

Pseudo-CT Generation by Conditional Inference Random Forest for MRI-based Radiotherapy Treatment Planning

Axel Largent^{1,2}, Jean-Claude Nunes^{1,2}, Hervé Saint-Jalmes^{1,2,4}, Antoine Simon^{1,2}, Nicolas Perichon³, Anais Barateau^{1,2}, Chloé Hervé³, Caroline Lafond^{1,2,3}, Peter B. Greer⁵, Jason A. Dowling⁶, Renaud de Crevoisier^{1,2,3} and Oscar Acosta^{1,2}.

¹Laboratory of Signal and Image Processing (LTSI), Rennes, France

²INSERM, UMR 1099, Rennes, France

³Radiotherapy department, Centre Eugène Marquis, Rennes, France

⁴Imagery department, Centre Eugène Marquis, Rennes, France

⁵School of Mathematical and Physical Sciences, University of Newcastle, Newcastle, Australia

⁶CSIRO Australian e-Health Research Centre, Herston, Queensland, Australia

Email: jean-claude.nunes@univ-rennes1.fr

Abstract—Dose calculation from MRI is a topical issue. New treatment systems combining a linear accelerator with a MRI have been recently being developed. MRI has good soft tissue contrast without ionizing radiation exposure. However, unlike CT, MRI does not provide electron density information necessary for dose calculation. We propose in this paper a machine learning method to simulate a CT from a target MRI and co-registered CT-MRI training set. Ten prostate MR and CT images have been considered. Firstly, a reference image was randomly selected in the training set. A common space has been built thanks to affine registrations between the training set and the reference image. Multiscale image descriptors such as spatial information, gradients and texture features were extracted from MRI patches at different levels of a Gaussian pyramid and used as voxel-wise characteristics in the learning scheme. A Conditional Inference Random Forest (CIRF) modelled the relation between MRI descriptors and CT patches. For validation, test images were spatially normalized and the same descriptors were computed to generate a new pCT. Leave-one out experiments were performed. We obtained a MAE = 45.79 (pCT vs CT). Dose volume histograms inside PTV and organs at risk are in close agreement. The D98% was 0.45 % (inside PTV) and the 3D gamma pass rate (1mm, 1%) was 99.2%. Our method has better results than direct bulk assignment. And the results suggest that the method may be used for dose calculations in an MR based planning system.

Keywords—Pseudo-CT; Radiotherapy; Magnetic Resonance Imaging; Treatment planning; Random Forest;

I. INTRODUCTION

CT-scans are the main imaging modality in external beam radiotherapy. They allow the definition of tissue electron density necessary for dose calculation. Nevertheless, CT-scans use ionizing radiation and they have particularly poor soft tissue contrast. Magnetic Resonance Imaging (MRI) has much better soft tissue contrast without ionizing radiation. Moreover, MRI can provide multi-parametric information. Consequently, treatment systems combining a linear accelerator and an MRI have been recently developed (without

additional CT) [1]. They allow real-time tracking and gating of tumors, patient repositioning and dose delivery. However, dose calculation only from MRI is still an outstanding problem as this modality does not provide tissue density information.

To cope with this issue, two strategies may be distinguished from the literature. Either a pseudo-CT (pCT) is generated from the MRI to plan the dose [2]–[10], or a dose is directly generated from the MRI based on physical underpinnings [11].

Within the first group, three approaches appear. *Bulk density assignment approaches*, aim to assign homogeneous densities to different volumes of interest (or tissue classes such as air, soft tissues, bones) defined on a target MRI [5], [9], [10]. Usually, densities are obtained from a CT set which have been manually or automatically segmented. This approach is simple, but it is user dependent and time consuming. Tissue heterogeneity is not considered, and so it could not be applied to bone tumor localizations [9]. There is no multimodal registration step in this approach.

Machine learning approaches model the local relationships between intensities of co-registered CT and MR images. In prostate localization, Kapanen et al. [6], [12] used a polynomial regression to obtain pelvis Hounsfield Unit (HU) values. Johanson et al. [4], [7] employed Gaussian mixtures to predict HU values of head images. Huynh et al. [8] used a patch-based method through Structured Random Forests and auto-context to generate pCT. These methods show interesting results, but they are not yet validated for all cancer localizations. For some of them, the training set size is too small because of memory limitations of modern computers. Also, bones and air cannot be distinguished in standard MRI sequences. And consequently, sequences such as ultra-short echo time (UTE) may be required.

