Physics and Imaging in Radiation Oncology 32 (2024) 100641

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/physics-and-imaging-in-radiation-oncology

Original Research Article

Automated plan generation for prostate radiotherapy patients using deep learning and scripted optimization

Cody Church [a,*], Michelle Yap [a], Mohamed Bessrour [a], Michael Lamey [a], Dal Granville [b]

[a] Department of Medical Physics, The Ottawa Hospital General Campus, Canada
[b] Department of Radiation Oncology and Department of Physics and Atmospheric Science, Dalhousie University, Canada

ARTICLE INFO

Keywords:
Deep learning
Autoplanning
Radiotherapy

ABSTRACT

Background and Purpose: Treatment planning is a time-intensive task that could be automated. We aimed to develop a "single-click" workflow, fully deployed within a commercial treatment planning system (TPS), for autoplanning prostate radiotherapy treatment plans using predictions from a deep learning model (DLM).
Materials and Methods: Automatically generated treatment plans were created with a single script, executed from within a commercial TPS scripting environment, that performed two stages sequentially. Initially, a 3D dose distribution was predicted with a ResUNet DLM. The DLM was trained and validated using previously treated datasets (n = 120) which used 3D contours as inputs. Following this, dose predictions were converted into treatment plans by extracting dose-volume metrics from the predictions to use as objectives for the inverse optimizer within the TPS. An independent test dataset (n = 20) was used to evaluate the similarity between automated and clinical plans.
Results: For planning target volumes, the median percentage difference and interquartile range between the automatically generated plans and clinical plans were 0.4% [0.2-1.1%] for the $V_{100\%}$, −0.5% [(−1.0)-(−0.2)%] for $D_{99\%}$ and −0.5% [(−1.0)-(−0.2)%] for $D_{95\%}$. Bladder and rectum volume-at-dose objectives agreed within −6.1% [(−12.5)-0.9%]. The conversion of the DLM prediction into a treatment plan took 15 min [13-16 min].
Conclusions: An automatic plan generation workflow that uses a DL model with scripted optimization was fully deployed in a commercial TPS. Autoplans were compared to previously treated clinical plans and were found to be non-inferior.

1. Introduction

Modern treatment planning in radiotherapy is commonly achieved with an inverse optimization process. Quality of treatment plans and the time taken to generate them can be affected by the skills and experience of the treatment planner [1]. As such, automating the inverse optimization process has the potential to improve efficiency and reduce variability [2,3].

Recent studies in automated treatment planning typically employ a knowledge-based planning (KBP) approach. As summarized in Babier et al., KBP workflows involve a two-step process [4]. In the first step, the dose distribution for a new patient treatment is predicted using knowledge of previous, similar patient treatments. In the second step, a deliverable treatment plan (hereinafter shortened to treatment plan) is automatically generated to deliver a dose distribution that closely matches the prediction. We define a deliverable plan in the context of volumetric modulated arc therapy (VMAT) as a plan that respects the physical limitations of the linac (*e.g.* dose rates, multileaf collimator (MLC) motions, jaw motions, gantry angles, collimator angles, couch angles, etc.).

The first step (*i.e.* dose prediction) is often accomplished using machine learning [5,6] or deep learning (DL). DL approaches may be preferable as they eliminate the need for feature engineering and can retain full three-dimensional (3D) dose information. Excellent performance of DL models for modulated dose distribution predictions has been demonstrated for a variety of treatment sites [7–14]. Notably, all entries in the 2020 AAPM OpenKBP Grand Challenge used DL for dose prediction [15].

The second step of the KBP process (*i.e.* generation of a treatment plan that closely matches the prediction) is often achieved using dose mimicking algorithms [15]. Dose mimicking algorithms are not common in commercial treatment planning systems (TPSs), although one is

* Corresponding author at: Department of Medical Physics, The Ottawa Hospital General Campus, 501 Smyth Rd, Ottawa, ON K1H 8L6, Canada.
E-mail address: cchurch@toh.ca (C. Church).

https://doi.org/10.1016/j.phro.2024.100641
Received 19 April 2024; Received in revised form 30 August 2024; Accepted 4 September 2024
Available online 8 September 2024

