RealDRR : rendering of realistic digitally reconstructed radiographs using locally trained image-to-image translation

RealDRR - rendering of realistic digitally reconstructed radiographs using locally trained image-to-image translation

Jennifer Dhont[1,2,3], Dirk Verellen[4,5], Isabelle Mollaert[4], Verdi Vanreusel[4], Jef Vandemeulebroucke[1,2]

1. Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel, Brussels, Belgium
2. Imec, Leuven, Belgium
3. Faculty of Medicine and Pharmaceutical Sciences, Vrije Universiteit Brussel, Brussels, Belgium.
4. Iridium Kankernetwerk, 2610 Wilrijk-Antwerp, Belgium
5. University of Antwerp, Faculty of Medicine and Health Sciences, 2610 Wilrijk-Antwerp, Belgium

Corresponding author:

Jennifer Dhont

Department of Electronics and Informatics (ETRO), VUB

Pleinlaan 9

BE-1050 Brussels

Belgium

e-mail: Jennifer.Dhont@vub.be

Keywords; deep learning, imaging, radiotherapy, motion management

Abstract

Introduction Digitally reconstructed radiographs (DRRs) represent valuable patient-specific pre-treatment training data for tumor tracking algorithms. However, using current rendering methods, the similarity of the DRRs to real X-ray images is limited, requires time-consuming measurements and/or are computationally expensive. In this study we present RealDRR, a novel framework for highly realistic and computationally efficient DRR rendering.

Materials and Methods RealDRR consists of two components applied sequentially to render a DRR. First, a raytracer is applied for forward projection from 3D CT data to a 2D image. Second, a conditional Generative Adverserial Network (cGAN) is applied to translate the 2D forward projection to a realistic 2D DRR.

The planning CT and CBCT projections from a CIRS thorax phantom and 6 radiotherapy patients (3 prostate, 3 brain) were split in training and test sets for evaluating the intra-patient, inter-patient and inter-anatomical region generalization performance of the trained framework. Several image similarity metrics, as well as a verification based on template matching, were used between the rendered DRRs and respective CBCT projections in the test sets, and results were compared to those of a current state-of-the-art DRR rendering method.

Results When trained on 800 CBCT projection images from two patients and tested on a third unseen patient from either anatomical region, RealDRR outperformed the current state-of-the-art with statistical significance on all metrics (two-sample t-test, $p < 0.05$). Once trained, the framework is able to render 100 highly realistic DRRs in under two minutes.

Conclusion A novel framework for realistic and efficient DRR rendering was proposed. As the framework requires a reasonable amount of computational resources, the internal parameters can be tailored to imaging systems and protocols through on-site training on retrospective imaging data.

Introduction

Stereotactic body radiotherapy (SBRT) has improved local control and overall survival for both primary and (oligo-)metastatic lung lesions [1-4]. A crucial factor in SBRT is geometrical accuracy, as high conformality and steep dose fall-off in the treatment plan are necessary to fulfill dose constraints on nearby organs at risk and avoid toxicity [5].

Respiratory induced motion, ranging from a few millimeters to several centimeters, can significantly degrade geometrical accuracy [6-8]. Multiple motion management strategies have therefore been developed, from dedicated treatment margins in the planning stage to real-time tumor tracking (RTTT) during treatment delivery [9-11]. A number of studies have highlighted the risk of passive motion management strategies in that they do not adapt to day-to-day and breath-to-breath variability [12-16]. Conversely, active strategies are able to adapt to this variability, but often have to rely on implanted fiducials for automatic target localization on planar X-ray images, a complication prone procedure [17-20]. For this reason, fiducial free or 'markerless' methods have been developed. Clinically implemented markerless methods are currently limited to those of the Cyberknife system (Accuray Incorporated, Sunnyvale, CA, USA) and Vero/ExacTrac system (BrainLab AG, Munich, Germany and Mitsubishi Heavy Industry, Tokyo Japan) [21,22]. For direct target localization, both use an intensity-based template matching strategy between digitally reconstructed radiograph (DRR) templates rendered from the planning (4D-)CT and real-time X-ray images. As such, the accuracy of the DRRs, e.g. their similarity to the actual X-ray images, likely influences the localization accuracy.

As DRRs provide pre-treatment information on the tumor location in different breathing phases pre-treatment, they continue to play a crucial role in the majority of newly proposed markerless tracking methods in literature [23-28]. Especially for those tracking algorithms which require training, DRRs represent valuable pre-treatment and patient-specific training data containing the target location without the need for time demanding manual annotations [29]. Recently, Hirai S. et al [30] developed such a promising markerless tracking method in which a neural network is trained on patient specific DRRs. However, while tracking errors in training were promising, results decreased when applying the

network to actual X-ray images, highlighting the importance of training on realistic and accurate DRR images [31].