The atlas-based methods, involve the registration of several MRI and CT atlases with a target MRI, followed by a CT atlas fusion step [2], [3], [13], [14]. Dowling et al. [2]

applied this methodology to prostate localization. Burgos et al. [3] extended this approach to images with different fields of view and employed an auto-context scheme to refine bone localization. One advantage of these methods is that they are fully automatized. Their dose results were in close agreement with the dose from standard CT planning. The main drawback is that they may not be robust to inter-individual variability [13] and several deformable registrations are necessary. In addition, registration errors could introduce bias in dose calculation.

The second group, of *physical-based approaches* aims to compute dose directly from the hydrogen contained within the tissues, quantified by MRI, such as proposed by Demol et al. [11]. pCT generation is not necessary with this approach. However, to quantify hydrogen from MRI still an unsolved problem. Furthermore, MRI can fail to capture signal of tissues with short relaxation time (such as cortical bones and collagen).

In this paper, we propose a machine learning approach based on Conditional Inference Random Forests (CIRF) and patch partitioning to generate pCT. Ten co-registered CT and MR prostate images have been considered. We evaluated the pCT by direct voxel comparison (Mean Absolute Error: MAE and Mean Error: ME from pCT and true CT) and with a dosimetric study (Dose Volume Histogram: DVH parameters and Gamma index).

II. METHOD

Our approach consists of the main components: i) Preprocessing step to deal with intra and inter patient intensity variability of MRI; ii) Registration step to bring all the training images to the same space; iii) Feature extraction to characterize relevant MRI information; iv) Patch partitioning to simplify and improve the modeling; v) Learning step with multi CIRF to model MRI-CT relationships. vi) pCT generation step, where the trained CIRF are applied to a new target MRI to generate the pCT.

This approach is close to the one of Huynh et al. [8]. However, we perform a voxel-wise prediction and a different registration scheme (which considers patient shape variability). A lower number of features are exploited during the building model step. The whole pelvis is modelled. Moreover, a simpler RF method, the CIRF, with a different prediction scheme has been implemented.

A. Data acquisition

Ethics approval for the study protocol was obtained from the local area health ethic committee, and informed consent

was obtained from all patients. Ten patients were considered, aged between 58 and 78 years, and all were diagnosed with tumors staged between T1 and T3. CT-scans were acquired on a GE LightSpeedRT large bore scanner (2.5 mm slices) or a Toshiba Aquilion (2.0 mm slices). T2-weighted 1.6-mm isotropic SPACE sequence with field of view to cover the entire pelvis were acquired with a Siemens Skyra 3T scanner (TE: 102 ms, TR: 1200 ms, flip angle: 135°). The MR scanner was equipped with a dedicated radiation therapy flat couch and coil mounts supplied by CIVCO Medical Solutions and a laser bridge from LAP Laser. Patients were MR imaged before treatment as close as possible to the acquisition of the planning CT-scan.

B. Data pre-processing and intra patient registration

The T2-weighted scans were preprocessed with the following steps built with the Insight Segmentation and Registration Toolkit. (i) N4 Bias field correction (B-spline fitting: [160, 3, 0, 0.5]; convergence: [100 x 100 x 100, 0.001]; shrink factor: 3). (ii) Histogram equalization (levels: 1024; match points: 7, threshold at mean intensity). (iii) Filtering via gradient anisotropic diffusion (10 iterations; time step: 0.03; conductance: 1.0). Further, each CT was co-registered to their related MRI (of the same patient) with a robust symmetric rigid registration [15], followed by structure-guided deformable registration (to promote bone rigidity while allowing high-quality bladder and rectum registration) [16].

C. CT and MRI model building

The data were firstly divided into a training and validation sets (leave-one out scheme) among pre-processed images of 10 patients. Further, inter-patient registrations have been performed to get all the training images in the same space. This involved a reference MRI being randomly selected and all images in the training set were registered by affine transforms.

