoulat, et al. (2020), was able to provide realistic deformations that enable plausible dose mapping. Vásquez Osorio, Kolkman-Deurloo, et al. (2015) have proposed a registration technique based on structure-based guidance designed for DIR-based dose accumulation for combined treatments. However, their method involves the integration of several transformations into a single vector field, which cannot guarantee one-to-one correspondence between images. Similarly, Fu et al. (2023) have proposed a multi-metric solution for combined treatments, emphasizing the lack of commercially available software able to handle the complex deformations common in such contexts. We believe that our proposed approach holds significant promise in the mentioned intricate registration cases and plan to further investigate its performance in such scenarios.

5. Conclusion

A structure-based term for deformable image registration which uses the soft edge representation of the structures was proposed. The aim was to achieve improved performance in the presence of large internal deformations. The new term was

implemented in an existing deformable registration algorithm as an additional term and validated for data sets where the original version of the algorithm had failed, namely extreme bladder volume reduction in reRT. The addition of the new term improved performance and was comparable to a complex non-clinical DIR pathway optimized to solve precisely that specific problem. Future work will be to explore the possible performance gains in other complex deformable image registration tasks such as dose accumulation for combined brachytherapy and external beam radiotherapy treatments. As the term is independent of the used algorithm, future studies also include the incorporation of the term into other DIR methods facing the same challenges when dealing with complex deformations, such as DL DIR.

Acknowledgments

A.A. and L.M. are currently Associate Professors funded by Yorkshire Cancer Research (award numbers L389AA and L389LM). The authors would also like to acknowledge Cancer Research UK funding for the Leeds Radiotherapy Research Centre of Excellence (RadNet; C19942/A28832) and funding for the DARIUS project (RRNPSF-Jan2100003).

References

Balakrishnan, Guha et al. (2019). "Voxelmorph: a learning framework for deformable medical image registration". In: *IEEE transactions on medical imaging* 38.8, pp. 1788–1800.

Bertelsen, Anders et al. (2011). "Radiation dose response of normal lung assessed by Cone Beam CT – A potential tool for biologically adaptive radiation therapy". In: *Radiotherapy and Oncology* 100.3, pp. 351–355.

Brock, KK et al. (2017). "Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132". In: *Medical Physics* 44.7, e43–e76.

Chen, Junyu et al. (2022). "Transmorph: Transformer for unsupervised medical image registration". In: *Medical image analysis* 82, p. 102615.

Cohen, Israel et al. (2009). "Pearson correlation coefficient". In: *Noise reduction in speech processing*, pp. 1–4.

De Vos, Bob D et al. (2019). "A deep learning framework for unsupervised affine and deformable image registration". In: *Medical image analysis* 52, pp. 128–143.

Fu, Qi et al. (2023). "Applying Multi-Metric Deformable Image Registration for Dose Accumulation in Combined Cervical Cancer Radiotherapy". In: *Journal of Personalized Medicine* 13.2, p. 323.

Ghose, Soumya et al. (June 2015). "A review of segmentation and deformable registration methods applied to adaptive cervical cancer radiation therapy treatment planning". In: *Artificial Intelligence in Medicine* 64.2.

Kang, Miao et al. (2022). "Dual-stream pyramid registration network". In: *Medical Image Analysis* 78, p. 102379.

Mannerberg, Annika et al. (2020). "Dosimetric effects of adaptive prostate cancer radiotherapy in an MR-linac workflow". In: *Radiation Oncology* 15, pp. 1–9.

Al-Mayah, A, J Moseley, and KK Brock (2007). "Contact surface and material nonlinearity modeling of human lungs". In: *Physics in Medicine & Biology* 53.1.

Murr, Martina et al. (2023). "Applicability and usage of dose mapping/accumulation in radiotherapy". In: *Radiotherapy and Oncology* 182, p. 109527.

Nenoff, Lena et al. (2023). "Review and recommendations on deformable image registration uncertainties for radiotherapy applications". In: *Physics in Medicine & Biology* 68.24, 24TR01.

Nix, Mike et al. (2022). "Dose summation and image registration strategies for radiobiologically and anatomically corrected dose accumulation in pelvic re-irradiation". In: *Acta Oncologica* 61.1, pp. 64–72.

O'Hara, Christopher J et al. (2022). "Assessment of CBCT–based synthetic CT generation accuracy for adaptive radiotherapy planning". In: *Journal of Applied Clinical Medical Physics* 23.11, e13737.

Oh, Seungjong and Siyong Kim (2017). "Deformable image registration in radiation therapy". In: *Radiation oncology journal* 35.2, p. 101.

Pace, Danielle F., Stephen R. Aylward, and Marc Niethammer (2013). "A Locally Adaptive Regularization Based on Anisotropic Diffusion for Deformable Image Registration of Sliding Organs". In: *IEEE Transactions on Medical Imaging* 32.11.

Rigaud, Bastien, Guillaume Cazoulat, et al. (2020). "Modeling complex deformations of the sigmoid colon between external beam radiation therapy and brachytherapy images of cervical cancer". In: *International Journal of Radiation Oncology* Biology* Physics* 106.5, pp. 1084–1094.

Rigaud, Bastien, Ann Klopp, et al. (2019). "Deformable image registration for dose mapping between external beam radiotherapy and brachytherapy images of cervical cancer". In: *Physics in Medicine & Biology* 64.11, p. 115023.

Sen, Anando et al. (2020). "Accuracy of deformable image registration techniques for alignment of longitudinal cholangiocarcinoma CT images". In: *Medical physics* 47.4, pp. 1670–1679.

Thörnqvist, Sara et al. (2010). "Propagation of target and organ at risk contours in radiotherapy of prostate cancer using deformable image registration". In: *Acta Oncologica* 49.7, pp. 1023–1032.

Vandemeulebroucke, Jef et al. (2012). "Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT". In: *Medical Physics* 39.2.

Vásquez Osorio, Eliana M, Inger-Karine K Kolkman-Deurloo, et al. (2015). "Improving anatomical mapping of complexly deformed anatomy for external beam radiotherapy and brachytherapy dose accumulation in cervical cancer". In: *Medical physics* 42.1, pp. 206–220.

Vásquez Osorio, Eliana M, Charles Mayo, et al. (2023). "Challenges of re-irradiation: A call to arms for physicists-and radiotherapy vendors". In: *Radiotherapy and Oncology* 182.

Veiga, Catarina et al. (2014). "Toward adaptive radiotherapy for head and neck patients: feasibility study on using CT-to-CBCT deformable registration for "dose of the day" calculations". In: *Medical physics* 41.3, p. 031703.

Velec, Michael et al. (2017). "Validation of biomechanical deformable image registration in the abdomen, thorax, and pelvis in a commercial radiotherapy treatment planning system". In: *Medical physics* 44.7, pp. 3407–3417.

Wang, Binbing et al. (2021). "Estimating the accumulative dose uncertainty for intracavitary and interstitial brachytherapy". In: *BioMedical Engineering OnLine* 20.1, pp. 1–11.

Weistrand, Ola and Stina Svensson (2015). "The ANACONDA algorithm for deformable image registration in radiotherapy". In: *Medical physics* 42.1, pp. 40–53.

White, Ingrid et al. (2019). "Geometric and dosimetric evaluation of the differences between rigid and deformable registration to assess interfraction motion during pelvic radiotherapy". In: *Physics and Imaging in Radiation Oncology* 9, pp. 97–102.