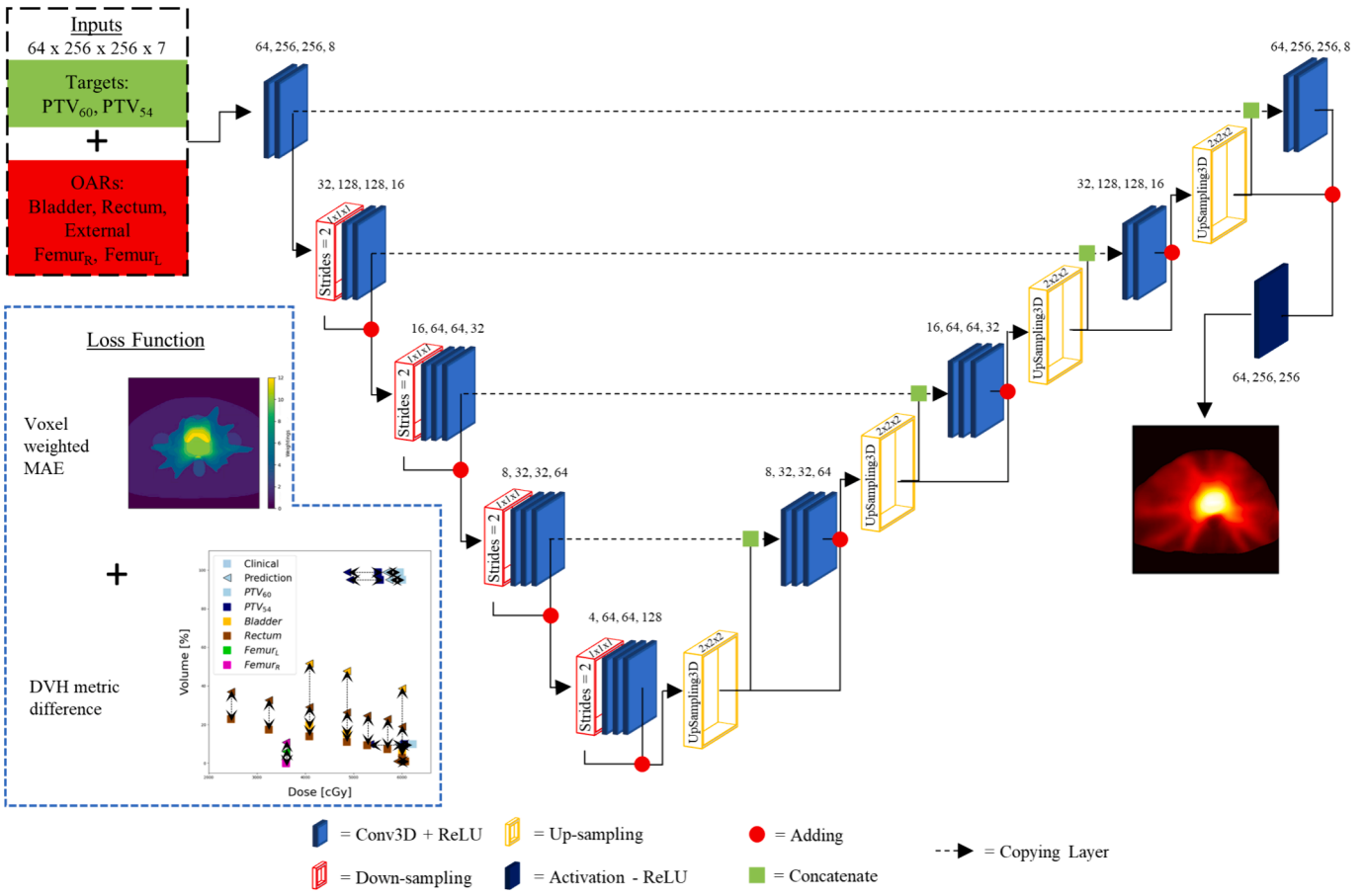

Fig. 1. A schematic representation of the residual U-Net architecture used in this work. The numbers above each block represent the spatial dimensions and number of feature channels along each convolutional layer. Dashed lines with green squares represent dense connections (concatenation) between the encoder and decoder regions. Skip connections (addition) are represented by red circles. The output of the network is a 3D dose distribution with the same spatial resolution as the input structures. The customized loss function is visualized in the bottom left which features a summation of the voxel-weighted mean absolute error (with the relative weightings shown) with a summation of all the difference between dose-volume metrics.

available for fallback planning [16]. A limitation of this approach is that the predicted dose distribution may not be optimal, although this can potentially be addressed using objective functions that attempt to improve upon the predicted dose distribution, rather than simply matching it [17].

An alternative to dose mimicking is to use predicted dose distributions to generate objectives and weights that can be used in standard inverse optimization algorithms. Although less common in recent KBP studies, this approach has advantages. Inverse optimization tools are readily available in commercial TPSs and are widely used in radiotherapy workflows. Additionally, prominent commercial TPSs contain scripting tools that can be used to automate the process. This method may offer a more human-centered approach to automation as it allows planners to evaluate plans in familiar software and continue optimization in an intuitive manner, if desired.

A recent investigation examined the direct prediction of VMAT MLC patterns using DL, foregoing the optimization step [18]. This approach, though novel and promising, has thus far resulted in plans of insufficient quality, and requires “in-house” software to generate plans, which makes clinical deployment challenging.

Although the use of DL dose predictions in KBP pipelines has been well-demonstrated [4,15], there are few studies that have used optimization algorithms within commercial TPSs to generate treatment plans. Using a commercial TPS for this step reduces barriers to clinical deployment as they are typically well-validated and approved by regulatory bodies. Xia et al. developed an automated method to generate rectum intensity-modulated radiation therapy (IMRT) plans that required exports from the TPS to an external DL server [19]. Lempart et al. derived optimization objectives using a nearest-neighbor search to identify the most similar dose distributions in an atlas of historical treatments [9]. This approach relied on availability of a robust atlas and included some manual steps. van de Sande et al. used a commercial dose mimicking algorithm designed for fallback planning to create breast IMRT plans based on two-dimensional (2D) DL dose predictions [20]. To our knowledge, this algorithm cannot be automated and requires manual intervention.

