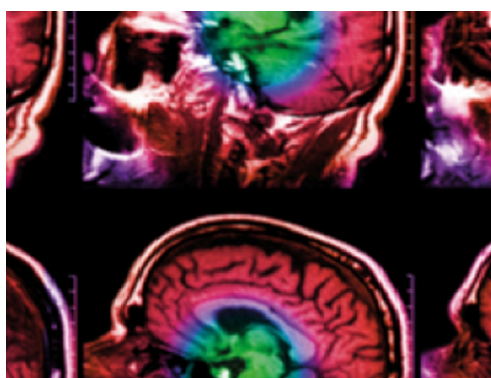# DeepDose: a robust deep learning-based dose engine for abdominal tumours in a 1.5 T MRI radiotherapy system

View the article online for updates and enhancements.

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

CrossMark

PAPER

DeepDose: a robust deep learning-based dose engine for abdominal tumours in a 1.5 T MRI radiotherapy system

G Tsekas , **G H Bol**, **B W Raaymakers** and **C Kontaxis**

Department of Radiotherapy, University Medical Center Utrecht, Heidelberglaan 100, Utrecht 3584CX, The Netherlands

E-mail: g.tsekas@umcutrecht.nl

Abstract

We present a robust deep learning-based framework for dose calculations of abdominal tumours in a 1.5 T MRI radiotherapy system. For a set of patient plans, a convolutional neural network is trained on the dose of individual multi-leaf-collimator segments following the DeepDose framework. It can then be used to predict the dose distribution per segment for a set of patient anatomies. The network was trained using data from three anatomical sites of the abdomen: prostate, rectal and oligometastatic tumours. A total of 216 patient fractions were used, previously treated in our clinic with fixed-beam IMRT using the Elekta MR-linac. For the purpose of training, 176 fractions were used with random gantry angles assigned to each segment, while 20 fractions were used for the validation of the network. The ground truth data were calculated with a Monte Carlo dose engine at 1% statistical uncertainty per segment. For a total of 20 independent abdominal test fractions with the clinical angles, the network was able to accurately predict the dose distributions, achieving 99.4% $\pm$ 0.6% for the whole plan prediction at the 3%/3 mm gamma test. The average dose difference and standard deviation per segment was 0.3% $\pm$ 0.7%. Additional dose prediction on one cervical and one pancreatic case yielded high dose agreement of 99.9% and 99.8% respectively for the 3%/3 mm criterion. Overall, we show that our deep learning-based dose engine calculates highly accurate dose distributions for a variety of abdominal tumour sites treated on the MR-linac, in terms of performance and generality.

1. Introduction

Magnetic resonance imaging (MRI) guidance directly from the treatment table for radiotherapy (RT) allows for real-time, high soft-tissue contrast visualization of malignant tumours, surrounding anatomies and organs at risk (OARs). Hybrid MRI-linear accelerators (MR-linacs) are being widely explored (Lagendijk *et al* 2014, Mutic and Dempsey 2014). They consist of a linear accelerator rotating under the presence of an external magnetic field and delivering the desired radiation to the targeted tumour cells while sparing the OARs. The introduction of the MR-linac in the clinic has significantly changed the treatment workflow of radiation therapy by offering reliable and personalized treatment plan adaptation, accounting for the latest anatomical changes of the patient on a daily basis (Winkel *et al* 2019), and by allowing the exploration of advanced delivery techniques (Kontaxis *et al* 2017).

A common characteristic of hybrid MRI RT systems is the presence of the electron return effect (ERE). This phenomenon is caused by the deflection of the released secondary electrons due to the Lorentz force from the static magnetic field. This deflection causes electrons to return to the skin surface after their exit from the body, thus resulting in a dose increase at tissue-air interfaces, like the skin or the lungs. The impact of the external magnetic field on the dose planning of MR-linacs has been extensively studied. Early research showed that the treatment planning and dose calculations are affected by the ERE (Raaymakers *et al* 2004, Raaijmakers *et al* 2007), however, using a multi-beam arrangement in combination with Intensity Modulated Radiation Therapy (IMRT) proved successful at compensating for the dose distortions caused (Raaijmakers *et al* 2007). Therefore,

an accurate dose modelling needs to take into account the impact of the magnetic field on the dose deposition to model effects such as ERE.

Conventionally, several dose calculation algorithms have been used in treatment planning systems including pencil beam (Mohan *et al* 1986) and Monte Carlo (MC) (Rogers 2006) based ones. Among contemporary state-of-the-art dose engines a widespread adoption of analytical MC simulations can be observed, which have been proven to yield the highest dose accuracy compared to other algorithms. Thanks to recent advances in computer hardware, GPU-based MC dose engine (GPUMCD) implementations (Hissoiny *et al* 2011) offer computationally boosted, highly-accurate dose calculations for dosimetry applications by simulating the full particle transport.

The evolution of deep learning (DL) methods has facilitated their use in dose calculations for radiation therapy treatment planning. Initial approaches included 'contour-to-dose' methods, where dose predictions from anatomical information (such as OAR contours) and computed tomography (CT) scans were explored (Kearney *et al* 2018, Nguyen *et al* 2019). While these first attempts assumed fixed beam angle setup, research has also been conducted towards developing more generic frameworks by incorporating beam arrangement information into the training (Barragán-Montero *et al* 2019).

