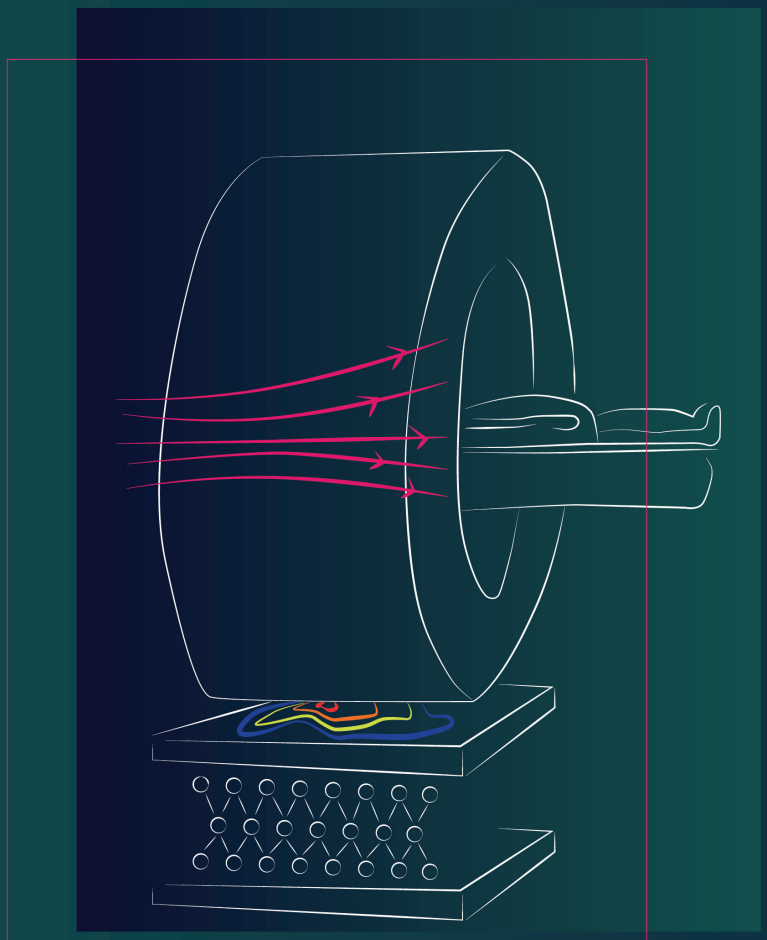
Patient-specific in-vivo QA in MRGRT: 3D EPID dosimetry for the Unity MR-linac

Torres Xirau, I.



6.

A DEEP LEARNING-BASED CORRECTION TO EPID DOSIMETRY FOR ATTENUATION AND SCATTER IN THE UNITY MR-LINAC SYSTEM

Igor Olaciregui-Ruiz [a,*]
Iban Torres-Xirau [a,*]
Jonas Teuwen [a,b]
Uulke A. van der Heide [a]
Anton Mans [a]

Department of Radiation Oncology,
The Netherlands Cancer Institute–Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

[b] Department of Radiology and Nuclear Medicine,
Radboud University Medical Center, Geert Grooteplein Zuid 10, 6525 GA Nijmegen, The Netherlands

*Both authors contributed equally.

Physica Medica, Volume 71, Pages 124-131
Published March 2020

Abstract

EPID dosimetry in the Unity MR-Linac system allows for reconstruction of absolute dose distributions within the patient geometry. Dose reconstruction is accurate for the parts of the beam arriving at the EPID through the MRI central unattenuated region, free of gradient coils, resulting in a maximum field size of ~10x22 cm^2 at isocenter. The purpose of this study is to develop a Deep Learning-based method to improve the accuracy of 2D EPID reconstructed dose distributions outside this central region, accounting for the effects of the extra attenuation and scatter.

A U-Net was trained to correct EPID dose images calculated at the isocenter inside a cylindrical phantom using the corresponding TPS dose images as ground truth for training. The model was evaluated using a 5-fold cross validation procedure. The clinical validity of the U-Net corrected dose images (the so-called DEEPID dose images) was assessed with *in vivo* verification data of 45 large rectum IMRT fields. The sensitivity of DEEPID to leaf bank position errors (±1.5 mm) and ±5% MU delivery errors was also tested.