The formation of an X-ray image can generally be attributed to primary and secondary processes. Primary processes refer to the attenuation of primary photons in the object, while secondary processes include scatter, beam hardening, detector response characteristics and noise. Conventional methods to render DRRs are either based on raytracing to model the primary process and/or statistical methods such as Monte-Carlo simulation to model the primary and/or secondary processes [32-36]. While raytracing is robust and can be computationally efficient, a significant amount of image pre- and post-processing is required to incorporate the effect of the X-ray source spectrum and secondary effects to render a realistic DRR. Statistical methods such as Monte-Carlo are able to render accurate virtual X-ray images when the correct device specific parameters and source spectra are available, but are computationally expensive and slow, making them a sub-optimal candidate to render large amounts of training data for deep learning-based applications [37,38]. More recently, Unberath et al proposed 'DeepDRR', a framework that aims to render realistic DRRs through sequential deep learning-based CT segmentation, forward projection, deep learning-based scatter estimation based on Monte-Carlo simulations and analytical noise injection [43,44]. A pre-trained network is available online with different source spectra and the possibility to tailor the framework to the local linear accelerator (LINAC) geometry.

In this paper we propose a novel DRR rendering framework called 'RealDRR' for fast and robust rendering of highly realistic DRRs, using a combination of raytracing and deep learning based image-to-image translation. The aim is to provide a DRR rendering framework which can be trained on site, requiring a reasonable amount of computational resources already available in most radiotherapy departments so the framework can be easily tailored to the local imaging systems and protocols to increase accuracy.

Materials and Methods

RealDRR Framework

The RealDRR framework consists of two components applied sequentially to render a virtual X-ray image or DRR, schematically shown in Fig. 1.

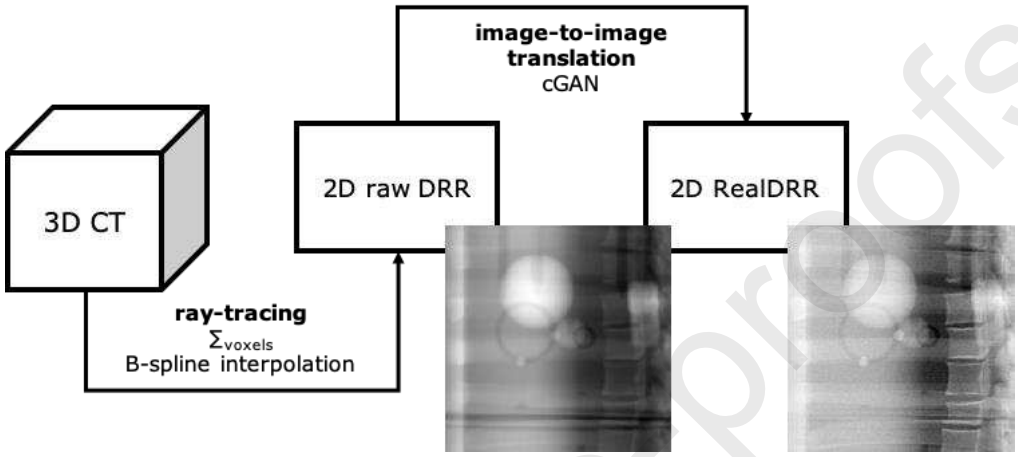

Figure 1: Schematic of the proposed RealDRR framework.

First, the transformation from 3D CT data to 2D projection image for any imaging angle is performed using an in-house developed raytracer implemented using the CUDA platform (Nvidia, Santa Clara, CA, USA) for fast forward projection through parallel computation on the graphics processing unit (GPU). Taking into account the C-arm LINAC and on-board imaging system geometry, together with the selected isocenter in the planning CT, a 512x512 pixels 2D projection image is created through the summation of interpolated CT voxel values along rays traveling from the virtual X-ray source to each detector pixel. This results in a raw 2D DRR displaying the correct geometry, but with poor pixel intensity accuracy as the source spectrum, correct linear attenuation coefficients and secondary phenomena have not been taken into account yet, as seen in Fig. 1.

As a second step, to take into account these effects in the intensity values, an image-to-image translation is performed using a U-net convolutional neural network architecture. The latter is part of a cGAN consisting of the previously mentioned U-net generator and a discriminator, as illustrated in

Supplementary Fig. 1 and described in detail by Isola et al [39]. The input to the U-net generator is a 2D forward projection image or raw DRR of size 512x512 pixels which is rescaled to 256x256 pixels, while the output is a realistic DRR of the same size, rescaled to 512x512 pixels. For training, the cGAN framework uses a structured loss function combining the GAN objective and a conventional L1-norm to obtain sharp and realistic results with minimal image artefacts [40,41].

$$\text{Loss} = \text{GAN} + \lambda \, . \, L_1 \quad (\lambda = 50)$$

Both the network topologies and hyperparameters were optimized for the specific application of realistic DRR rendering. Specifically, the number of layers and kernel sizes in the generator were adapted resulting in the network architectures illustrated in Supplementary Fig. 1. Optimization was performed using minibatch (size = 4) stochastic gradient descent (SGD) and Adam solver (learning rate = 0.0002, $\beta_1 = 0.9$) through alternating gradient descent steps on the discriminator and generator, as described by Goodfellow et al [42]. While the original pix2pix framework applies the discriminator on an image patch level, in this study the full image was used as input to avoid tiling artifacts. A λ value of 50 empirically resulted in the best trade-off between image sharpness and absence of image artefacts related to the GAN objective.