Three kinds of descriptors were calculated from the registered training MRI computed at three different levels of a Gaussian pyramid. The shrink factor of the pyramid (for all dimensions) was equal to $2^{(\text{Number of level} - 1)}$ and the variance of the discrete Gaussian function was equal to $\left(\frac{\text{shrink factor}}{2}\right)^2$.

Spatial information. In order to account for spatial information, we used the Cartesian coordinates of the voxels (in the common space).

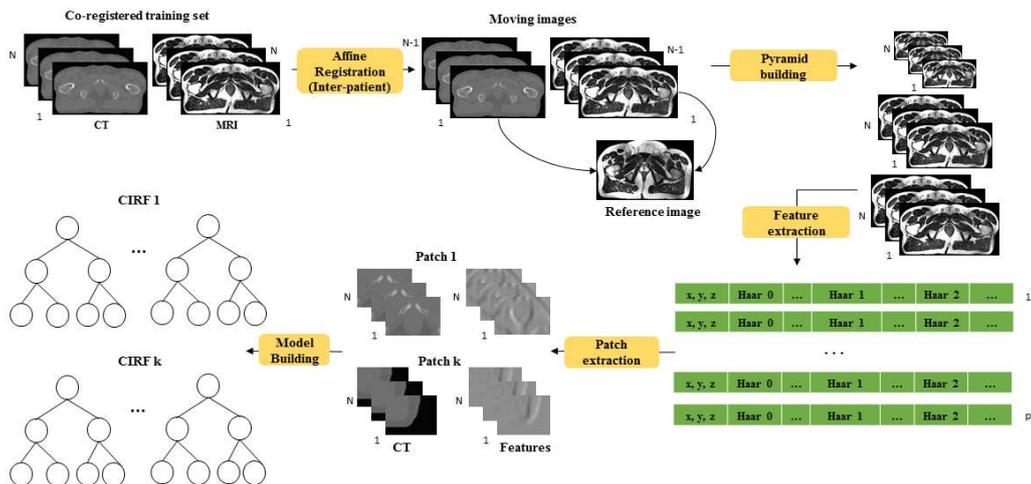


Fig. 1. Workflow model building

Gradient information. Edge information is important to distinguish different tissue classes or organs inside the images. For this purpose, we used Prewitt and Sobel edge detectors and the Histogram of Oriented Gradients (HOG). The HOG aims to extract the shape of structures inside a region by capturing information about gradients. The image is divided in small cells where magnitude and orientation of the gradients are computed. Then, histograms are built in these areas by voxel weighted votes related to their orientation and magnitude gradient. In our study, four orientation bins have been considered (0-45, 45-90, 90-135, 135-180 degrees).

Texture features. We additionally used 3D Haar-like, Local Binary Pattern (LBP) and Gabor filters. Haar-like features consist of two (or more) rectangular areas which are summed up and subtracted from each other. In our study, we used the 3D version of the Haar. The LBP principle is to compare, for each voxel, its intensity in relation to voxels belonging to its neighborhood. Thus, a binary pattern is obtained for each voxel. We considered a circular neighborhood of 8 voxels for the LBP building. Moreover, the patterns have been transformed into hexadecimal values for computational purpose. The Gabor filters are the product of Gaussians by sin or cosine functions. They aim to extract oriented and localized frequency information to describe texture. Four orientations have been employed for the wavelet computation (0, 45, 90, 135 degrees) and the Gaussian sigma has been fixed to $11/3$.

In total, 96 features were computed: 3 Cartesian coordinates (x, y, z), 9 Haar-like features, 2 LBP features, 8 Gabor filters, 4 HOG features, 4 Prewitt filters, 4 Sobel filters (for each level of the image pyramid except the Cartesian coordinate). They are calculated from a sliding window with a size of $11*11*11 \text{ mm}^3$.

Conditional Inference Random Forests (CIRF). The CIRF is a Random Forest method which aims to aggregate the results of several decision trees, built in a well-defined statistical framework. In classical decision trees, variable selection is realized with current information criteria (e.g. Gini index, entropy, etc.). These criteria tend to select variables

with many possible splits or missing values. This can lead to a decrease in model accuracy. To overcome this issue, a significance test procedure has been proposed by Horthon et al. [14] during the fitting process. This procedure selects the most significant explained variable and uses this as a better split instead of selecting the variable that maximizes only an information measure. In our study, several CIRF have been used to get the relations between MRI and CT voxels (cf. Fig. 1). We extracted patches from MRI descriptors and their related CT, and we trained the models from them.