Compared to the TPS, *in vivo* 2D DEEPID dose images showed an average γ-pass rate of 90.2% (72.6%-99.4%) outside the central unattenuated region. Without DEEPID correction, this number was 44.5% (4.0%-78.4%). DEEPID correctly detected the introduced delivery errors.

DEEPID allows for accurate dose reconstruction using the entire EPID image, thus enabling dosimetric verification for field sizes up to ~19x22 cm^2 at isocenter. The method can be used to detect clinically relevant errors.

Keywords: Unity MR-Linac, Deep Learning, *in vivo* EPID dosimetry

6.1. Introduction

The Unity MR-Linac system offers real-time soft-tissue visualization to allow for more precise delivery and online plan adaptation [164–166]. Independent quality assurance (QA) tools are required for the verification of the online adaptive strategies in magnetic resonance image-guided radiotherapy [167]. MR-compatible detector devices are currently employed to perform patient plan specific QA [97,98,100,168]. However, these solutions are typically time-consuming and not directly suitable for online adaptive workflows. Alternative solutions have been proposed, such as fast sanity checks for each adapted plan [169] or the use of online independent dose calculations [159]. The limitation of such checks is that they verify only parts of the workflow.

For conventional linacs, Electronic Portal Image Devices (EPIDs) are commonly used as an independent end-to-end dosimetric check of the Radiotherapy (RT) chain [148–150,170–175]. The process can be fully automated, which is essential to reduce the number of labor-intensive and error prone tasks [87]. The Unity MR-Linac is also equipped with an EPID which is mounted on the rotating gantry, opposite to the accelerator head [156]. For each plan adaptation, EPID images are acquired automatically containing information about the dose absorbed by the patient. Therefore, the implementation of an automated EPID-based dosimetric verification solution for the MR-Linac adapted workflow would be within reach once EPID dosimetry becomes feasible. To that purpose, our conventional dose back-projection algorithm [58,176] has been adapted to the Unity MR-Linac [177]. The algorithm utilizes pixel values of EPID images acquired during delivery to estimate the dose delivered to the patient. Patient plan specific QA is performed by comparing EPID-reconstructed dose distributions with those calculated by the Treatment Planning System (TPS) for each adapted plan. The main limitation of the method is that dose reconstruction is only accurate for

the parts of the beam that arrive at the EPID through the MRI central unattenuated region free of gradient coils, resulting in a maximum field size of ~10x22 cm^2 at isocenter. EPID-based dosimetric verification for larger fields is currently not possible, excluding, for instance, Intensity Modulated radiotherapy (IMRT) rectum plans.

Recent advances in deep learning are highly impacting various fields in science including healthcare and medical imaging [178]. In particular, convolutional neural networks (CNNs) have been largely applied in image detection and recognition [179], image segmentation [180], image registration [181] or image reconstruction [182]. The U-Net architecture, which comprises decoder, encoder and skip-connection modules integrated in a single network, has become the de-facto standard for image segmentation [183]. Recent studies have also modified the original design of the U-Net for dose prediction [184].

In this study, we present a Deep Learning-based method to improve the accuracy of 2D EPID reconstructed dose distributions outside the central unattenuated region, accounting for the effects of the extra attenuation and scatter. A U-Net was trained using 2D EPID and TPS dose images calculated at the isocenter inside a cylindrical phantom as training data. The goal was to use U-Net corrected dose images (the so-called DEEPID dose images) to accurately reconstruct *in vivo* 2D patient dose distributions using the entire EPID image, thus allowing for dosimetric verification of field sizes up to ~19x22 cm^2 at isocenter. The clinical validity of DEEPID was assessed with *in vivo* verification data of 45 rectum IMRT fields. To ensure that the presented Deep Learning-based method is capable of detecting clinically relevant errors, the sensitivity of DEEPID to leaf bank position errors and monitor unit (MU) errors was also tested.