In this work, we developed a fully-automated workflow for generating VMAT treatment plans for prostate radiotherapy within a commercial TPS. To our knowledge, this is the first study demonstrating a fully-automated workflow that uses a custom DL dose prediction model and scripted optimization to automatically generate VMAT plans within a commercial TPS. Potential benefits of this approach include the use of common optimization objectives (which planners can intuitively interact with to tweak the plan), the elimination of data transfers, full deployment within a commercial TPS scripting environment, and improvements to dose distributions beyond mimicking predictions.

2. Materials and methods

2.1. Data description

Data from 140 previously treated prostate radiotherapy patients were used in this study. All patients were treated with dual-arc VMAT plans that delivered 60 Gy to the prostate and 54 Gy to the proximal

Table 1
Dose volume metrics that were included in the custom loss function. In this table, D_X refers to the dose received by X% structure volume and V_Y refers the volume receiving Y Gy.

Structure	Dose Volume Metrics
PTV_{60}	$D_{99\%}$, $D_{95\%}$, $D_{10\%}$
PTV_{54}	$D_{99\%}$, $D_{95\%}$, $D_{10\%}$
Bladder	V_{60Gy}, $V_{48.6Gy}$, $V_{40.8Gy}$, D_{max}
Rectum	V_{60Gy}, V_{57Gy}, $V_{52.8Gy}$, $V_{48.6Gy}$, $V_{40.8Gy}$, $V_{32.4Gy}$, $V_{24.6Gy}$, D_{max}
$Femur_{Right}$	V_{36Gy}
$Femur_{Left}$	V_{36Gy}

seminal vesicles in 20 fractions. Each dataset included the 3D dose distribution and contours of the planning target volume (PTV) PTV_{60} (prostate + margins), PTV_{54} (proximal 1 cm of seminal vesicles + margins), bladder, rectum, left/right femurs (contoured inferiorly to the extent of the ischial tuberosities) and the external body contour. Patients with hip prostheses were not included. Treatment plans in this dataset were created in the Monaco TPS (Elekta AB, Stockholm) and dose was calculated using Monaco's Monte Carlo (MC) dose engine. This study was approved by the Ottawa Health Science Network Research Ethics Board (ID#: 20200300 01*H*).

2.2. Deep learning dose prediction model

A subset of the dataset described in section 2.1 was used to train (n=100) and validate (n=20) a 3D residual U-Net to predict 3D dose distributions [21,22]. Structure contours and dose data were extracted from digital imaging and communications in medicine (DICOM) files exported from Monaco and used to train the DL model. Structures were converted to binary masks and dose distributions were normalized such that $D_{95\%}$ of the PTV_{60} was equal to 60 Gy, which is a common normalization method used in our clinic. Structure and dose arrays were resized and cropped to extend 72 mm inferiorly and 120 mm superiorly about the isocentre resulting in 64×256×256 arrays with 1.2×1.2×3 mm voxel size to match the computed tomography (CT) images. This cropping was chosen to balance the intent to minimize the dataset size (reducing computational memory burden) while maintaining the full extent of the key organs at risk (OARs) (*e.g.* rectum and bladder).

The first layer of the DL model used 8 filters and each subsequent layer used $f_i = 2^i * f_0$, where $f_0 = 8$. The model was trained with 1000 epochs (batch size = 1) and a learning rate of 0.0001 with the Adam optimizer. The model architecture is shown in Fig. 1. TensorFlow [23] and an NVIDIA RTX 8000 GPU were used for model training. The model was trained using a custom, domain-specific loss function inspired by Soomro et al. and Sun et al. [24,25]. This loss function summed a voxel-wise weighted mean absolute error (MAE) function with absolute dose-volume histogram (DVH) metric differences between predicted and delivered dose distributions. Weights in the MAE calculation varied by structure and dose level. Voxels within targets (*i.e.* PTV_{60} and PTV_{54}) and key OARs (*i.e.* bladder and rectum) were assigned higher weights, while voxels in low-dose areas and less important structures were assigned lower weights. The weight parameters were tuned based upon the expertise of two independent observers to preserve the dose distributions expected with local plans using the validation dataset. A visual depiction of the loss function is included in Fig. 1. DVH metrics used in the loss calculation were derived from internal institutional protocols based on the CHHiP trial [26,27] and represent important metrics in local plan-quality evaluation. These are summarized in Table 1.

2.3. Scripted optimization

The RayStation (RaySearch Laboratories, Stockholm) Python scripting application programming interface (API) was used to automatically optimize treatment plans (autoplans) using DVH metrics extracted from the predicted dose distributions. This scripted routine was developed by using simple DVH-based optimization objectives in RayStation to mimic (and, if possible, improve upon) DVH curves from clinical plans (created in Monaco) in the validation dataset (n=20). Once the routine was able to reproduce the clinical plans adequately, it was then used to produce plans that match predicted dose distributions in the autoplanning pipeline.

Table 2
Dose volume metrics that were calculated from the dose predicted by the deep learning model. These metrics were used as dose-volume objectives in the automated TPS optimization process. Organs at risk include the rectum, bladder, and individual femurs.