Furthermore, recent DL approaches were focused on replicating dose engines by using inputs in the multi-leaf-collimator (MLC) domain, as the term 'dose engine' is typically used to describe the unit responsible for forward calculating the dose of a single MLC shape/element. The feasibility of a DL-based beamlet dose calculation method by predictive denoising for MRI-guided RT (MRIgRT) was proven in DeepMC (Neph *et al* 2020), where a denoised low-noise dose was predicted from an extremely noisy dose input and CT data while at the same time speeding up significantly the dose calculations. In addition, dose prediction using pre-calculated low-accuracy dose distributions from IMRT fluence maps has been examined (Xing *et al* 2020). The method we present in this work fits to this latter application category, as we developed a dose engine framework operating directly on MLC segments.

As presented above, a broad use of DL-based dose calculation methods can be found in literature. Nonetheless, one reason why artificial intelligence-based methods are often criticised, is the fact that they strongly rely on the data used for the training of the network and are usually restricted to solving a narrowly defined task of the treatment workflow. Thus, developing more generic DL-based solutions would require the presence of a framework robust enough to incorporate a broad amount of patient data while serving a highly significant clinical purpose. To that end, this paper examines the potential use of a DL-based dose engine capable of generating accurate clinical plans for a variety of cancer cases and patient anatomies.

The feasibility of a DL-based standalone dose calculation engine for prostate IMRT was proven in DeepDose (Kontaxis *et al* 2020), where dose distributions were generated per segment from a set of physics-based inputs and MLC segment shapes for each targeted anatomy. The patient cohort consisted of prostate cancer clinical plans from a conventional linear accelerator. The concept of DeepDose was able to generate highly accurate dose distributions, demonstrating a potential of being introduced to clinical setups in the future. Yet the framework was not previously tested on patients treated with MRIgRT under the presence of external magnetic field or for multiple tumour sites.

In this work, we aim to prove that a DL-based dose engine can be robust for anatomical variations, as well as robust for gantry positions and that it can include the impact of magnetic field dose effects. We propose an improved version of DeepDose by introducing cancer data from three anatomical sites: prostate, rectal and oligometastatic abdominal nodules. Also, the effects of the 1.5 T magnetic field on the delivered dose distribution are included to further expand our method. Additionally, we designed a generic framework by decoupling the dependency of beam angle configuration from the training process. This is accomplished by training the network with radiation beams emerging from randomly generated gantry angles. With this reliable DL-based dose engine we aim to investigate its potential future use in an online workflow, serving as the main dose engine for a broad range of clinical sites.

2. Materials and methods

2.1. Patient data

The patient cohort we used consisted of prostate, rectal and oligometastatic fractions, previously treated in our clinic with fixed-beam IMRT on the Elekta MR-linac with 6 MV FFF beam and MLC collimator with 22×57 cm^2 field size, 80 leaf pairs and 7.15 mm leaf width. The tumour sites included in the dataset were all located in the low abdominal area. Based on the treatment protocol, the prostate tumours were treated with 36.25 Gy split over 5 fractions, the rectal treatment consisted of 25 Gy delivered in 5 fractions and the oligometastatic tumours received 35 Gy in 7 fractions. The prescribed linac gantry angles were set to 0°, 50°, 100°, 155°, 205°, 260°, 310° for prostate, 80°, 145°, 180°, 215°, 280° for rectal and 45°, 80°, 120°, 145°, 180°,

215°, 235° for oligometastatic nodules respectively. In some oligometastatic tumour cases, an additional foam mattress was used for positioning purposes. For each of the patient fractions, the MLC shapes were extracted from the clinical plans, resulting in a total of 10279 segments.

Regarding the included data, we aimed at using only the first fraction of the treatment for each patient in our database, in order to increase the variability of the anatomies. Thus, unique patient plans were added to the dataset for all oligometastatic and prostate patients. However, the rectal plans available were fewer, therefore additional fractions of some rectal patients were used, yielding in total 72 unique prostate, 72 (43 unique) rectal and 72 unique oligometastatic patients in the dataset.

The data were uniformly distributed into training, validation and inference sets based on criteria of equivalent square area per segment, patient fractions per treatment site and beam angle configuration. After the final split, the distribution of data was 176 patient fractions for training, 20 for validation and 20 for inference, and thus 8368 training, 916 validation and 996 inference segments. Their average equivalent square area per segment was 33.8 ± 41.7 cm^2, 31.1 ± 37.5 cm^2 and 34.2 ± 42.6 cm^2 respectively.

Before training the model, random beam angles were assigned to each training sample to ensure its independence from the clinical beam angle configuration. This was achieved by substituting the prescribed angle of each segment with a randomly generated beam angle (synthetic segments), while maintaining the distribution of the assigned angles as uniform as possible. However, the various network inputs for both synthetic and clinical angles were calculated in order to compare the prediction of the total doses in the end.