6.2 Methods

6.2.1. Equipment

The Unity MR-linac system combines a 7 MV flattening filter free (FFF) beam linac (Elekta AB, Stockholm, Sweden) with an integrated wide bore 1.5 T MRI scanner (Philips Medical Systems, Best, the Netherlands) [126]. An a-Si flat panel X-ray detector (XRD 1642 AP, Perkin Elmer Optoelectronics, Wiesbaden, Germany) is mounted on the ring gantry built around the MRI scanner at a fixed source-to-detector distance (SDD) of 265.3 cm, and with a detection area of 41x41 cm^2 [156], see **Fig. 6.1**. EPID images were down sampled from their original size of 1024x1024 pixels to 256x256 pixels yielding a pixel pitch of 1.6 mm. The central region of the magnet is free of gradient coils, allowing for minimal and homogenous attenuation of the beam for field sizes up to ±11 cm in the cranial-caudal (CC) direction at isocenter (±20.2 cm at EPID level). However, since the effective size of the beam exiting the MRI scanner is larger due to divergence, the EPID acquisition of un-attenuated beams is limited to ±4.8 cm in the CC direction at the isocenter (±8.8 cm at EPID level). The EPID image is therefore divided into a central region receiving un-attenuated signal and an outer region receiving signal with extra attenuation and scatter due to exceeding the free-coils region. Since the detector is displaced 5.7 cm in the cranial direction with respect to the beam axis, fields exceeding 8.1 cm in the caudal direction at isocenter plane cannot be entirely acquired by the EPID and parts of the beam fall outside the panel. After the EPID images were cropped to remove the region where beams cannot be received (> 20.5 cm in cranial direction), the size of the input EPID images for dose reconstruction was 35.3 x 41 cm^2 (224x256 pixels).

Plans were generated using the Monaco 5.4 (Elekta AB, Stockholm, Sweden) treatment planning system (TPS). EPID images were measured

using Elekta's MVIC acquisition software.

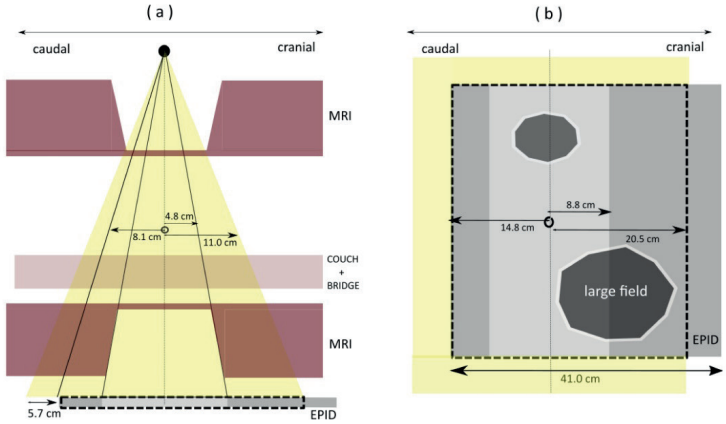


Fig. 6.1. (a) Unity MR-Linac sagittal cross section showing in yellow the maximum field size of 22 cm in the cranial-caudal (CC) direction at isocenter (black circle). The scale of the diagram has been adapted for viewing purposes. (b) Beams at the level of the EPID are received in a central unattenuated region (light grey) and in an outer attenuated region (dark grey). In the context of this study, a field is considered 'large' if the corresponding acquired EPID image contains signal in the outer region. The dashed black rectangle represents the cropped EPID image used for dose reconstruction.

6.2.2. Limitations of EPID dosimetry for large fields

For EPID dose reconstruction in this study, an adapted version of IViewDose software (Elekta, AB, Stockholm, Sweden) was used, which incorporates the adaptation of the conventional back-projection algorithm to the MR-Linac geometry [177]. The algorithm has two modes of operation: non-transit and transit EPID dosimetry [115,185]. In non-transit mode, *in air* EPID images acquired without a phantom/patient in the beam are utilized to reconstruct dose in any arbitrary geometry [186]. This mode is commonly used for pre-treatment verification of the reference plan. In transit mode, EPID images acquired behind the

patient are utilized to reconstruct dose within the patient. This mode is commonly used for *in vivo* verification of adapted plans.