Structure	Dose Volume Metrics
PTV_{60}	V_{60Gy}, V_{59Gy}, $D_{1.0\%}$, D_{max}
PTV_{54}	V_{54Gy}, $V_{55.5Gy}$, V_{57Gy}, $D_{1.0\%}$, D_{max}
Organs At Risk	V_{60Gy}, V_{50Gy}, V_{40Gy}, V_{30Gy}, V_{20Gy}, V_{10Gy}, $D_{1.0\%}$, D_{max}

The DVH-based optimization objectives and weights used in this routine were chosen for their simplicity and were adjusted through a trial-and-error approach. Because there is not a simple, consensus metric to determine similarity or superiority of dose distributions, we relied on the expertise of two independent observers to inform adjustments to the objectives to improve the routine. This process was analogous to common treatment planning workflows, in which planners first start with ideal optimization objectives and adjust until they receive an adequate result based on their clinical expertise.

The DVH-based objectives used in this routine are summarized in Table 2. Because predicted dose distributions are not necessarily optimal, we attempted to improve upon them by setting DVH objectives that were slightly better than those predicted. For all OARs, the predicted $D_{1.0\%}$ and D_{max} values were reduced by 1.0% to create 'MaxDVH' and 'MaxDose' objectives. Similarly, the volumes predicted to receive a given dose had the doses reduced by 1.0 Gy to create 'MaxDVH' objectives (*e.g.* the predicted V_{60Gy} was used to create a V_{59Gy} 'MaxDVH' objective). For target volumes (*i.e.* PTV_{54} and PTV_{60}), the predicted DVH metrics that quantify coverage (*i.e.* V_{54gy}, V_{59Gy}, V_{60Gy}) were increased by 1.0% to a maximum of 100% of the volume to create 'MinDVH' objectives. The predicted DVH metrics that quantify hot spots (*i.e.* $V_{55.5Gy}$, V_{57Gy}, $D_{1.0\%}$, and D_{max}) were reduced by 1.0% to create 'MaxDVH' and 'MaxDose' objectives. In addition to the DVH objectives, a single 'DoseFallOff' objective was used to control conformality.

For the initial optimization, all objectives were given a weight of 1. After each iteration, objectives that were not met had their weight increased by 50. This process was repeated for 10 iterations. The choice of 10 iterations was found to balance improvements in the plan with total optimization time.

2.4. Fully automated planning deployment

The autoplanning process described in sections 2.2 and 2.3 was fully deployed within the RayStation Python scripting API. A Python virtual environment was created in RayStation, and the packages necessary for executing the DL dose prediction model were installed (*e.g.* TensorFlow and its dependencies). Execution of a single script within RayStation resulted in an autoplan being generated according to the following automated steps (also shown in Fig. 2). Initially, the patient's 3D contours were converted into binary arrays for input into the DL model described in section 2.2 to generate a predicted 3D dose distribution. DVH metrics were then extracted from the predicted dose distribution and used to create a patient-specific optimization template. Following this, the VMAT treatment plan characteristics were initialized (*e.g.* energy, arc geometry, couch, and collimator positions) and the iterative optimization process described in section 2.3 was executed to create a treatment plan.