The CIRF parameters are the following: test = univariate, p-value = 0.1, number of trees = 30, maximum depth = 12, minimum sum of weights in terminal nodes = 5, number of explanatory variables randomly chosen = 15. The size of the patches extracted from the CT and MRI descriptors was $20*20*20$ voxels. For the CIRF prediction, an aggregation scheme works by averaging observation weights extracted from each tree. Notice in classical Random Forest, the aggregation is realized only by averaging the prediction directly.

D. Model application

For this step, an MR test image was registered into the same coordinate space as the training set by affine transform. The features have been computed from this MRI. Then, patches were extracted on the feature images, and the trained CIRF were applied on them to generate the related pCT. Finally, the inverse affine transform is applied to the pCT to return in the native space (cf. Fig. 2).

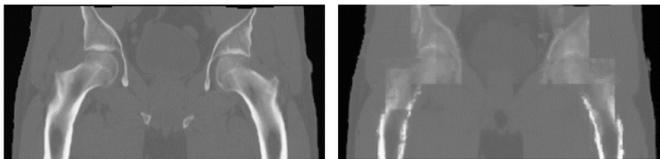
III. EXPERIMENTS AND RESULTS

A. Model validation: New MRI images

For the validation, a leave one out scheme was used on the 10 MR images. In average, we obtain a MAE = 45,79 HU and ME = 9.11 HU between the pCT and the related true CT (inside the whole pelvis). We compared these results through MAE and ME with those of an atlas-based method with the same data set and we obtained close results (cf. Table 1).

Table 1. Mean of the MAE and ME (pCT vs true CT)

	MAE	ME
CIRF	45.79 ± 10.02 HU	9.11 ± 9.91 HU
Atlas	44.42 ± 5.65 HU	15.15 ± 11.82 HU

**Fig. 2.** CT and pCT images (left to right)

B. Dosimetric study

A dosimetric study was realized to compare our method (CIRF) with a bulk density method. For the bulk density method, water density was assigned to soft tissues and the mean bone density to the pelvis bones. A VMAT treatment (1 arc, 38 fractions of 2 Gy) was planned on the CT and doses were recalculated on the pCT from both methods. The results for the DVH points D98%, D2%, V95% (Planning Target Volume: PTV) and D2% (rectum and bladder) are shown in

Table 2. For all volumes of interest, the relative dose differences are smaller than 2%. Regarding these DVH points, our method appears dosimetrically in close agreement with the standard CT planning and so clinically acceptable. In total, it depicts better results than the bulk density method (except at D2% in the PTV). To compare two dose distributions (CT vs pCT) the Gamma Index has been used [17], [18]. This index combines a dose difference (DD) criterion and a distance-to-agreement (DTA) criterion. The dose difference criterion is defined as

$$\delta(\vec{r}_e, \vec{r}_r) = D_e(\vec{r}_e) - D_r(\vec{r}_r)$$

and the distance to agreement criterion as

$$r(\vec{r}_e, \vec{r}_r) = |\vec{r}_e - \vec{r}_r|$$

where $D_e(\vec{r}_e)$ is the evaluated dose distribution and $D_r(\vec{r}_r)$ is the reference dose difference, at given grid points \vec{r}_e and \vec{r}_r .

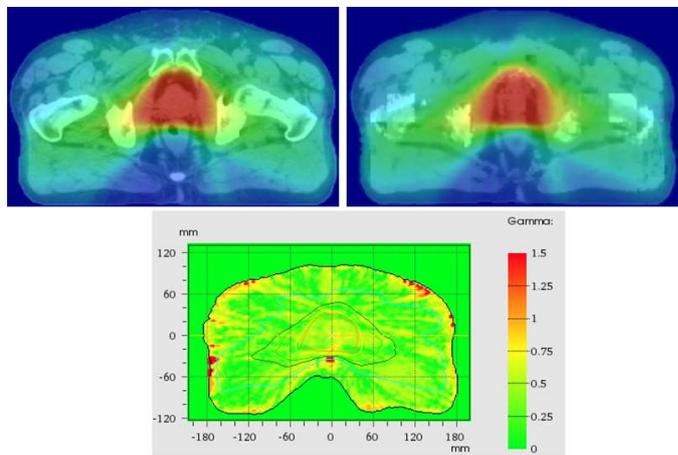
The gamma index is finally defined as

$$\Gamma(\vec{r}_e, \vec{r}_r) = \sqrt{\frac{\delta^2(\vec{r}_e, \vec{r}_r)}{\Delta D^2} + \frac{r^2(\vec{r}_e, \vec{r}_r)}{\Delta d^2}}$$