The experiments were performed at 3 mm^3 grid spacing, which is the clinical resolution used for treatment planning and dosimetric evaluation. All patient grids included the treatment couch and were cropped to $216 \times 160 \times 80$ while making sure that the whole dose distribution of each patient is included in the grid. The ground-truth doses for each of the segments were calculated using the research version of the GPUMCD dose engine (Hissoiny *et al* 2011) using a statistical uncertainty of 1% per segment. Each segment was calculated with 100 Monitor Units (MU) in order to ensure a standardized dose output.

2.2. Network
2.2.1. Input
After splitting the data, the required input grids were calculated for each segment. The inputs of the network are the same as presented in DeepDose. The network expects five different 3D volumes, each one modelling a physical feature also used implicitly by conventional dose engines. These inputs are the mask of the segment, the distance from the source, the central beamline distance, the radiological depth and the volume density. These inputs, apart from adding information from a physics perspective, also allow patch-based training and localization instead of using full 3D anatomies. In order to gain a deeper insight of the intuitive physics-based inputs used, the reader is advised to inspect their analysis in the initial DeepDose implementation (Kontaxis *et al* 2020).

2.2.2. Model architecture
The network used for the dose prediction is based on the initial implementation of DeepDose, which follows the 3D UNet, originally published by Çiçek *et al* (2016) (figure 1). For the experiments NiftyNet (Gibson *et al* 2018), a medical image analysis platform was used and its original 3D UNet implementation was modified accordingly to match our objectives.

For the training of the network input grids of $3 \times 3 \times 3$ mm^3 grid spacing were used. The model was then trained in a supervised way using patch-based training, with patches of size $32 \times 32 \times 32$. A batch size of 32 patches was used and the root mean squared error was selected as a loss function. The Adam optimizer was chosen with a learning rate of and beta1, beta2 and epsilon parameters set to 0.9, 0.999 and 10^{-8} respectively. During training a validation step was performed every 5 iterations on the validation data in order to evaluate the performance of the network.

2.3. Analysis
The training of the network was performed on a workstation with a dual Intel® Xeon® E5-2690 v4, 128 GB RAM. The GPUMCD dose calculations used a single NVIDIA® Quadro GP100 card.

For the evaluation of the performance of our method the dose differences of the predicted dose distributions of each individual test segment were calculated for the voxels within the 10%–100% of the dose maximum. Moreover, gamma analysis of the predicted individual segments at 1%/1 mm, 2%/2 mm and 3%/3 mm was performed. The aforementioned evaluation was performed on both the clinical and the synthetic segments and the results were compared.

After summing up all clinical segments of each patient, weighted using their original MU values in the clinical plan, and constructing the total predicted dose, the dose differences were reported for each of the whole

Figure 1. 3D UNet architecture.
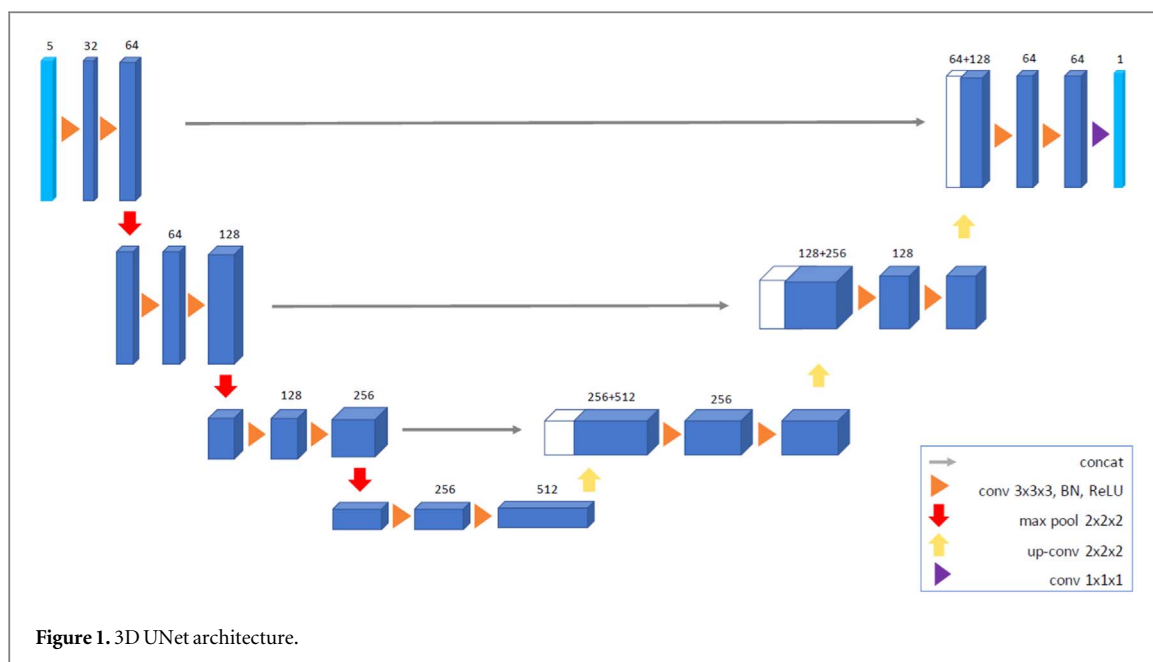
treatment plans for the voxels lying within the 50%–100% of the dose maximum and the same gamma analysis stats were calculated. PTV and OAR coverage was also evaluated by plotting some dose-volume-histogram (DVH) parameters.