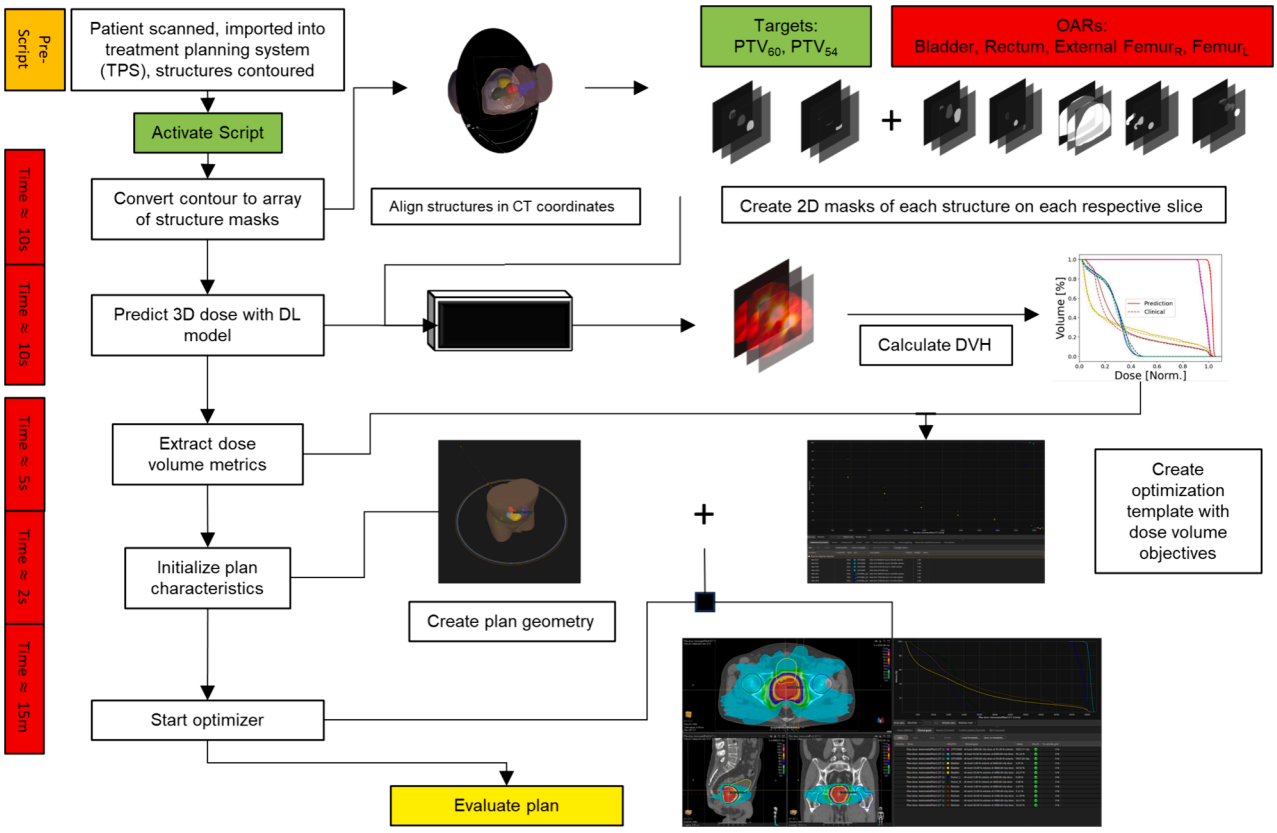

Fig. 2. The workflow for automatic treatment plan generation using the RayStation scripting API. The native Python language and built-in functions within this environment allow for dose prediction with a DL model. Dose-volume metrics from the predicted dose were used to create and execute a patient-specific optimization and generate a plan with the built-in TPS optimizer.

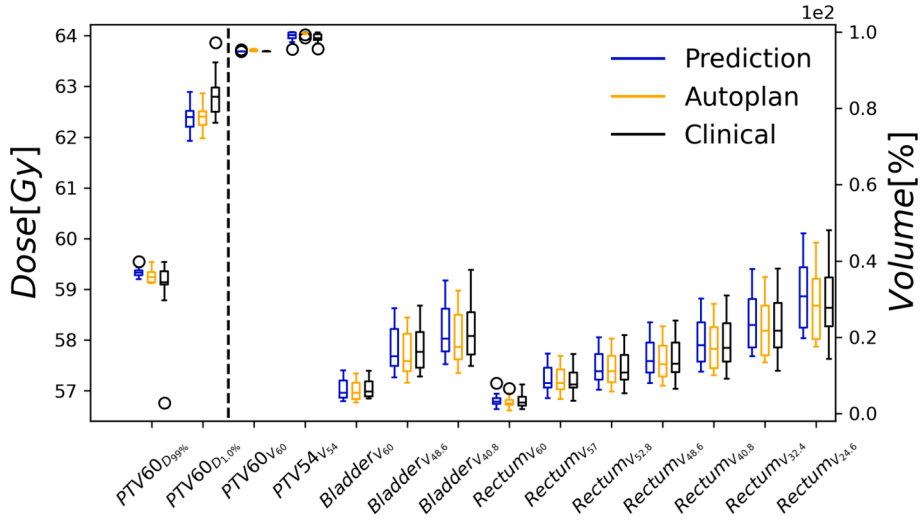

Fig. 3. Comparison of various dose-volume metrics between the deep learning dose prediction (Prediction), the deliverable plan generated with the automated workflow (Autoplan), and the previously delivered clinical plan (Clinical). The centre tick of the bars represents the median value across all patients (n = 20) and the box represents the upper and lower quartile range. The left-axis references all dosimetric metrics and the right axis references all volumetric metrics.

2.5. Statistical evaluation and analysis

An independent test dataset (n=20) was used to evaluate the full autoplanning process. A two-sided Wilcoxon signed-rank test with a significance level of 0.05 was performed to compare DVH metrics used to assess plan quality in our institution (outlined in Fig. 3) between the autoplan and the clinical plan. To adjust for multiple comparisons, a Bonferroni correction was used with n=18 comparisons to adjust the significance level to 0.0028. This correction is for the comparisons between the metrics shown in Fig. 3 (metrics not shown for femurs as they were equal to zero) and the conformity index [28] for the PTV_{60} and PTV_{54}. All metrics are reported as the median and interquartile range.

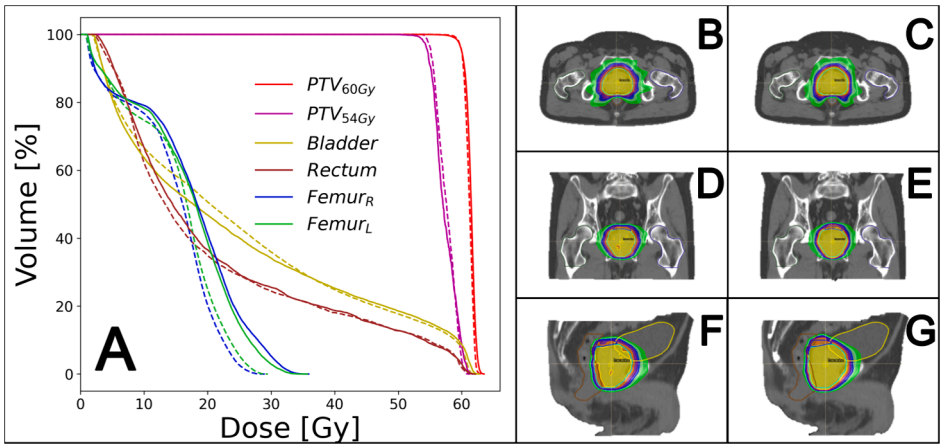

Fig. 4. A characteristic example comparing a clinical plan (previously treated) to automatically generated plan derived from a dose predicted by the deep learning model. (A) Shows the dose volume histograms for all relevant structures where the solid lines are from the clinical plan and the dashed lines are from the autoplan. Subfigures (B), (D) and (F) shown an axial, coronal and sagittal slice from the clinical dose distribution overlayed on the CT. Subfigures (C), (E) and (G) shown an axial, coronal and sagittal slice from the autoplan dose distribution overlayed on the CT.