To illustrate the limitations of transit EPID dosimetry for large fields, square fields (5x5, 10x10, 15x15 and 20x20 cm^2, 100MU) were irradiated at gantry angle 0° to a 20 cm thick slab phantom consisting of 30x30x1 cm^3 polystyrene slabs. Note that the measurements behind the phantom were used as a *surrogate* for *in vivo* measurements made behind the *patient.* TPS and EPID reconstructed dose profiles through the isocenter of the phantom were calculated and compared in both left-right (LR) and CC directions.

6.2.3. Deep learning architecture

Fig. 6.2 displays the U-Net architecture used in this study. A U-Net of depth 4 was used, where each down-sampling block consists of two blocks containing a convolution layer, followed by batch normalization [187] and a rectified linear unit (ReLu) activation function [188]. This block was subsequently followed by a max pooling layer of size 2x2. Convolutional filters in same mode with kernel 5x5 were selected. The number of filters started at 16 and was doubled for each subsequent block. The encoder was connected to the decoder by a last layer involving two 5x5 convolution layers. At the decoder parts, the up-convolutions were concatenated with the feature maps from the same layer of the encoding path [189]. The concatenated features followed the same sequence of convolution layers as in the decoder parts. This succession was repeated also four times. The final layer used a 1×1 convolution with a sigmoid as activation function. Each max pooling and concatenated layer was followed by a dropout layer with a drop rate of 0.3 [190]. The network was optimized using Adam [191], an algorithm for first-order gradient-based optimization of stochastic objective functions using default values of 0.9 and 0.999 for β_1 and β_2, respectively and a learning rate of 0.001. Mean squared error was utilized as loss

function for the training. The algorithm was trained for a maximum of 150 epochs but training could stop earlier if no improvement was found after 50 epochs. The U-Net was implemented in Python 3.6 with TensorFlow v1.7, an open source deep learning software library [192]. Training and evaluation of the network was performed on a GTX 1080 NVIDIA GPU. The hyperparameters of the network were tuned manually using the results of one of the cross-validation folds. As this showed the relative insensitivity to the selected hyperparameters, they were fixed for the rest of the study.

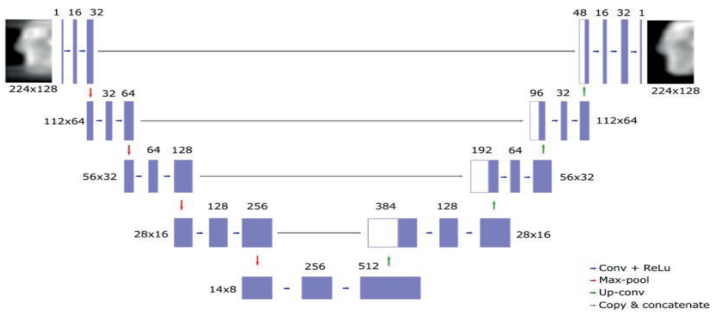


Fig. 6.2. Encoder and decoder pathways of our U-Net architecture.

6.2.4. Model training and validation

In-air EPID measurements were acquired for 90 IMRT fields corresponding to 12 adapted rectum plans of 6 patients. The *in air* EPID raw images were utilized in combination with the CT data set of the Octavius^{MR} 4D phantom (PTW, Freiburg, Germany) to reconstruct 2D EPID dose distributions at the isocenter plane in the phantom geometry using non-transit EPID dosimetry. More details regarding the employed methodology can be found elsewhere [185]. These EPID reconstructed dose images were utilized as input data for the training.

The corresponding dose distributions recalculated by the TPS on the phantom geometry were used as ground truth, or ideal output of the U-Net. The training of the model is performed with phantom data in order to have a well-controlled environment. The use of *in vivo* dose distributions for training would be problematic as the actual delivered dose to the patient is unknown (i.e. no 'ground truth' for training). Non-transit EPID reconstructed dose distributions were chosen for training to eliminate the need for phantom positioning. The choice of the Octavius^{MR} 4D phantom was made because TPS dose calculations were already available for that geometry.