Additionally, the model was used to predict the dose distributions of two extra anatomical sites. One cervical and one pancreatic tumour case were used to demonstrate the robustness of our method. For these additional cases, previously treated at our clinic with the MR-linac, the dose differences and 3%/3 mm gamma pass rate of the predicted dose grids were calculated. The performance of the model on these supplementary anatomical sites is presented by DVH curves. Compared to the clinical plans of our dataset, these additional cases included different OARs and bony anatomy, for example spinal cord, and different beam arrangement.

The accurate modelling of the ERE was assessed by comparing the ground-truth and predicted segment central profiles.

3. Results

3.1. Training results

The model was trained for approximately 14 days until convergence. It was trained for a total of 5.3×10^5 iterations with 32 patches per batch, resulting in over 10^8 processed random patches before being stopped early in order to avoid a potential overfitting (figure 2). The trained network was then used to generate the dose per segment for the fractions of the test set.
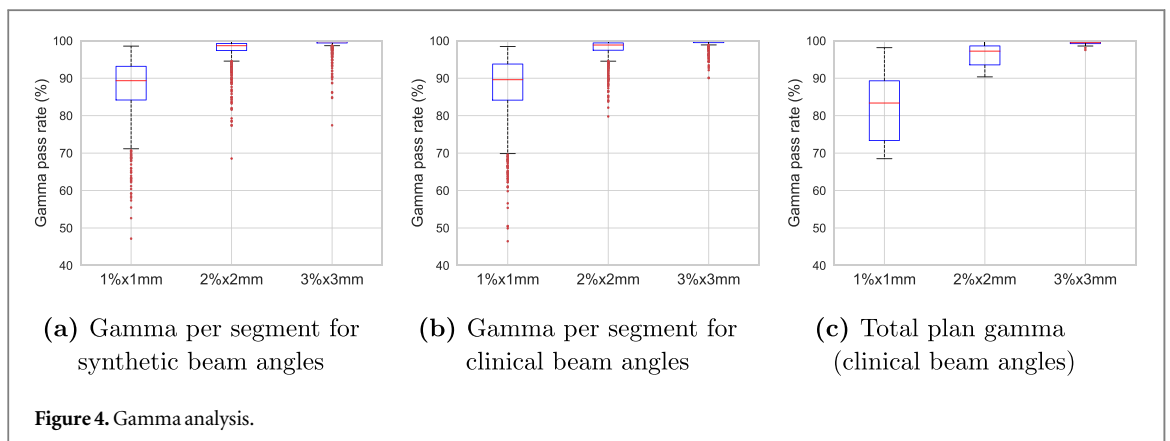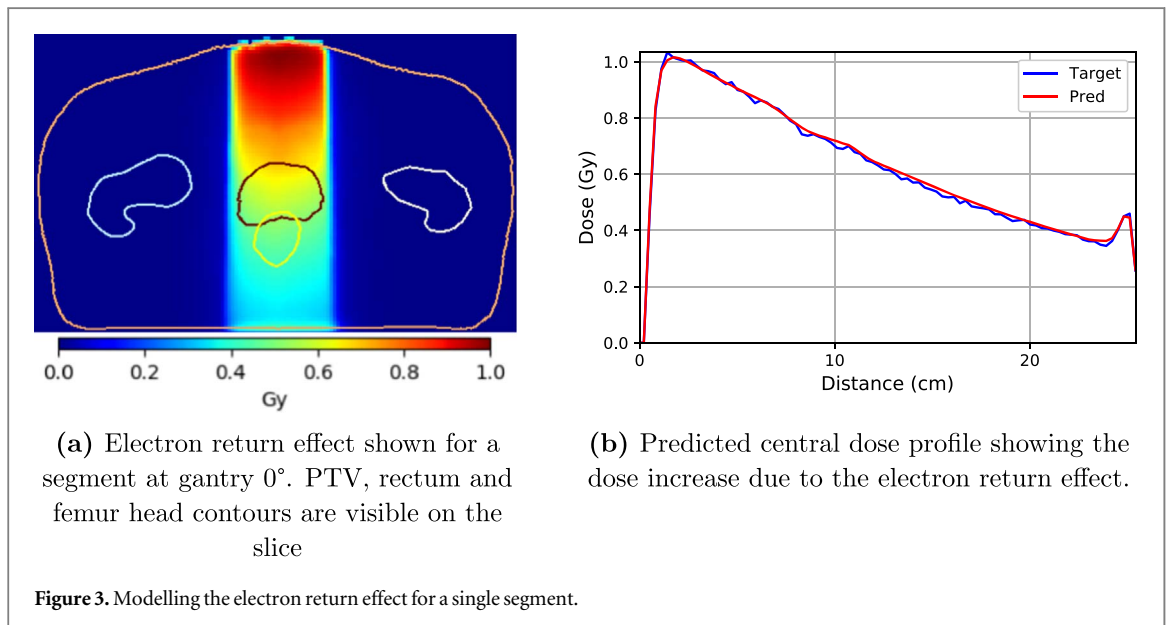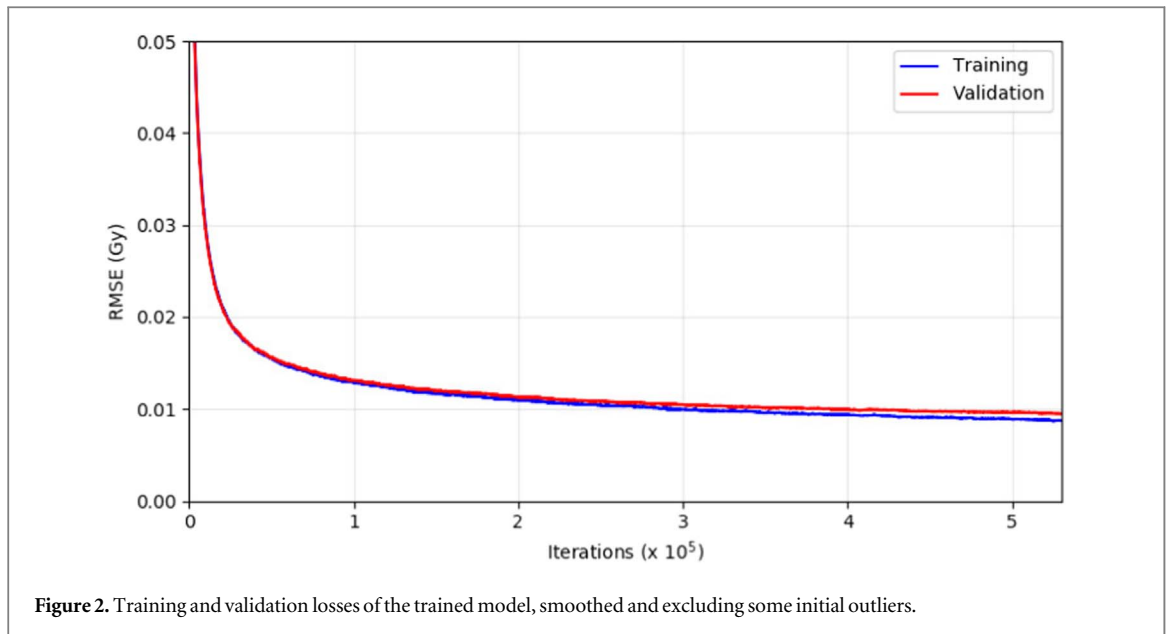
3.2. ERE modelling