where ΔD and Δd are the acceptance criteria for the dose difference and DTA, respectively (notice these criteria could be locally or globally computed). For our dose distributions, the 3D Gamma pass rate (local, $\Delta d = 1\text{mm}$, $\Delta D = 1\%$) was equal to 99,2% and the mean of this indicator was equal to 0.425 (cf. Fig. 3).

Table 2. Absolute mean of the relative dose difference (pCT vs CT)

	D98% (PTV)	D2% (PTV)	V95% (PTV)	D2% (rectum)	D2% (bladder)
CIRF	0.45 ± 0.60	0.71 ± 1.03	0.18 ± 0.30	0.91 ± 1.89	0.37 ± 0.30
Bulk density	0.53 ± 0.28	0.61 ± 0.58	0.34 ± 0.28	0.97 ± 1.41	0.89 ± 0.78

**Fig. 3.** Dosimetric results: CT and pCT doses, gamma index (left to right)

IV. DISCUSSION AND CONCLUSION

In this paper, a multi-CIRF and patch-based approach has been proposed for generating pseudo-CT from MRI. Ten pseudo-CTs have been generated from 10 co-registered MR and CT prostate images.

Results suggest that our approach may be used for MRI-based radiotherapy treatment planning. This approach is dosimetrically more accurate than a bulk density method and is a good competitor for atlas-based method. It is fully automatic and organ of interest delineation has not required. Tissue density heterogeneity is considered and the patch-based approach copes with memory limitations of modern computers. During the model application step, with our method, only two affine registrations are necessary to generate a pCT (going to the common space and applying the inverse of the previous transform to return in the native space). For an atlas-based method, the number of deformable registrations depend on the number of training atlases (in this case 9 deformable registrations are necessary). Consequently, our approach is less computationally expensive. Moreover, no complex RF method as structured Random Forest is required [8], and few features (3D Haar, HOG, Gabor filter, ...) need to be computed during the fitting process.

We observed that CIRF has some difficulties predicting HU values of unbalanced tissue classes inside patches. Indeed, we obtained good results for soft tissues (larger classes) and poor results for bones, air cavities, and bone-soft tissue interfaces (smaller classes). To overcome this limitation, non-rigid registration methods could be used to better balance the classes. New patients could be added to the training set

thereby improving the variability. Moreover, another perspective may be to use weighted CIRF. As with all machine learning methods, our approach is heavily dependent on the number of patients and anatomical variability in the training set. To improve the robustness, we could also integrate UTE sequences to differentiate air cavities and bones. An auto-context scheme (using features computed from the CT) could be also used to distinguish air and bones. Another limitation of our method is the patch effect (aliasing) which appears in areas close to the bones (cf. Fig. 2). This patch effect is the consequence of distinct prediction errors of the patch-based CIRF in these localizations. Sliding windows (instead of this splitting process) would be a solution to suppress the patch effect, but with a much higher computational cost. Models more parsimonious than Random Forest could be also proposed, to deal with this issue, without increasing the computation time. Finally, a registration-less method exploiting patch similarity in the line of Wang et al. [19] is part of future work. Image retrieval strategies using patch similarity, which are currently used in medical image segmentation, would also extend and improve our method.