Data

To develop, validate the hyperparameters and perform an initial evaluation of the RealDRR framework, imaging data from an anthropomorphic CIRS thorax phantom (CIRS, Norfolk, VA, USA) was used. This included a reconstructed 3D-CT as would be used for treatment planning, from which the DRRs are rendered (Toshiba system, 512x512x328 pixels, 1.07x1.07 mm^2 in-plane resolution, 1 mm slice thickness) and 895 X-ray projection images from a single cone-beam CT (CBCT) acquired on a conventional RT system as the objective images (Varian Truebeam, full-fan mode, 1024x768 pixels, 0.39x0.39 mm^2 resolution, 100kV-15mA-20ms). Pre-processing of CBCT projection images consisted of cropping each image with respect to the image center to 512x512 pixels and image normalization between [0,1] using the minimum and maximum pixel intensities over the entire dataset.

To evaluate the clinical performance of the RealDRR framework, imaging data from 6 patients treated with radiotherapy was used. Patients 1-3 were treated in the pelvic region and the data per patient included the 3D planning CT (Toshiba system, 512x512x119 pixels, 1.12x1.12 mm^2 in-plane resolution, 3 mm slice thickness) and 895 CBCT projection images (Varian Truebeam, full-fan mode, 1024x768 pixels, 0.39x0.39 mm^2 resolution, 125:140kV-72:85mA-25ms) acquired during patient positioning on the radiotherapy treatment system. Patients 4-6 were treated with stereotactic radiotherapy (STX), wearing a thermoplastic fixation mask throughout treatment and image acquisition. Similarly, but with the use of different imaging protocols, the data per patient included the 3D planning CT (Toshiba system, 512x512x399 pixels, 0.78x0.78 mm^2 in-plane resolution, 1 mm slice thickness) and 895 CBCT projection images (Varian Truebeam, full-fan mode, 1024x768 pixels, 0.39x0.39 mm^2 resolution, 100kV-15mA-20ms) acquired during patient positioning. For all patients, pre-processing of CBCT projection images was performed identical to the CIRS data.

Training and evaluation

The RealDRR network topologies and (hyper)parameters were optimized and the image-to-image translation network was trained using a training and validation set of 800 CBCT projection images from the CIRS dataset. The remaining 95 projection images were used as a test set to perform an initial evaluation of the framework performance.

Next, keeping the optimized network topology and hyperparameters fixed, the framework was trained on 800 randomly selected CBCT projection images from patients 1-2 (pelvic), and tested on 100 unseen CBCT projection images from the same patients. The same trained network was also tested on 100 randomly selected CBCT projection images of patient 3 (pelvic) and patients 4-6 (STX), respectively. Similarly, the framework was trained on 800 randomly selected CBCT projection images from patients 4-5 (STX) and tested on 100 unseen projection images of the same patients, as well as on 100 projection images of patient 6 (STX) and patients 1-3 (pelvic), respectively. This to evaluate the intra-patient, inter-patient and inter-anatomical region (including different imaging protocols) generalization performance of the trained framework.

Evaluating the quality of synthetic 2D images in terms of realistic appearance is a complex problem. Evaluation was performed by comparing the rendered DRR images to their respective ground-truth CBCT projection images using the mean absolute error (MAE), the normalized root-mean-square error (NRMSE), the structural similarity index (SSIM) and the peak signal-to-noise ratio (PSNR). Several of these metrics assume a perfectly identical anatomy between the DRRs, rendered from the planning-CT, and the CBCT projection images which are acquired on a different day. Therefore, pixel intensity profiles along the central horizontal and central vertical axis as well as intensity histograms of both rendered DRR images and corresponding ground-truth CBCT projection images were rendered for qualitative evaluation. The performance of the proposed framework was also evaluated through a template matching analysis between synthetic DRR templates and their ground-truth kV CBCT projections, explained in detail in Appendix A (supplementary material). Lastly, as the cGAN is not necessarily a spatially invariant transformation, the location of anatomical landmarks in images was evaluated before and after passing through the trained networks, as explained in Appendix B (supplementary material).

All results were benchmarked against the state-of-the-art DRR rendering framework 'DeepDRR' proposed by Unberath et al [43,44]. The pre-trained network available online was implemented and tailored to the CBCT imaging system by selecting the best matching source spectra and implementing the correct LINAC and imaging system geometry. Per patient (1-6), 100 DRR images were rendered, cropped to 512x512 pixels and normalized between [0,1] using the minimum and maximum pixel intensities over the entire dataset.

Results

Training the image-to-image translation component of the RealDRR framework for 200 epochs (batchsize = 4) lasted 23h on a desktop PC equipped with an Nvidia GeForce GTX 1080/PCIe/SSE2 GPU (Nvidia, Santa Clara, CA, USA), using a dataset of 800 raw DRR and CBCT projection image pairs. On the same desktop PC, rendering 100 realistic DRRs (512x512 pixels, 0.39x0.39 mm^2

resolution) from 100 different imaging angles through the RealDRR framework took on average 91.4s (raytracing: 82.6s, image-to-image translation: 8.8s).

In an initial evaluation of the framework performance, after training on 800 images from the CIRS dataset and testing on 95 unseen images, the MAE ± 1SD, NRMSE, SSIM and PSNR equaled 0.01 ± 0.01, 0.01, 0.98 and 39.70, respectively. Figure 2 shows a representative example of the performance during testing, including a 2D raw DRR rendered through raytracing which is used as input to the image-to-image translation network, the respective RealDRR output and ground-truth CBCT projection image.