The presence of the ERE was taken into account and successfully modelled from the network. Figure 3(a) illustrates one indicative view of a predicted prostate segment at 0°, where the ERE is most evident. On figure 3(b), the respective predicted central dose profile is compared to the target one. The ERE is evident at the posterior part of the patient body as depicted clearly with a dose increase spike, prior to reaching the body/couch mattress interface.

3.3. Dose calculation results

The overall agreement between the target and predicted dose distributions was very good. The network was capable of predicting the dose deposition in the patient anatomies from the generated input quantities regardless of the beam angle arrangement.

Among all individual synthetic segments, an average dose difference and standard deviation of 0.3% ± 0.7% (0.002 ± 0.006 Gy) was reported for the voxels within 10%–100% of the dose maximum, while the per segment average dose difference for the clinical segments was 0.3% ± 0.7% (0.002 ± 0.007 Gy), demonstrating an equally accurate behaviour.

For each of the synthetic and clinical segments a gamma analysis was performed and the gamma pass rates for different criteria were calculated (figure 4). Synthetically generated segments averaged gamma pass rates of

Figure 2. Training and validation losses of the trained model, smoothed and excluding some initial outliers.

(a) Electron return effect shown for a segment at gantry 0°. PTV, rectum and femur head contours are visible on the slice

(b) Predicted central dose profile showing the dose increase due to the electron return effect.

Figure 3. Modelling the electron return effect for a single segment.

(a) Gamma per segment for synthetic beam angles

(b) Gamma per segment for clinical beam angles

(c) Total plan gamma (clinical beam angles)

Figure 4. Gamma analysis.

87.7% ± 7.7%, 97.7 ± 2.9% and 99.3 ± 1.5% for the 1%/1 mm, 2%/2 mm and 3%/3 mm respectively. For each of the segments following the clinical protocol, the average pass rates were 87.6% ± 8.3%, 97.9 ± 2.6% and 99.5 ± 1.0% for the 1%/1 mm, 2%/2 mm and 3%/3 mm, respectively, demonstrating highly accurate agreement with the reference dose values.

For the clinical total dose distributions of the 20 inference patients, very good agreement was reported, with average dose difference of 0.6% ± 0.6% (0.2 ± 0.2 Gy). Average gamma pass rates of 82.2% ± 9.7%, 96.1% ± 3.1% and 99.4% ± 0.6% for the 1%/1 mm, 2%/2 mm and 3%/3 mm criteria respectively were achieved. Figure 5 depicts the DVH and the central transversal slice of a predicted total dose distribution for a prostate case of the test set.