The iViewDose software stored the EPID reconstructed 2D dose images in AVS field file format (Advance Visual Systems Inc, Waltham, MA, USA). Planned dose distributions were exported from the TPS in DICOM format and imported into iViewDose to calculate 2D dose distributions at the isocenter plane which were also stored in AVS format. The AVS files were then converted into numpy arrays for network training using in-house developed python code.

Regarding data augmentation, the EPID dose images were flipped around the cranial-caudal axis first and then split into nine crops of 128x224 pixels along the same axis. A total of 1620 (90x2x18) image patches were available as training pairs (EPID, TPS) to the network. The model was evaluated using a 5-fold cross validation procedure. The evaluation was scored by comparing U-Net corrected DEEPID dose images with TPS dose images by γ-analysis using 3% of the maximum dose and 3 mm as dose-difference and distance-to-agreement criteria, respectively. The results were calculated within the region surrounded by the 10% isodose line. γ-statistics were obtained for the outer attenuated region separately.

6.2.5. in vivo DEEPID dosimetry

in vivo EPID images were acquired behind the patient for 45 IMRT fields corresponding to 5 adapted rectum plans of 3 patients. Approximately half of these IMRT fields were large fields. The EPID raw images were utilized in combination with the CT data set of the patient to reconstruct *in vivo* 2D EPID dose distributions at the isocenter plane in the patient geometry using transit EPID dosimetry. The resulting 2D EPID dose images were passed to the U-Net to calculate the corrected DEEPID dose images. *in vivo* EPID and DEEPID 2D dose distributions at the isocenter were then compared with TPS dose distributions by γ-analysis. γ-statistics were obtained for the central and for the outer regions separately.

6.2.6. Introduction of errors

Machine delivery errors were introduced by manual modification of the treatment prescription file before import into the record-and-verify system (MOSAIQ version 2.65, Elekta Inc., Sunnyvale, CA, USA). An adapted rectum plan was delivered correctly first and then with leaf bank position errors introduced: leaves moved 1.5 mm inwards (closing the fields) and leaves moved 1.5 mm outwards (opening the fields). Similarly, another adapted rectum plan corresponding to a different treatment was also delivered correctly first and then with an increase and a decrease in the number of MUs of 5%. *In air* EPID raw measurements were acquired for all cases and were utilized in combination with the CT data set of the patient to reconstruct 2D EPID dose distributions at the isocenter plane in the patient using non-transit EPID dosimetry. The resulting EPID 2D dose images were passed to the U-Net to calculate the corrected DEEPID dose images.

6.3. Results

6.3.1. Limitations of EPID dosimetry for large fields

Fig. 6.3 displays TPS and EPID-reconstructed dose profiles through the isocenter in a phantom for four square fields. In the LR direction, the agreement was good for all field sizes. In the CC direction, however, for the 10x10 cm2 and 15x15 cm2 fields there was an overestimation of the EPID reconstructed dose outside the field. For the 15x15 cm^2 field, there was an underestimation of the EPID reconstructed dose inside the field. The agreement was again good for the 20x20 cm^2 field, see discussion section for more details.

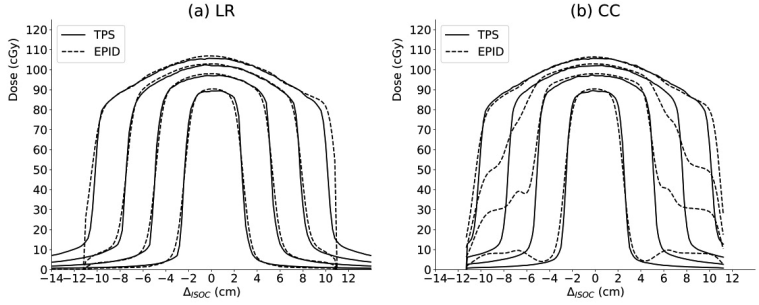


Fig. 6.3. LR and CC profiles of EPID-reconstructed and TPS dose distributions through the isocenter in a 20cm thick slab phantom consisting of 30x30x1 cm^3 polystyrene slabs for four representative square fields (5x5, 10x10, 15x15 and 20x20 cm^2, 100MU).