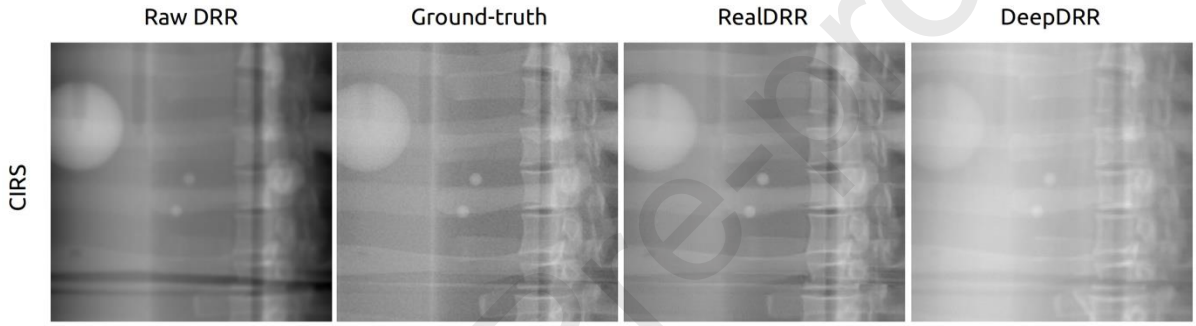

Figure 2: A representative example of the RealDRR performance when applied on the CIRS test set, after training the image-to-image translation network on 800 CBCT projection images from the same phantom.

Using clinical patient data, the performance of the RealDRR framework was evaluated in terms of intra- and inter-patient, as well as inter-anatomical region generalization. Table 1 comprises all quantitative performance results for the different test sessions, together with a comparison to the DeepDRR framework. The accuracy of the raw DRRs, rendered through raytracing but without image-to-image translation, are also included to show the added value of the image-to-image translation part of the framework. All evaluation was performed by comparing the rendered DRR images in each test set to their respective ground-truth CBCT projection images using several metrics.

	MAE ± 1SD [max]	NRMSE	SSIM	PSNR
Trained on patient 1-2 (pelvis)				
Test patient 1-2 (pelvis)	0.01 ± 0.00 [0.28]	0.05	0.98	26.75
Test patient 3 (pelvis)	0.02 ± 0.02 [0.19]	0.03	0.96	31.03
Test patient 4-6 (STX)	0.05 ± 0.11 [0.75]	0.13	0.92	18.59
Raw DRR patient 1-3 (pelvis)	0.27 ± 0.12 [0.65]	0.31	0.72	10.41
DeepDRR patient 1-3 (pelvis)	0.06 ± 0.05 [0.48]	0.11	0.78	22.25
Trained on patient 4-5 (STX)				
Test patient 4-5 (STX)	0.01 ± 0.01 [0.36]	0.02	0.99	35.58
Test patient 6 (STX)	0.03 ± 0.04 [0.32]	0.05	0.93	26.89
Test patient 1-3 (pelvis)	0.02 ± 0.04 [0.69]	0.05	0.94	25.93
Raw DRR patient 4-6 (STX)	0.38 ± 0.09 [0.71]	0.43	0.68	8.24
DeepDRR patient 4-6 (STX)	0.23 ± 0.12 [0.75]	0.28	0.86	11.82

Table 1: Quantitative results of the RealDRR framework and comparison to the raw DRRs and DeepDRR framework per anatomical region using the mean absolute error (MAE) ± 1 standard deviation (SD) [maximum MAE], the normalized root-mean-square error (NRMSE), the structural similarity index (SSIM) and the peak signal-to-noise ratio (PSNR). All image intensities were normalized between [0,1].

When the network is tested on unseen imaging data coming from the same patients of whom data was used for training, the RealDRR framework shows good (intra-patient) generalization performance. Taking into account that image intensities range between [0,1], pixel intensity accuracy is high with MAE ± 1SD equal to 0.01 ± 0.00 and 0.01 ± 0.01 for the pelvic and cranial region, respectively. Further, in both anatomical regions, the SSIM ranges between 98% and 99%.

When trained and tested on different patients but representing the same anatomical region, MAE ± 1SD and SSIM equal 0.02 ± 0.02 and 96% for the pelvic region and 0.03 ± 0.04 and 93% for the cranial

region. While the inter-patient generalization performance is slightly lower than intra-patient, the RealDRR framework shows better results compared to the DeepDRR framework for the respective anatomical regions with statistical significance on all metrics (two-sample t-test, $p < 0.05$). Fig. 3 shows a representative example of the inter-patient generalization performance of RealDRR for both the pelvic and cranial region. Per anatomical region, the figure includes a 2D forward projection image or raw DRR rendered through raytracing which is used as input to the image-to-image translation network, the respective RealDRR output and ground-truth CBCT projection image.

The decrease in accuracy continues when evaluating the inter-anatomical region generalization performance, which also includes the generalization between different imaging protocols. MAE ± 1SD equal 0.05 ± 0.11 and 0.02 ± 0.04 when trained/tested on the pelvic/cranial and cranial/pelvic region, respectively. Nevertheless, for the respective anatomical regions, RealDRR continues to outperform the DeepDRR framework on all metrics with statistical significance (two-sample t-test, $p < 0.05$).