3.4. Additional testing

For the additional cervical and pancreatic cases, a very good agreement was observed. After predicting the dose distribution for the cervical patients, the gamma pass rate reported was 91.9 % for the 1%/1 mm, 99.3 % for the 2%/2 mm and 99.9% for the 3%/3 mm criteria. The dose difference calculated was at 0.2 Gy (0.6%), while the overall organ coverage of the DVH was assessed as highly acceptable (figure 6).

The pancreatic case showed similar performance, the dose difference lied at 0.5 Gy (1.3%), also with a convincing OAR coverage. The according gamma pass rates were 83.3%, 98.0% and 99.8% for the 1%/1 mm, 2%/2 mm and 3%/3 mm analysis.

4. Discussion

In this paper we proposed a robust DL-based solution for accurate dose calculations of abdominal tumours in IMRT. We extended DeepDose by introducing segments of varying shapes and sizes, using random beam angles, from three different anatomical sites treated clinically on a 1.5 T MR-linac The network was trained on a set of intuitive physics-based inputs per segment and was then used to infer whole dose distributions. Our method was successful at modelling the particle interactions and dose deposition under the external magnetic field, including the ERE. For the clinically used 3 mm³ grid spacing we demonstrated convincing dose predictions for a set of previously unseen abdominal plans.
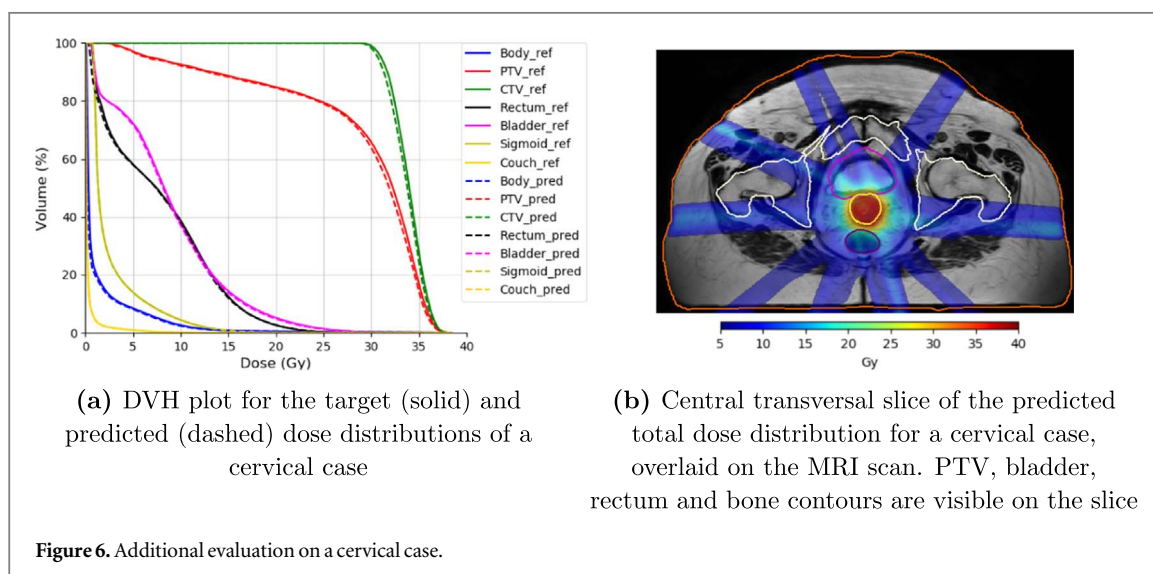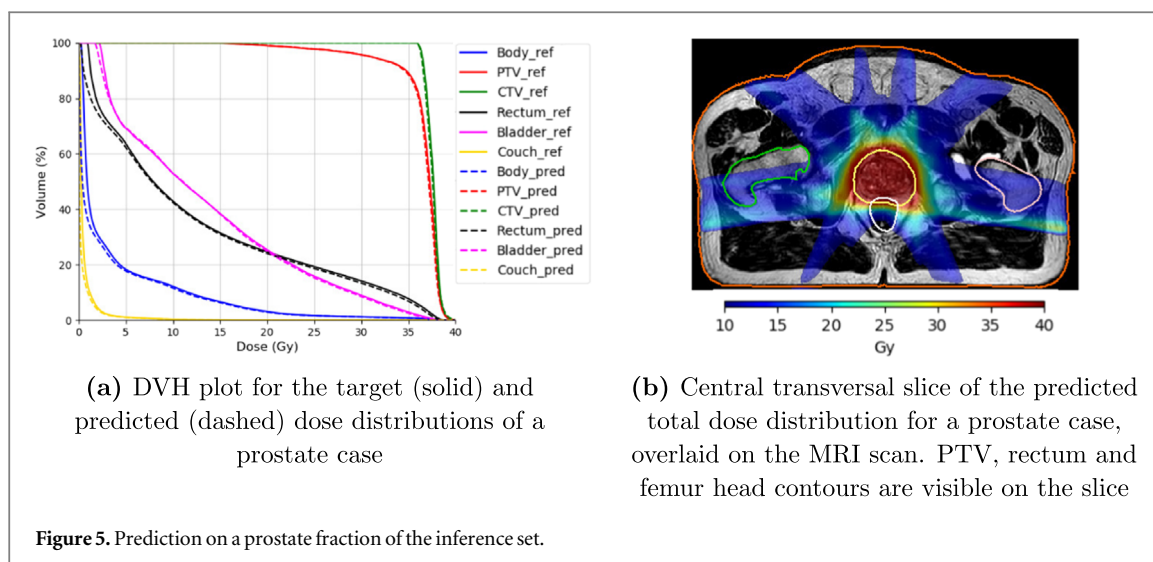
A major enhancement of this extended DeepDose framework is the generality of segment shapes and sizes included in the dataset. In particular, each abdominal tumour site offered segments with highly heterogeneous size, with an average equivalent square area of 20.2 ± 12.0 cm², 69.5 ± 52.9 cm² and 8.1 ± 7.3 cm² for the prostate, rectal and oligometastatic cases respectively. Additionally, compared to the 40 leaf pairs present in a conventional Elekta linac, MR-linac has 80 leaf pairs, thus offering higher spatial modulation and yielding more complex segment shapes. The proposed model was able to accurately reproduce the delivered dose of the various MLC patterns, proving the robustness of our approach.