6.3.2. Model validation

Fig. 6.4.a illustrates the results of the 5-fold cross validation by presenting boxplots of γ-pass rate values for each fold calculated for the outer attenuated region only. The average median and interquartile range (IQR) values were 91.9±1.3%(1SD) and 8.2±1.7%(1SD), respectively. For comparison, **Fig. 6.4.b** presents the corresponding results obtained with EPID dosimetry before the DEEPID correction. The average

γ-pass rate values corresponding to the results of all folds combined were 91.6% (80.7%-100%) and 42.2% (4.9%-98.3%), for DEEPID and EPID respectively. Over the entire field, the average γ-pass rate values were 92.2% (80.8%, 99.6%) and 76.4% (51.7%,98.9%), for DEEPID and EPID respectively.

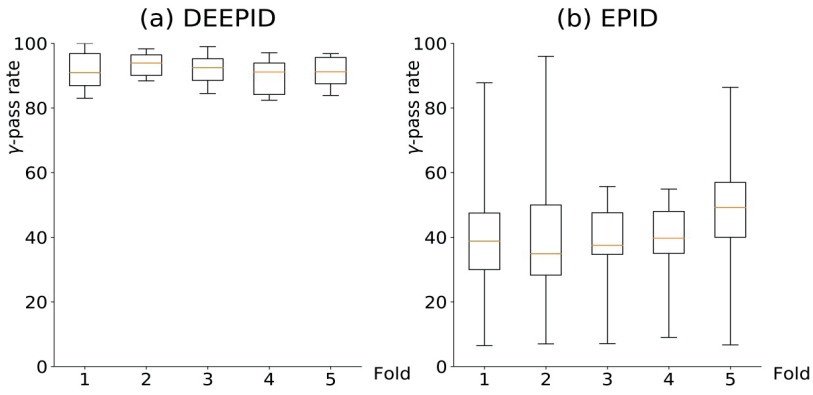


Fig. 6.4. 5-fold validation results presented as boxplots of γ-pass rate values corresponding to the comparison for the outer attenuated region between TPS dose distributions with (a) U-Net corrected DEEPID dose images and (b) EPID dose images before the correction. The box extends from the lower to upper quartile values of the data, with a line at the median. The whiskers are set at the 5th and 95th percentiles of the data. Outliers are not shown.

6.3.3. In vivo DEEPID dosimetry

Table 6.1 presents the results of the comparison between DEEPID and EPID with TPS dose distributions corresponding to the *in vivo* verification of 45 rectum IMRT fields. The agreement between DEEPID and TPS was similar for the central attenuated and the outer unattenuated regions. This can be also observed in the example presented in **Fig. 6.5** where *in vivo* EPID and DEEPID cranio-caudal dose profiles of a large rectum IMRT field are shown. Without the DEEPID correction, similar to the results of Fig. 3, there was an

underestimation of the EPID-reconstructed dose for the parts that were within the field. A good agreement with the TPS was only observed in the central unattenuated region. With DEEPID, the agreement with the TPS is also good in the outer attenuated region. TPS, EPID and DEEPID 2D dose distributions and γ-distributions for five rectum IMRT fields are displayed in **Fig. 6.6**.

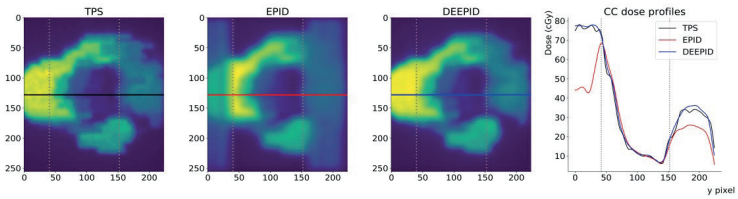


Fig. 6.5. TPS, EPID and DEEPID 2D dose distributions at the isocenter corresponding to the *in vivo* verification of a large rectum IMRT field. Cranial-caudal dose profiles are displayed in black, red and blue, respectively. The dotted grey lines represent the boundaries between central and outer regions.