REFERENCES

- [1] J. J. W. Lagendijk *et al.*, "MRI/linac integration," *Radiother. Oncol.*, vol. 86, no. 1, pp. 25–29, Jan. 2008.
- [2] J. A. Dowling *et al.*, "Automatic Substitute Computed Tomography Generation and Contouring for Magnetic Resonance Imaging (MRI)-Alone External Beam Radiation Therapy From Standard MRI Sequences," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 93, no. 5, pp. 1144–1153, Dec. 2015.
- [3] N. Burgos *et al.*, "Robust CT Synthesis for Radiotherapy Planning: Application to the Head and Neck Region," in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Springer International Publishing, 2015, pp. 476–484.
- [4] A. Johansson, A. Garpebring, M. Karlsson, T. Asklund, and T. Nyholm, "Improved quality of computed tomography substitute derived from magnetic resonance (MR) data by incorporation of spatial information – potential application for MR-only radiotherapy and attenuation correction in positron emission tomography," *Acta Oncol.*, vol. 52, no. 7, pp. 1369–1373, Oct. 2013.
- [5] J. Lambert *et al.*, "MRI-guided prostate radiation therapy planning: Investigation of dosimetric accuracy of MRI-based dose planning," *Radiother. Oncol.*, vol. 98, no. 3, pp. 330–334, Mar. 2011.
- [6] M. Kapanen and M. Tenhunen, "T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning," *Acta Oncol.*, vol. 52, no. 3, pp. 612–618, Apr. 2013.
- [7] A. Johansson, M. Karlsson, and T. Nyholm, "CT substitute derived from MRI sequences with ultrashort echo time," *Med. Phys.*, vol. 38, no. 5, pp. 2708–2714, May 2011.
- [8] T. Huynh *et al.*, "Estimating CT Image From MRI Data Using Structured Random Forest and Auto-Context Model," *IEEE Trans. Med. Imaging*, vol. 35, no. 1, pp. 174–183, Jan. 2016.
- [9] S. J. Hoogcarpsel, J. M. Van der Velden, J. J. Lagendijk, M. van Vulpen, and B. W. Raaymakers, "The feasibility of utilizing pseudo CT-data for online MRI based treatment plan adaptation for a stereotactic radiotherapy treatment of spinal bone metastases," *Phys. Med. Biol.*, vol. 59, no. 23, p. 7383, 2014.
- [10] Y. K. Lee *et al.*, "Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone," *Radiother. Oncol.*, vol. 66, no. 2, pp. 203–216, Feb. 2003.
- [11] B. Demol, R. Viard, and N. Reynaert, "Monte Carlo calculation based on hydrogen composition of the tissue for MV photon radiotherapy," *J. Appl. Clin. Med. Phys.*, vol. 16, no. 5, Sep. 2015.
- [12] J. Korhonen, M. Kapanen, J. Keyriläinen, T. Seppälä, and M. Tenhunen, "A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer," *Med. Phys.*, vol. 41, no. 1, p. 011704, 2014.
- [13] B. Demol, C. Boydev, J. Korhonen, and N. Reynaert, "Dosimetric characterization of MRI-only treatment planning for brain tumors in atlas-based pseudo-CT images generated from standard T1-weighted MR images," *Med. Phys.*, vol. 43, no. 12, pp. 6557–6568, Dec. 2016.
- [14] J. A. Dowling *et al.*, "An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 83, no. 1, pp. e5–e11, 2012.
- [15] D. Rivest-Hénault, N. Dowson, P. B. Greer, J. Fripp, and J. A. Dowling, "Robust inverse-consistent affine CT-MR registration in MRI-assisted and MRI-alone prostate radiation therapy," *Med. Image Anal.*, vol. 23, no. 1, pp. 56–69, Jul. 2015.
- [16] D. Rivest-Hénault, P. Greer, J. Fripp, and J. Dowling, "Structure-Guided Nonrigid Registration of CT-MR Pelvis Scans with Large Deformations in MR-Based Image Guided Radiation Therapy," in *Clinical Image-Based Procedures. Translational Research in Medical Imaging*, M. Erdt, M. G. Linguraru, C. O. Laura, R. Shekhar, S. Wesarg, M. A. G. Ballester, and K. Drechsler, Eds. Springer International Publishing, 2013, pp. 65–73.
- [17] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," *Med. Phys.*, vol. 25, no. 5, pp. 656–661, 1998.
- [18] M. Wendling *et al.*, "A fast algorithm for gamma evaluation in 3D," *Med. Phys.*, vol. 34, no. 5, pp. 1647–1654, May 2007.
- [19] Z. Wang, C. Donoghue, and D. Rueckert, "Patch-based segmentation without registration: application to knee MRI," in *International Workshop on Machine Learning in Medical Imaging*, 2013, pp. 98–105.