3. Results

The conversion of the prediction into the autoplan took 15 min [13-16 min] across the test set. A comparison of the DVH-based plan-quality metrics used in our local institutional protocols is shown for the DL dose prediction, autoplan, and clinical plan in Fig. 3. With respect to these metrics, n=15 patients passed all PTV metrics in the clinical plans with n=5 cases exceeding the primary threshold for maximum dose (63 Gy). For the autoplans, all patients passed all metrics for the PTVs. There was no significant difference between the conformity index found for the PTV_{60} (clinical: 0.92 [0.91-0.94], autoplan: 0.93 [0.91-0.93]) or the PTV_{54} (clinical: 0.20 [0.17-0.26], autoplan: 0.22 [0.18-0.27]). With respect to metrics for OARs, 91.5% were met across all clinical plans and 91.9% were met across all autoplans. The plan-quality metrics that most frequently failed were V_{60} for the bladder and rectum. Failure rates for the bladder V_{60} were 55% in both the clinical and autoplans. Failure rates for the rectum V_{60} were 45% for the clinical plans and 40% for the autoplans. Between the autoplan and the clinical plan, statistically significant differences were seen for V_{60} and $D_{1.0\%}$ in PTV_{60}. Significant differences were seen for V_{54} in PTV_{54}. For OARs, significant differences were only seen in the bladder $V_{48.6}$ (p=4.8e-04) and $V_{40.8}$ (p=7.1e-04). The most notable differences between the clinical and autoplans were hotter doses received by PTV_{54} (increase of 6% [3-13%] in volume receiving 57 Gy) and smaller volumes receiving doses in the ~12–20 Gy range.

4. Discussion

In this work, a fully automated, “single-click” planning process deployed within a commercial TPS was presented. This process produced prostate VMAT plans that were non-inferior to previous clinically approved plans. There are several benefits to this automated process: A treatment plan is generated with a “single-click.” The model used for predicting dose is potentially sharable and does not contain any personal health information. Specialized hardware/software is not needed (beyond the TPS). The scripting solution is designed with freely available software (Python). The treatment plans are created within the confines of a validated commercial TPS. The final plan and optimization template is in a familiar format that planners can further improve and/or tweak. No data transfers are required.

In the first stage of autoplan generation, a 3D dose distribution was predicted with a residual U-Net trained with 3D contours and previous clinical dose distributions. A potential limitation of this work is the combination of model architecture and loss function selected. Rigorous testing of multiple architecture and loss function pairings was not explored and is beyond the scope of this work. Our goal was to generate a model with sufficient performance for our purposes, rather than rigorously comparing multiple modelling methods. There is no clear consensus on the “best” choice of architecture and loss function for dose prediction models, though many combinations have been studied [4,7,8,10,11,15,20,29–31]. While the speed of model training could be impacted by the model parameters and data configuration, this step is likely inconsequential from a clinical-utility perspective; once a model is trained, calling it to predict a 3D dose, even on a CPU, should only take seconds to a few minutes with modern hardware.

Strategies to automatically generate treatment plans from DL dose predictions within clinical TPSs have varied in the literature. Lempart et al. [9] derived optimization objectives using a nearest neighbor search that compared predictions to an atlas of previously treated cases. This work required the availability of a sizeable and robust atlas and relied on manual intervention for DICOM transfers and setting of some optimization objectives. In the RayStation TPS, van de Sande et al. [20] created deliverable doses for left-sided breast cancer patients using a commercial dose-mimicking algorithm on doses predicted by a 2D U-Net model. They found statistically significant differences in the mean dose received by the PTV – although the magnitude of the difference may not be clinically significant. Xia et al. [19] performed autoplanning using a similar strategy to this work, in that DL dose predictions were used to generate optimization objectives in a clinical TPS in a fully-automated manner. It differed, however, in that autoplans had a single target (rectum), less complex treatment modality (IMRT), used a different TPS (Pinnacle, Philips Radiation Oncology Systems), and required DICOM data transfers to a dedicated DL server.

One limitation of this study is the unavailability of clinical plans developed in RayStation to compare to the autoplans. RayStation was used to generate autoplans due to its mature Python scripting API, but available clinical plans were generated using Monaco. Dose distributions in Monaco were calculated using MC, while RayStation used collapsed cone superposition-convolution. Because of the statistical uncertainty in MC calculations, there are larger maximum doses [32,33] and less-steep DVHs. The magnitude of the impact is larger in smaller volumes [34]. A characteristic example comparing an autoplan and clinical plan is shown in Fig. 4. As seen in this example, the DVHs for PTV_{60} and PTV_{54} exhibit a more homogenous dose in the autoplan. This observation is generally consistent across all patients in the independent test set and is at least partially attributable to differences in dose engines and beam models, rather than genuine differences in plan quality. Similarly, we could not compare manually generated RayStation plans

to autoplans. However, we have no reason to expect manually generated RayStation plans to substantially differ from the Monaco plans, outside of the differing dose engines. Favourable comparisons with the clinically approved Monaco plans suggest that the autoplans are of sufficient quality to be deemed clinically acceptable.