	DEEPID		EPID	
	γ-pass rate %	γ-mean	γ-pass rate %	γ-mean
Central region	91.6 ± 6.6	0.47 ± 0.11	88.8 ± 7.5	0.56 ± 0.12
Outer region	90.2 ± 8.9	0.52 ± 0.14	44.5 ± 22.4	2.22 ± 2.70
Entire field	91.0 ± 6.7	0.49 ± 0.11	79.2 ± 13.3	0.84 ± 0.35

Table 6.1. Results of the comparison between DEEPID and EPID with TPS 2D dose distributions corresponding to the *in vivo* verification of 45 rectum IMRT fields. Results are presented as AVG±(1SD).

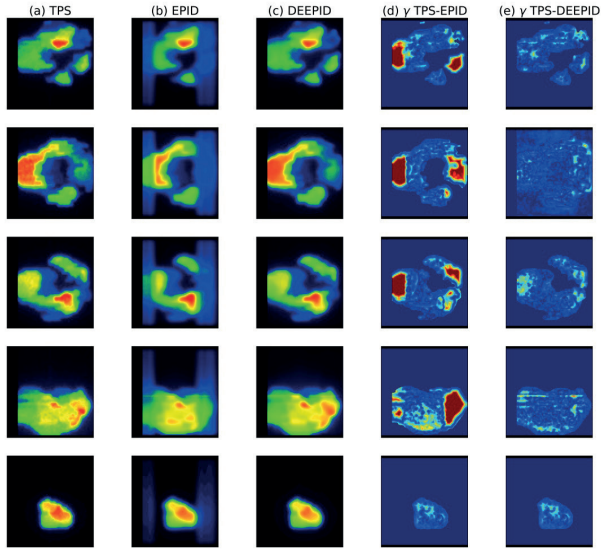


Fig. 6.6. TPS, EPID and DEEPID 2D dose distributions and γ-maps for five representative rectum IMRT fields. Only the last one is not a large field.

6.3.4. Introduction of errors

Results after ±1.5 mm movements of the leaf bank are displayed in **Fig. 6.7** for both EPID and DEEPID dose reconstructions. As in the example of Fig. 5, the DEEPID correction accurately reconstructs the dose in the outer region. The sensitivity to the error is also higher with DEEPID, as can be deduced from the presented γ-maps. Regarding the effect of -5% and +5% MU errors on DEEPID dose images, the isocenter dose difference for all fields combined with respect to the no-error situation was found to be -4.4% and +5.3%, respectively. The sensitivity of DEEPID to this error is further illustrated in **Fig. 6.8.** where dose profiles corresponding to one rectum field are displayed. Note that, in this case, the DEEPID correction improved also the agreement with the TPS in the central region for the no-error case, see Figs. **6.8.c** and **6.8.d**.

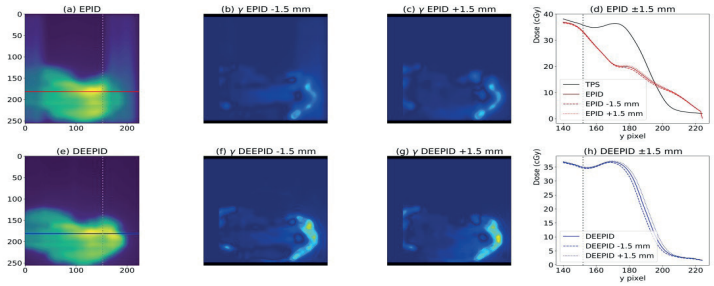


Fig. 6.7. Effect of ±1.5 mm leaf bank errors on EPID and DEEPID dose distributions. (a) EPID dose distributions for the no-error delivery of a large rectum field, (b,c) γ-distributions (2% / 1 mm) between no-error and error EPID dose distributions and (d) dose profiles. The same is displayed in (e,f,g,h) for DEEPID. The TPS dose profile for the no-error delivery is displayed for comparison.

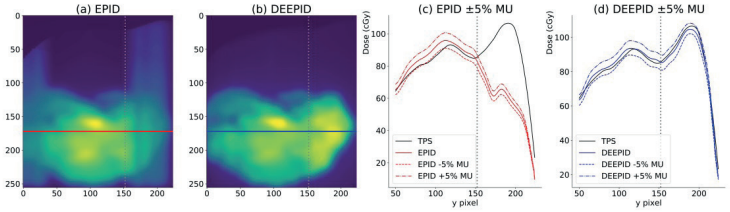


Fig. 6.8. Effect of ±5% MU errors on EPID and DEEPID dose distributions. (a,b) EPID and DEEPID 2D dose distributions for the no-error delivery and (c,d) EPID and DEEPID dose profiles. The TPS dose profile for the no-error delivery is displayed for comparison.