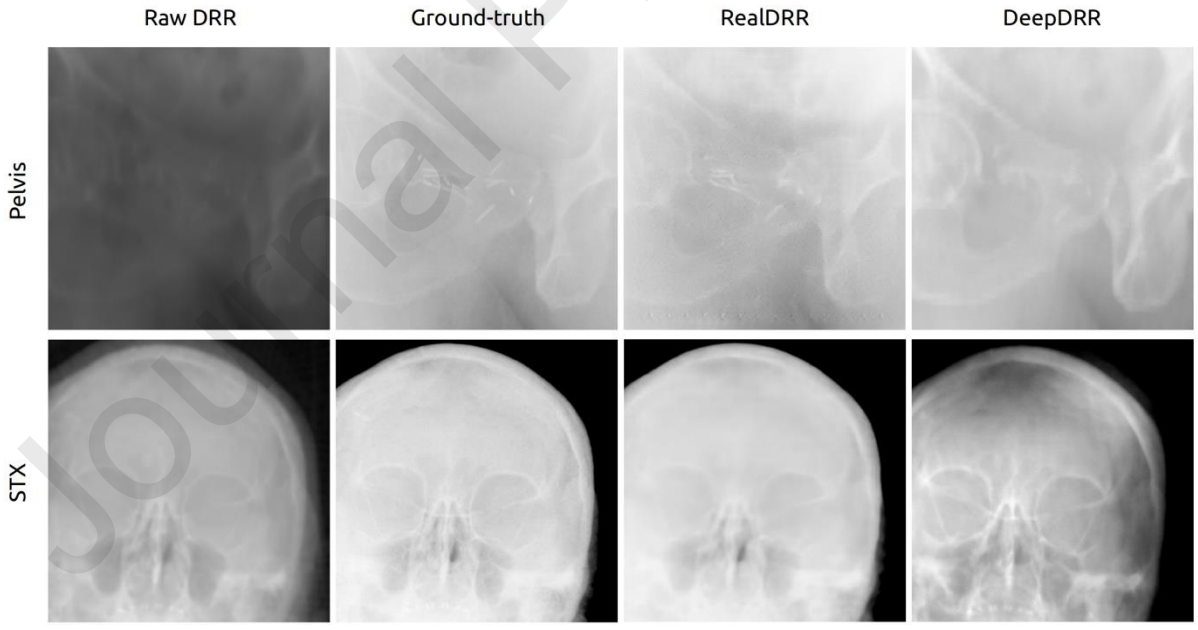

Figure 3: A representative example of the RealDRR inter-patient generalization performance within the same anatomical region for the pelvic (top) and cranial (bottom) region.

From the RealDRR and corresponding ground-truth CBCT projection images presented in Fig. 2 and Fig. 3, pixel intensity profiles along the central horizontal and central vertical axis are plotted in supplementary Fig. 2, while supplementary Fig. 3 includes the pixel intensity histograms of the same images.

Table 2 contains the results of the template matching analysis. Using templates generated through the proposed RealDRR framework results in a significant improvement in matching accuracy with the ground-truth kV CBCT projections compared to raw DRR (two-sample t-test, $p < 0.001$) or DeepDRR (two-sample t-test, $p < 0.05$).

	50x50	**55x55**	**60x60**	**65x65**
Pelvis				
RealDRR	1.5 ± 3.1	1.3 ± 2.4	1.1 ± 1.5	1.0 ± 1.0
Raw DRR	48.6 ± 24.5	45.2 ± 43.2	44.1 ± 24.4	41.2 ± 26.0
DeepDRR	3.3 ± 6.5	2.4 ± 4.9	1.6 ± 1.4	1.7 ± 1.7
Head				
RealDRR	1.81 ± 1.2	1.9 ± 1.2	1.8 ± 1.1	1.9 ± 0.9
Raw DRR	65.7 ± 19.2	62.4 ± 19.6	59.2 ± 20.5	55.3 ± 22.1
DeepDRR	29.4 ± 4.0	16.7 ± 6.0	10.5 ± 5.7	11.5 ± 5.6

Table 2: Results of the template matching analysis, performed between templates rendered from the synthetic images through raytracing (raw DRR), the proposed framework (RealDRR) and a state-of-the-art method (DeepDRR) and the ground-truth kV CBCT projections. Results are presented as the MAE ± 1 SD in millimeter between the ground-truth location and matched template location averaged over the different locations and 100 images, for the different template sizes.

Results of the spatial invariance analysis are included in the supplementary materials (Appendix B) and show that the geometrical location of anatomical landmarks is not always maintained when the network is trained on a single patient. This because the cGAN learns to correct for the anatomical differences between the planning CT and the CBCT acquisition several days later, and as such

introduces geometrical changes. However, when trained on the combined data from two patients, geometrical locations remain stable.

Discussion

In this paper, a novel framework for realistic DRR rendering called RealDRR was proposed, consisting of a combination of fast raytracing and deep learning-based image-to-image translation. The experiments in this study have shown that training the framework to reach state-of-the-art results requires a reasonable amount of computational resources which are already available in most radiotherapy departments. This makes it feasible to train the framework on site and with retrospective on site imaging data to easily tailor the internal parameters to the local CT and CBCT imaging systems and imaging protocols, resulting in highly realistic DRRs. Once trained, the framework is able to render 100 highly realistic DRRs in under two minutes, making it an ideal tool to render large amounts of patient-specific pre-treatment training data with automatic annotations.