In addition, we tested the trained model on one cervical and one pancreatic tumour case in order to demonstrate the generality of our method and reproduce the delivered dose of these MLC shapes. The generated dose predictions passed all clinical QA gamma tests with an excellent agreement score. The diverse beam arrangement and 3D dose distributions of these additional cases show the potential of extending the treatment sites currently handled by our framework.

To our knowledge, this is the first DL-based dose engine trained on MR-linac data from various anatomical sites that can additionally be robust to beam angle variations. While other dose calculation approaches, such as DeepMC (Neph *et al* 2020), operate on individual MLC beamlet data, our proposed network architecture differs as it uses MLC segments, namely shapes containing multiple beamlets. To that end, the network we use relies on a novel set of physics-based quantities that encode all MLC segment shape information for the dose deposition inside different tissue types. The use of random angles during the training phase significantly enhances the generality of our method and its robustness to different anatomies.

As further shown in the initial implementation of DeepDose (Kontaxis *et al* 2020), different input combinations can be explored in order to reduce the number of input grids needed and to investigate a potential speedup of the training procedure. The average total time needed for the dose calculation of one segment with DeepDose is approximately 3 sec (around 2 sec for input generation and 0.8 sec for the inference). Also, for the whole plan calculations DeepDose is currently faster than GPUMCD for an uncertainty value of 1% per segment. Future research will focus on optimizing both the model and the hardware components, thus giving higher computational boost to our framework.

The accurate modelling of the particle interactions in a 1.5 T transverse magnetic field alongside with the prior application of DeepDose at 0 T using prostate plans (Kontaxis *et al* 2020) proves the potential applicability of our approach on different RT systems, such as a 0.35 T environment (Mutic and Dempsey 2014). However, in order to use data from linear accelerators with a different MRI field, an additional training of the network will be required to

(a) DVH plot for the target (solid) and predicted (dashed) dose distributions of a prostate case

(b) Central transversal slice of the predicted total dose distribution for a prostate case, overlaid on the MRI scan. PTV, rectum and femur head contours are visible on the slice

Figure 5. Prediction on a prostate fraction of the inference set.

(a) DVH plot for the target (solid) and predicted (dashed) dose distributions of a cervical case

(b) Central transversal slice of the predicted total dose distribution for a cervical case, overlaid on the MRI scan. PTV, bladder, rectum and bone contours are visible on the slice

Figure 6. Additional evaluation on a cervical case.

ensure optimal results. In addition, future research could focus on applying DeepDose on dynamic Volumetric Modulated Arc Therapy (VMAT) treatment plans. This type of therapy typically features a wide variety of MLC shapes and each treatment plan consists of multiple control points per arc. The appropriate modifications to our framework in order to incorporate more complex VMAT treatment plans are considered to be future work.

Until now the limitations of our method have not yet been thoroughly explored. Apart from its highly accurate performance on tumour sites of the abdomen, it is expected that introducing additional tumour anatomies which include a significant amount of previously unseen tissue type, for instance air cavities in lung tumours, will demand additional training of the network. Nevertheless, this work shows that our network can be trained to correlate arbitrary gantry angles and segment shapes with the corresponding dose distribution and as such that indeed adding more sites to the training data will increase the applicability of the DL-based dose engine.

For this purpose, our future work will focus on establishing a working envelope that will set boundaries on the plan parameters that DeepDose can accurately handle. Thus, a precise range of segment types will be assessed as acceptable, based on their shape, equivalent square area and other structural features. In that way, the creation of a safe environment for the introduction of DeepDose in a clinical setting will be facilitated and we will be able to move forward towards an application targeting whole body dose calculations.

In the near future we aim to demonstrate further progress in clinical research by introducing DeepDose into the MR-linac treatment workflow. An initial goal would be to take advantage of its reduced inference time and include it in the treatment pipeline of abdominal tumours as a fast secondary dose check. The primary goal of a secondary dose engine is to assess the quality of the generated plans prior to treatment. DeepDose is expected to surpass the accuracy of the current secondary dose check software for the clinical 3%/3 mm criterion and

therefore we are confident that it will enhance the current plan evaluation procedure. Furthermore, no external software is needed in order to run a whole plan inference, thus indicating that our approach could serve as a standalone dose application.