In conclusion, a single-click automated plan generation routine fully deployed within the scripting environment of a commercial TPS was presented. Automated plans were non-inferior to manually generated plans when evaluated using common DVH-based metrics.

Financial disclosures

This work was supported by the Harold E. Johns Scholarship.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CRediT authorship contribution statement

Cody Church: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Michelle Yap:** Methodology, Software, Validation. **Mohamed Bessrour:** Methodology, Software. **Michael Lamey:** Conceptualization, Supervision, Funding acquisition, Writing – review & editing. **Dal Granville:** Conceptualization, Data curation, Supervision, Funding acquisition, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Cody Church – No declaration. Michael Lamey – No declaration. Michelle Yap - Funded by the Harold E. Johns Scholarship. Mohamed Bessrour - Funded by the Harold E. Johns Scholarship. Dal Granville - Currently Employed by Nova Scotia Health Authority in the Department of Radiation Oncology and appointed with the Department of Physics and Atmospheric Science, Dalhousie University.

References

[1] Nelms BE, Robinson G, Markham J, Velasco K, Boyd S, Narayan S, et al. Variation in external beam treatment plan quality: an inter-institutional study of planners and planning systems. Pract Radiat Oncol 2012;2:296–305. https://doi.org/10.1016/j.prro.2011.11.012.

[2] Scaggion A, Fusella M, Roggio A, Bacco S, Pivato N, Rossato MA, et al. Reducing inter-and intra-planner variability in radiotherapy plan output with a commercial knowledge-based planning solution. Phys Med 2018;53:86–93. https://doi.org/10.1016/j.ejmp.2018.08.016.

[3] van Gysen K, Kneebone A, Le A, Wu K, Haworth A, Bromley R, et al. Evaluating the utility of knowledge-based planning for clinical trials using the TROG 08.03 post prostatectomy radiation therapy planning data. Phys Imaging Radiat Oncol 2022;22:91–7. https://doi.org/10.1016/j.phro.2022.05.004.

[4] Babier A, Mahmood R, Zhang B, Alves VG, Barragán-Montero AM, Beaudry J, et al. OpenKBP-Opt: an international and reproducible evaluation of 76 knowledge-based planning pipelines. Phys Med Biol 2022;67:185012. https://doi.org/10.1088/1361-6560/ac8044.

[5] McIntosh C, Welch M, McNiven A, Jaffray DA, Purdie TG. Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method. Phys Med Biol 2017;62:5926. https://doi.org/10.1088/1361-6560/aa71f8.

[6] Shiraishi S, Moore KL. Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy. Med Phys 2016;43:378–87. https://doi.org/10.1118/1.4938583.

[7] Nguyen D, Jia X, Sher D, Lin M-H, Iqbal Z, Liu H, et al. 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture. Phys Med Biol 2019;64:065020. https://doi.org/10.1088/1361-6560/ab039b.

[8] Mahmood R, Babier A, McNiven A, Diamant A, Chan TC. Automated treatment planning in radiation therapy using generative adversarial networks. PMLR: Proceedings of the 3rd machine learning for healthcare conference; 2018; Palo Alto, California. California: 2018; p. 484–499.

[9] Lempart M, Benedek H, Gustafsson CJ, Nilsson M, Eliasson N, Bäck S, et al. Volumetric modulated arc therapy dose prediction and deliverable treatment plan generation for prostate cancer patients using a densely connected deep learning model. Phys Imaging Radiat Oncol 2021;19:112–9. https://doi.org/10.1016/j.phro.2021.07.008.

[10] Bakx N, Bluemink H, Hagelaar E, van der Sangen M, Theuws J, Hurkmans C. Development and evaluation of radiotherapy deep learning dose prediction models for breast cancer. Phys Imaging Radiat Oncol 2021;17:65–70. https://doi.org/10.1016/j.phro.2021.01.006.

[11] Ahn SH, Kim E, Kim C, Cheon W, Kim M, Lee SB, et al. Deep learning method for prediction of patient-specific dose distribution in breast cancer. Radiat Oncol 2021;16:1–13. https://doi.org/10.1186/s13014-021-01864-9.

[12] Koike Y, Takegawa H, Anetai Y, Ohira S, Nakamura S, Tanigawa N. Patient-specific three-dimensional dose distribution prediction via deep learning for prostate cancer therapy: improvement with the structure loss. Phys Med 2023;107:102544. https://doi.org/10.1016/j.ejmp.2023.102544.

[13] Murakami Y, Magome T, Matsumoto K, Sato T, Yoshioka Y, Oguchi M. Fully automated dose prediction using generative adversarial networks in prostate cancer patients. PloS one 2020;15:e0232697.

[14] Nguyen D, Long T, Jia X, Lu W, Gu X, Iqbal Z, et al. A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning. Sci Rep 2019;9:1076. https://doi.org/10.1038/s41598-018-37741-x.