To evaluate the clinical performance of the RealDRR framework, imaging data from three diverse anatomical regions was used, including an anthropomorphic thorax phantom and 6 patients treated with radiotherapy in the pelvic and cranial region. Several experiments were performed to evaluate the intra-patient, inter-patient and inter-anatomical region generalization performance. In all scenarios and for every anatomical region, the RealDRR framework rendered highly realistic DRRs, resulting in better performance both quantitatively and qualitatively compared to the current state-of-the-art. The results with highest accuracy were obtained when the framework was both trained and tested on imaging data coming from the same patients, but this solution is rather unpractical in a clinical setting. Therefore, training and applying the RealDRR framework is recommended per anatomical region to obtain the optimal balance between accuracy and practical applicability, as the framework can be tailored to common anatomy-specific protocols by training on retrospective imaging data from patients treated in the same anatomical region. Further, as it was shown that training the framework on a single patient is not spatially invariant as the cGAN learns to correct for possible anatomical differences between the planning CT and CBCT acquisition, it is highly recommended that the training set contains retrospective data from at least two patients.

Previous studies have addressed the speed of DRR rendering when applying raytracing, taking advantage of the powerful parallel computation capability of the GPU as was done in this study [45-49]. However, most of these studies aimed to render DRRs solely for 3D/2D image registration for patient positioning, focusing on the appearance of bony anatomy resulting in low soft tissue contrast and less realistic images.

Other studies have specifically addressed the realistic appearance of DRR images. Moore et al [50] replaced the use of rays with 3D pencil beams to increase realness and avoid aliasing artefacts. Staub et al [51] aimed to render realistic DRRs solely based on first principles, e.g. the fundamental physical principles underlying image formation to replicate absolute image intensities. By first removing secondary effects such as scatter, beam hardening and veiling glare, a CT number to linear attenuation coefficient function was determined before applying forward raytracing. While both methods were able to render highly realistic DRRs, several phantom measurements were necessary to quantify the system-specific secondary effects before they could be added in a post-processing step. With the RealDRR framework, these measurements are avoided completely by training the network with actual X-ray images so the framework learns how to correct for secondary effects by itself. Unberath et al [44] proposed one of the first DRR rendering frameworks using deep learning (DeepDRR) and provide the framework pre-trained online for fast and practical implementation without requiring expert knowledge. While the network is able to render realistic DRRs without requiring any measurements or training and is able to render a large number of DRRs in a reasonable time frame, the framework is not designed for detailed tailoring to on site imaging protocols and systems as it requires Monte Carlo simulations to render training data. As such, it renders more generic images resulting in less accurate intensity values compared to RealDRR, making it less appropriate to render tumor templates or pre-treatment training data for tracking algorithms.

To evaluate the extent of realistic appearance, multiple metrics as well as qualitative comparisons using intensity profiles and histograms were used. The latter have shown improvements can be made regarding under- and overestimation of pixel intensity values, either through post-processing in case of

fixed systematic deviations per anatomical region or by further improving the image-to-image translation part of the framework. To avoid interference due to anatomical changes between the planning CT and the CBCT projections, the use of CycleGANs and unpaired training could provide an interesting alternative [52].

The performance of the proposed framework was also evaluated by applying the synthetic images in one of the intended applications, which highlighted the possible improvements the framework can accomplish. As the analysis in deep learning-based applications was outside the scope of this paper, the code of the RealDRR framework will be made publicly available so novel markerless tumor tracking methods can be trained and/or re-evaluated using RealDRR images. As DRRs provide valuable pre-treatment and patient-specific information, we believe RealDRR will prove to be a helpful tool in facilitating a number of novel applications in the field of radiotherapy and personalized medicine.

References

[1] Haque W. et al. “Stereotactic body radiation therapy versus conventionally fractionated radiation therapy for early stage non-small cell lung cancer”. Radiotherapy and Oncology (2018). doi: 10.1016/j.radonc.2018.07.008.

[2] McGarry R.C. et al. “Stereotactic body radiation therapy of early stage non-small-cell lung carcinoma: Phase I study”. International Journal of Radiation Oncology Biology Physics (2005). doi:10.1016/j.ijrobp.2005.03.073.

[3] Lagerwaard F.J. et al. “Outcomes of Risk-Adapted Fractionated Stereotactic Radiotherapy for Stage I Non-Small-Cell Lung Cancer”. International Journal of Radiation Oncology Biology Physics (2008). doi: 10.1016/j.ijrobp.2007.10.053.

[4] Timmerman R.D. et al. “Long-term Results of RTOG 0236: A Phase II Trial of Stereotactic Body Radiation Therapy (SBRT) in the Treatment of Patients with Medically Inoperable Stage I Non-Small Cell Lung Cancer”. International Journal of Radiation Oncology*Biology*Physics (2014). doi: 10.1016/j.ijrobp.2014.05.135.