On the long term, the ability of our method to accurately predict a variety of patient plans in the range of a few seconds makes it a promising candidate for a clinical dose engine used in the online plan optimization itself. Therefore, after establishing its performance on various anatomies, for instance lung and brain tumours, we aim to gradually move towards exploring DeepDose as the primary dose calculation engine for a variety of tumours in MRI-guided radiation therapy.

5. Conclusion

We presented a robust DL-based framework for dose calculations of abdominal tumours in IMRT. The network was trained per segment with MR-linac data from three different abdominal tumour sites using random beam angle configuration. The trained network was then able to generate 3D dose predictions for whole patient plans. This approach will increase the efficiency of dose checks in the online workflow by initially serving as a secondary dose engine for MRIgRT of the abdominal area in our clinic. Subsequently, we aim to introduce DeepDose as the primary DL-based dose engine for online plan optimization.

Acknowledgments

ORCID iDs

G Tsekas ⓘ https://orcid.org/0000-0001-5834-738X
C Kontaxis ⓘ https://orcid.org/0000-0002-7300-6513

References

Barragán-Montero A M, Nguyen D, Lu W, Lin M H, Norouzi-Kandalan R, Geets X, Sterpin E and Jiang S 2019 Three-dimensional dose prediction for lung imrt patients with deep neural networks: robust learning from heterogeneous beam configurations *Med. Phys.* **46** 3679–91

Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T and Ronneberger O 2016 3D U-Net: learning dense volumetric segmentation from sparse annotation *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* ed S Ourselin *et al* (Berlin: Springer) pp 424–32

Gibson E *et al* 2018 Niftynet: a deep-learning platform for medical imaging *Comput. Methods Programs Biomed.* **158** 113–22

Hissoiny S, Raaijmakers A J, Ozell B, Després P and Raaymakers B W 2011 Fast dose calculation in magnetic fields with gpumcd *Phys. Med. Biol.* **56** 5119–29

Kearney V, Chan J W, Haaf S, Descovich M and Solberg T D 2018 Dosenet: a volumetric dose prediction algorithm using 3d fully-convolutional neural networks *Phys. Med. Biol.* **63** 235022

Kontaxis C, Bol G H, Lagendijk J J and Raaymakers B W 2020 Deepdose: towards a fast dose calculation engine for radiation therapy using deep learning *Phys. Med. Biol.* **65** 075013

Kontaxis C, Bol G H, Stemkens B, Glitzner M, Prins F M, Kerkmeijer L G, Lagendijk J J and Raaymakers B W 2017 Towards fast online intrafraction replanning for free-breathing stereotactic body radiation therapy with the mr-linac *Phys. Med. Biol.* **62** 7233–48

Lagendijk J J, Raaymakers B W, Berg C A V D, Moerland M A, Philippens M E and Vulpen M V 2014 Mr guidance in radiotherapy *Phys. Med. Biol.* **59** R349–69

Mohan R, Chui C and Lidofsky L 1986 Differential pencil beam dose computation model for photons *Med. Phys.* **13** 64–73

Mutic S and Dempsey J F 2014 The viewray system: magnetic resonance-guided and controlled radiotherapy *Semin. Radiat. Oncol.* **24** 196–9

Neph R, Lyu Q, Huang Y, Yang Y M and Sheng K 2020 Deepmc: a deep learning method for efficient monte carlo beamlet dose calculation by predictive denoising in magnetic resonance-guided radiotherapy *Phys. Med. Biol.* **66** 035022

Nguyen D, Jia X, Sher D, Lin M H, Iqbal Z, Liu H and Jiang S 2019 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected u-net deep learning architecture *Phys. Med. Biol.* **64** 065020

Raaijmakers A J E, Raaymakers B W, van der Meer S and Lagendijk J J W 2007 Integrating a mri scanner with a 6 mv radiotherapy accelerator: impact of the surface orientation on the entrance and exit dose due to the transverse magnetic field *Phys. Med. Biol.* **52** 929–39

Raaijmakers A J, Hardemark B, Raaymakers B W, Raaijmakers C P and Lagendijk J J 2007 Dose optimization for the mri-accelerator: Imrt in the presence of a magnetic field *Phys. Med. Biol.* **52** 7045–54

Raaymakers B W, Raaijmakers A J E, Kotte A N T J, Jette D and Lagendijk J J W 2004 Integrating a mri scanner with a 6 mv radiotherapy accelerator: dose deposition in a transverse magnetic field *Phys. Med. Biol.* **49** 4109–18

Rogers D W 2006 Fifty years of monte carlo simulations for medical physics *Phys. Med. Biol.* **51** R287

Winkel D *et al* 2019 Adaptive radiotherapy: the elekta unity mr-linac concept *Clin. Transl. Radiat. Oncol.* **18** 54–9

Xing Y, Nguyen D, Lu W, Yang M and Jiang S 2020 Technical note: a feasibility study on deep learning-based radiotherapy dose calculation *Med. Phys.* **47** 753–8