[15] Babier A, Zhang B, Mahmood R, Moore KL, Purdie TG, McNiven AL, et al. OpenKBP: the open-access knowledge-based planning grand challenge and dataset. Med Phys 2021;48:5549–61.

[16] RayStation BD. External beam treatment planning system. Med Dosim 2018;43:168–76.

[17] Yousefi A, Ketabi S, Abedi I. A novel mathematical model to generate semi-automated optimal IMRT treatment plan based on predicted 3D dose distribution and prescribed dose. Med Phys 2023;50:3148–58. https://doi.org/10.1002/mp.16236.

[18] Heilemann G, Zimmermann L, Schotola R, Lechner W, Peer M, Widder J, et al. Generating deliverable DICOM RT treatment plans for prostate VMAT by predicting MLC motion sequences with an encoder-decoder network. Med Phys 2023;50:5088–94. https://doi.org/10.1002/mp.16545.

[19] Xia X, Wang J, Li Y, Peng J, Fan J, Zhang J, et al. An artificial intelligence-based full-process solution for radiotherapy: a proof of concept study on rectal cancer Front. Oncol 2021;10:616721. https://doi.org/10.3389/fonc.2020.616721.

[20] van de Sande D, Sharabiani M, Bluemink H, Kneepkens E, Bakx N, Hagelaar E, et al. Artificial intelligence based treatment planning of radiotherapy for locally advanced breast cancer. Phys Imaging Radiat Oncol 2021;20:111–6. https://doi.org/10.1016/j.phro.2021.11.007.

[21] Han X. Automatic liver lesion segmentation using a deep convolutional neural network method arXiv:170407239 2017. https://doi.org/10.48550/arXiv.1704.07239.

[22] Milletari F, Navab N, Ahmadi S-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: IEEE: Proc. 2016 Fourth International Conference on 3D Vision (3DV); 2016; Stanford, California. California: 2016; p. 565-571.

[23] Chollet F. Deep learning with Python. Simon and Schuster; 2021.

[24] Soomro MH, Alves VGL, Nourzadeh H, Siebers JV. DeepDoseNet: a deep learning model for 3D dose prediction in radiation therapy arXiv preprint arXiv:211100077 2021. https://doi.org/10.48550/arXiv.2111.00077.

[25] Sun Z, Xia X, Fan J, Zhao J, Zhang K, Wang J, et al. A hybrid optimization strategy for deliverable intensity-modulated radiotherapy plan generation using deep learning-based dose prediction. Med Phys 2022;49:1344–56. https://doi.org/10.1002/mp.15462.

[26] Khoo V, Dearnaley D. Question of dose, fractionation and technique: ingredients for testing hypofractionation in prostate cancer–the CHHiP trial. Clin Oncol 2008;20:12–4. https://doi.org/10.1016/j.clon.2007.10.008.

[27] Morgan SC, Hoffman K, Loblaw DA, Buyyounouski MK, Patton C, Barocas D, et al. Hypofractionated radiation therapy for localized prostate cancer: an ASTRO, ASCO, and AUA evidence-based guideline. J Clin Oncol 2018;36:3411. https://doi.org/10.1016/j.prro.2018.08.002.

[28] Paddick I, Lippitz B. A simple dose gradient measurement tool to complement the conformity index. J Neurosurg 2006;105:194–205. https://doi.org/10.3171/sup.2006.105.7.194.

[29] Babier A, Mahmood R, McNiven AL, Diamant A, Chan TC. Knowledge-based automated planning with three-dimensional generative adversarial networks. Med Phys 2020;47:297–306. https://doi.org/10.1002/mp.13896.

[30] Cardenas CE, McCarroll RE, Court LE, Elgohari BA, Elhalawani H, Fuller CD, et al. Deep learning algorithm for auto-delineation of high-risk oropharyngeal clinical target volumes with built-in dice similarity coefficient parameter optimization function. Int J Radiat Oncol Biol Phys 2018;101:468–78. https://doi.org/10.1016/j.ijrobp.2018.01.114.

[31] Gronberg MP, Gay SS, Netherton TJ, Rhee DJ, Court LE, Cardenas CE. Dose prediction for head and neck radiotherapy using a three-dimensional dense dilated U-net architecture. Med Phys 2021;48:5567–73. https://doi.org/10.1002/mp.14827.

[32] Bukulmez T, Ozdemir B. The Effect of Pencil Beam, Collapsed Cone, and Monte Carlo Algorithms on Dose-Volume Parameters in Esophagus Cancer. The Digital

Phantom Study Turk J Oncol 2021;36:191–9. https://doi.org/10.5505/tjo.2021.2517.

[33] Hasenbalg F, Neuenschwander H, Mini R, Born E. Collapsed cone and analytical anisotropic algorithm dose calculations compared to VMC++ Monte Carlo simulations in clinical cases. IOP Publishing: J Phys Conf Ser; 2007; Het Pand, Gent, Belgium. Belgium: 2007; p. 021007.

[34] Feygelman V, Latifi K, Bowers M, Greco K, Moros EG, Isacson M, et al. Maintaining dosimetric quality when switching to a Monte Carlo dose engine for head and neck volumetric-modulated arc therapy planning. J Appl Clin Med Phys 2022;23:e13572.