[5] Timmerman R.D. et al. “Excessive toxicity when treating central tumors in a phase II study of stereotactic body radiation therapy for medically inoperable early-stage lung cancer”. Journal of Clinical Oncology (2006). doi: 10.1200/JCO.2006.07.5937.

[6] Seppenwoolde Y. et al. “Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy”. International Journal of Radiation Oncology Biology Physics (2002). doi: 10 .1016/S0360-3016(02)02803-1.201

[7] Cusumano D. et al. “Predicting tumour motion during the whole radiotherapy treatment: a systematic approach for thoracic and abdominal lesions based on real time MR”. Radiotherapy and Oncology (2018). doi: 10.1016/j.radonc.2018.07.025.

[8] Waghorn B.J. et al. “Analyzing the impact of intrafraction motion: Correlation of different dose metrics with changes in target D95”. Medical Physics (2011). doi: 10.1118/1.3605633.

[9] Langen K. M. et al. Organ motion and its management. 2001. doi: 10.1016/S03603016(01)01453-5.

[10] Keall P.J. et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. 2006. doi: 10.1118/1.2349696.

[11] Depuydt T. et al. "Geometric accuracy of a novel gimbals-based radiation therapy tumor tracking system". Radiotherapy and Oncology (2011). doi: 10.1016/j.radonc.2011.01.015.

[12] Guckenberger M. et al. "Is a Single Respiratory Correlated 4D-CT Study Sufficient for Evaluation of Breathing Motion?" International Journal of Radiation Oncology Biology Physics (2007). doi: 10.1016/j.ijrobp.2006.11.025.

[13] Ge J. et al. "Planning 4-dimensional computed tomography (4DCT) cannot adequately represent daily intrafractional motion of abdominal tumors". International Journal of Radiation Oncology Biology Physics (2013). doi: 10.1016/j.ijrobp.2012.09.014.

[14] Lens E. et al. "Differences in respiratory-induced pancreatic tumor motion between 4D treatment planning CT and daily cone beam CT, measured using intratumoral fiducials". Acta Oncologica (2014). doi: 10.3109/0284186X.2014.905699.

[15] Dhont J. et al. "The long- and short-term variability of breathing induced tumor motion in lung and liver over the course of a radiotherapy treatment". Radiotherapy and Oncology (2018). doi: 10.1016/j.radonc.2017.09.001.

[16] Steiner E. et al. "Both four-dimensional computed tomography and four-dimensional cone beam computed tomography under-predict lung target motion during radiotherapy". Radiotherapy and Oncology 135 (2019), pp. 65–73. doi: 10.1016/j.radonc.2019.02.019.

[17] Yousefi S. et al. "Complications of thoracic computed tomography guided fiducial placement for the purpose of stereotactic body radiation therapy". Clinical Lung Cancer (2007). doi: 10.3816/CLC.2007.n.002.

[18] Bhagat N. et al. "Complications associated with the percutaneous insertion of fiducial markers in the thorax". CardioVascular and Interventional Radiology (2010). doi:10.1007/s00270-010-9949-0.

[19] Seppenwoolde Y. et al. “Treatment precision of image-guided liver SBRT using implanted fiducial markers depends on marker-tumour distance”. Physics in Medicine and Biology (2011). doi:10.1088/0031-9155/56/17/001.

[20] Bahig H. et al. “Predictive parameters of cyberknife fiducial-less (XSight Lung) applicability for treatment of early non-small cell lung cancer: A single-center experience”. International Journal of Radiation Oncology Biology Physics (2013). doi: 10 . 1016 / j .ijrobp.2013.06.2048.

[21] Fu D. et al. “Xsight lung tracking system: A fiducial-less method for respiratory motion tracking”. Treating Tumors that Move with Respiration. 2007. doi:10.1007/978-3-540-69886-9_26.

[22] Ziegler M. et al. “Performance of Makerless Tracking for Gimbaled Dynamic Tumor Tracking”. Zeitschrift fur Medizinische Physik (2019). doi: 10.1016/j.zemedi.2019.10.003.

[23] Schweikard A. et al. Respiration tracking in radiosurgery without fiducials. 2005. doi: 10.1002/rcs.38.

[24] Li R. et al. “A feasibility study of markerless fluoroscopic gating for lung cancer radiotherapy using 4DCT templates”. Physics in Medicine and Biology (2009). doi: 10.1088/0031-9155/54/20/N03.

[25] Yang Y. et al. “A novel markerless technique to evaluate daily lung tumor motion based on conventional cone-beam CT projection data”. International Journal of Radiation Oncology Biology Physics (2012). doi: 10.1016/j.ijrobp.2011.11.035.

[26] Teske H. et al. “Real-time markerless lung tumor tracking in fluoroscopic video: Handling overlapping of projected structures”. Medical Physics (2015). doi: 10.1118/1.4917480.

[27] Chun Chien Shieh et al. “A Bayesian approach for three-dimensional markerless tumor tracking using kV imaging during lung radiotherapy”. Physics in Medicine and Biology (2017). doi:10.1088/1361-6560/aa6393.

[28] Mueller M. et al. “The first prospective implementation of markerless lung target tracking in an experimental quality assurance procedure on a standard linear accelerator”. Physics in Medicine and Biology (2020). doi: 10.1088/1361-6560/ab5d8b.