6.4. Discussion

In this study, we have used deep learning to correct the limitations of the EPID dose back-projection algorithm in the outer attenuated region of EPID images of the Unity MR-linac, making *in vivo* dosimetric validation feasible using the entire EPID image. The deficiencies of EPID dosimetry for the Unity MR-Linac were presented with square field measurements in **Fig. 6.3**. They illustrate how the EPID dose modelling of the panel fails to account for the differences in scatter

and attenuation in a range of field sizes for beams that traverse the non-coil free regions. During the last step of the pixel to portal dose response fit in the commissioning process [177], the parameters of a kernel are fitted to minimize the difference between the reconstructed dose at the EPID level and array measurements profiles in both LR and CC directions for a set of square fields irradiated at gantry angle 0°. The kernel is a rotationally symmetric gaussian kernel with no directional bias, producing incorrect results in the outer attenuated region. The use of asymmetric kernels and/or dedicated correction masks in the dose modelling of the panel was unsuccessfully explored prior to the investigation of the Deep Learning-based solution presented in this study.

The results of the 5-fold cross validation demonstrate how DEEPID improves the accuracy of dose reconstructions in the outer attenuated region. This was also corroborated with the *in vivo* verification results presented in **Table 6.1** and with the example of **Fig. 6.5.** The results presented in section 6.3.4. prove that DEEPID is sensitive to errors introduced in the outer attenuated region as it was the case with the ±1.5 mm leaf bank errors introduced for a large rectum IMRT field. The dosimetric effects of ±5% MUs errors were also correctly considered by DEEPID. This demonstrates the capability of the method to detect clinically relevant errors.

The EPID dose images utilized for training were reconstructed using non-transit EPID dosimetry from raw portal data acquired without a phantom/patient in the beam. However, the intended use of the model is for *in vivo* verification using transit EPID dosimetry from raw portal data acquired behind the patient. Although the dose engines for non-transit and transit EPID dosimetry are similar, they are not identical. Therefore, a more accurate model is expected if the EPID dose images utilized for training were reconstructed using transit EPID dosimetry from raw portal data acquired behind a phantom.

Secondly, TPS dose calculations on the Octavius^{MR} phantom were used as the ideal output of the U-Net during training. These calculations are believed to be accurate enough for the purpose of the study. Absolute dose measurements made with an MR-compatible Octavius 1500 2D detector array (PTW, Freiburg, Germany) were alternatively considered as ground truth for training. However, this option was disregarded due to the low resolution of the detector array (1405 vented ionization chambers with 7.1 mm center-to-center distance). Finally, the attenuation of the cryostat, couch and bridge (at the exit of the phantom) varies considerably with gantry angle [177]. Results should improve if a dedicated U-Net model was trained for each gantry angle separately. Finally, beams exceeding 8.1 cm in the caudal direction at isocenter plane, or 11 cm in the cranial direction, or ±11 cm in the lateral direction, will have parts that fall outside the panel. Evidently, no reconstruction method will be able to detect errors when these occur in these undetected parts of the beam.

The study demonstrates that Deep Learning can become a very powerful technique to correct deficiencies of portal dosimetry algorithms. A first estimate of the delivered dose distribution is generated using EPID dosimetry and then we pass the results to a convolutional neural network for a final correction.

6.5. Conclusions

A Deep Learning-based method corrects the deficiencies of EPID dosimetry to account for the extra attenuation and scatter in the outer attenuated region. The method allows for accurate dose back-projection at the isocenter plane using the entire EPID image. With this method, dosimetric verification becomes possible for field sizes up to ~19x22 cm^2 at isocenter. The method can be used to detect clinically relevant errors.

6.6. Acknowledgements

The authors wish to thank B. Vivas-Maiques, A. Torres Valderrama and S. v.d. Velden for their fruitful discussions.

6