[29] Takahashi W. et al. “Real-time markerless tumour tracking with patient-specific deep learning using a personalised data generation strategy: proof of concept by phantom study”. The British Journal of Radiology (2020). doi: 10.1259/bjr.20190420.

[30] Hirai R. et al. “Real-time tumor tracking using fluoroscopic imaging with deep neural network analysis”. Physica Medica (2019). doi: 10.1016/j.ejmp.2019.02.006.

[31] Mori S. et al. Comments on “Novel real-time tumor-contouring method using deep learning to prevent mistracking in Xray fluoroscopy” by Terunuma et al. 2018. doi: 10.1007/s12194-018-0447-4.

[32] Sherouse G.W. et al. “Computation of digitally reconstructed radiographs for use in radiotherapy treatment design”. International Journal of Radiation Oncology, Biology, Physics (1990). doi: 10.1016/0360-3016(90)90074-T.

[33] Milickovic N. et al. “CT imaging based digitally reconstructed radiographs and their application in brachytherapy”. Physics in Medicine and Biology (2000). doi: 10.1088/0031-9155/45/10/305.

[34] Li X. et al. “Digitally reconstructed radiograph generation by an adaptive Monte Carlo method”. Physics in Medicine and Biology (2006). doi: 10.1088/0031-9155/51/11/004.

[35] Shen A. et al. “Point-based digitally reconstructed radiograph”. Proceedings - International Conference on Pattern Recognition. 2008. doi: 10.1109/icpr.2008.4761753.

[36] Pallotta S. et al. “A simple method to test geometrical reliability of digital reconstructed radiograph (DRR)”. Journal of Applied Clinical Medical Physics (2010). doi: 10.1120/jacmp.v11i1.3128.

[37] Zhao W. et al. “Incorporating imaging information from deep neural network layers into image guided radiation therapy (IGRT)”. Radiotherapy and Oncology (2019). doi: 10.1016/j.radonc.2019.06.027.

[38] Zhao W. et al. “Toward markerless image-guided radiotherapy using deep learning for prostate cancer”. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2019. doi: 10.1007/978-3-030-32486-5_5.

[39] Isola P. et al. “Image-to-image translation with conditional adversarial networks”. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. 2017. doi: 10.1109/CVPR.2017.632. arXiv: 1611.07004.

[40] Pathak D. et al. “Context Encoders: Feature Learning by Inpainting”. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2016. doi: 10.1109/CVPR.2016.278. arXiv: 1604.07379.

[41] Zhang R. et al. “Colorful image colorization”. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2016. doi: 10.1007/978-3-319-46487-9_40. arXiv: 1603.08511.

[42] Goodfellow I.J. et al. “Generative adversarial nets”. Advances inNeural Information Processing Systems. 2014.

[43] Unberath M. et al. “DeepDRR – A Catalyst for Machine Learning in Fluoroscopy-Guided Procedures”. Lecture Notes in Computer Science 2018. doi: 10.1007/978-3-030-00937-3_12. arXiv: 1803.08606.

[44] Unberath M. et al. “Enabling machine learning in X-ray-based procedures via realistic simulation of image formation”. International Journal of Computer Assisted Radiology and Surgery (2019).doi: 10.1007/s11548-019-02011-2.

[45] Russakoff D.B. et al. “Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration”. IEEE Transactions on Medical Imaging (2005). doi: 10.1109/TMI.2005.856749.

[46] Ruijters D. et al. “GPU accelerated Digitally Reconstructed Radiographs”. Proceedings of the 6th IASTED International Conference on Biomedical Engineering, BioMED 2008. 2008.

[47] Dorgham O.M. et al. “GPU accelerated generation of digitally reconstructed radiographs for 2-D/3-D image registration”. IEEE Transactions on Biomedical Engineering (2012). doi: 10.1109/TBME.2012.2207898.

[48] Vidal P.F. et al. “Development and validation of real-time simulation of X-ray imaging with respiratory motion”. Computerized Medical Imaging and Graphics (2016). doi: 10.1016/j.compmedimag.2015.12.002.

[49] Mu Z. et al. “A Fast DRR Generation Scheme for 3D-2D Image Registration Based on the Block Projection Method”. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. (2016). doi: 10.1109/CVPRW.2016.82.

[50] Moore C.S. et al. “A method to produce and validate a digitally reconstructed radiograph-based computer simulation for optimisation of chest radiographs acquired with a computed radiography imaging system”. British Journal of Radiology (2011). doi: 10.1259/bjr/30125639.

[51] Staub D. et al. “A digitally reconstructed radiograph algorithm calculated from first principles”. Medical Physics (2013). doi: 10.1118/1.4769413.

[52] Zhu J. et al. “Unpaired Image-to-Image translation Using Cycle-Consistent Adversarial Networks”. IEEE International Conference on computer Vision (ICCV). (2017). doi: 10.1109/ICCV.2017.244

Highlights

- DRRs represent valuable patient-specific pre-treatment information and training data.
- A novel framework for highly realistic and efficient DRR rendering is proposed, termed RealDRR.
- RealDRR allows on-site training to tailor the parameters to local imaging systems and protocols.
- RealDRR is able to render 100 highly realistic DRRs in under 2 minutes.

Conflict of interests statement

All authors declare to have no conflict of interests related to the work in this study.