HHS Public Access

Author manuscript

Comput Med Imaging Graph. Author manuscript; available in PMC 2021 June 07.

Published in final edited form as:

Comput Med Imaging Graph. 2021 June ; 90: 101907. doi:10.1016/j.compmedimag.2021.101907.

Development and dosimetric assessment of an automatic dental artifact classification tool to guide artifact management techniques in a fully automated treatment planning workflow

Soleil Hernandez[a,b,*], **Carlos Sjogreen**[a,b], **Skylar S. Gay**[b], **Callistus Nguyen**[b], **Tucker Netherton**[a,b], **Adenike Olanrewaju**[b], **Lifei Joy Zhang**[b], **Dong Joo Rhee**[a,b], **José David Méndez**[b], **Laurence E. Court**[a,b], **Carlos E. Cardenas**[a,b]

[a]The University of Texas MD Anderson Cancer Center Graduate School of Biomedical Sciences, Houston, TX, USA

[b]The University of Texas MD Anderson Cancer Center, Department of Radiation Physics, Houston, TX, USA

Abstract

Purpose: We conducted our study to develop a tool capable of automatically detecting dental artifacts in a CT scan on a slice-by-slice basis and to assess the dosimetric impact of implementing the tool into the Radiation Planning Assistant (RPA), a web-based platform designed to fully automate the radiation therapy treatment planning process.

Methods: We developed an automatic dental artifact identification tool and assessed the dosimetric impact of its use in the RPA. Three users manually annotated 83,676 head-and-neck (HN) CT slices (549 patients). Majority-voting was applied to the individual annotations to determine the presence or absence of dental artifacts. The patients were divided into train, cross-validation, and test data sets (ratio: 3:1:1, respectively). A random subset of images without dental artifacts was used to balance classes (1:1) in the training data set. The Inception-V3 deep learning model was trained with the binary cross-entropy loss function. With use of this model, we automatically identified artifacts on 15 RPA HN plans on a slice-by-slice basis and investigated three dental artifact management methods applied before and after volumetric modulated arc

[*]Corresponding author at: Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX, 77030, USA. shernandez6@mdanderson.org (S. Hernandez).

Author statement

Soleil Hernandez: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing-Original Draft, Writing-Review & Editing, Visualization. **Carlos Sjogreen, MS**: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing-Review &Editing. **Skylar S. Gay**: Methodology, Software, Resources, Data Curation, Writing-Review & Editing. **Callistus Nguyen, PhD**: Software, Resources, Writing-Review and Editing. **Tucker Netherton, DMP**: Software, Resources, Writing-Review and Editing, **Adenike Olanrewaju, M.B.A., CM.D.:** Resources, Data Curation, Supervision, Writing-Review & Editing. **Lifei Joy Zhang, PhD**: Software, Resources, Writing-Review & Editing. **Dong Joo Rhee, MS**: Software, Resources, Writing-Review & Editing. **Jose Mendez**: Data Curation, Writing-Review & Editing. **Laurence E. Court, PhD**: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing-Review & Editing, Supervision, Project administration, funding acquisition. **Carlos E. Cardenas, PhD**: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing-Original Draft, Writing-Review & Editing, Visualization, Supervision, Project administration, funding acquisition.

Declaration of Competing Interest

The authors report no declarations of interest.

therapy (VMAT) plan optimization. The resulting dose distributions and target coverage were quantified.

Results: Per-slice accuracy, sensitivity, and specificity were 99 %, 91 %, and 99 %, respectively. The model identified all patients with artifacts. Small dosimetric differences in total plan dose were observed between the various density-override methods (±1 Gy). For the pre- and post-optimized plans, 90 % and 99 %, respectively, of dose comparisons resulted in normal structure dose differences of ±1 Gy. Differences in the volume of structures receiving 95 % of the prescribed dose (V95[%]) were ≤0.25 % for 100 % of plans.

Conclusion: The dosimetric impact of applying dental artifact management before and after artifact plan optimization was minor. Our results suggest that not accounting for dental artifacts in the current RPA workflow (where only post-optimization dental artifact management is possible) may result in minor dosimetric differences. If RPA users choose to override CT densities as a solution to managing dental artifacts, our results suggest segmenting the volume of the artifact and overriding its density to water is a safe option.

Keywords

Automation; Deep learning; Dental artifact; CT scan; Treatment planning; Radiation therapy

1. Introduction

A dental artifact is a streaking present on CT images caused by dental fillings, crowns, or other hardware. Since about 2.3 million implant-supported crowns are made annually in the United States, streaking artifacts occur frequently in routine CT imaging (American College of Prosthodontists, 2020). Such artifacts negatively affect the visibility of anatomical structures and the definition of their density on CT scans. Consequently, such factors complicate a physician’s ability to identify, delineate, diagnose, and treat anatomical structures near the oral cavity.

Specifically, in radiation oncology, CT image quality influences the quality of photon and proton radiation therapy planning. Dental artifacts result in incorrect Hounsfield unit (HU) values, which propagate to tissue density errors and consequently can introduce contouring and dose calculation errors in the treatment planning system (TPS) (Chu et al., 2000; Kilby et al., 2002). Kim et al. (Kim et al., 2006) reported that in HN photon radiation therapy plans, dental artifacts led to hot spots in organs at risk (OARs) and to cold spots in controlled tumor volumes. In the head-and-neck (HN) region, Hansen et al. (Hansen et al., 2017) found that dental artifacts that obstructed normal anatomy led to increased contoured volume of the gross tumor volume by 22 % and of the parotid glands by 7 %. In addition, the presence of dental artifacts led to increased contouring variability among individual physicians (Hansen et al., 2017).

Dental artifacts can introduce errors in proton beam radiation therapy as well. In this therapy, the calculations of proper beam range, sharp dose falloff, and optimal target coverage depend heavily on the density of tissues within the modeled beam path. Incorrect HUs in a proton therapy simulation CT scan can propagate to incorrect calculations of the

relative linear stopping power. Such errors in turn propagate to dose calculation errors and negatively affect the quality of the proton therapy (Branco et al., 2020; Richard et al., 2015). In a study by Huang et al., dental artifacts led to a 10 % error in the calculation of the dose to clinical target volumes located near the oral cavity (Huang et al., 2016).

The clinical effect that dental artifacts have on the quality of radiation therapy reveals a need for robust management techniques for dental artifacts. Several methods have been developed to reduce the effect that dental artifacts have on CT scan visualization and tissue density calculation. Kim et al. proposed acquiring an additional CT scan to obtain supplementary data to address the dental artifacts present in the original scan (Kim et al., 2019). However, acquiring additional imaging data is not cost-efficient. Alternatively, streaking artifacts may be reduced by surgically replacing metal implants with an inert composite material to better match the density of native teeth. However, this process is time-consuming, (1 h per tooth), and introduces an additional surgical procedure for the patient (Richard et al., 2015).

Additional vendor-specific artifact management techniques include dental artifact reduction reconstruction software, the use of dual energy CT, or a combination of these techniques (Cha et al., 2017; Hegazy et al., 2018; Huang et al., 2015). The magnitude of the benefit of applying dental artifact reduction techniques is still being investigated. Kidoh et al. (Kidoh et al., 2014) found that metal artifact reduction algorithms such as O-MAR significantly lowered the amount of image noise of the soft tissue of the oral cavity in corrected HN images compared with the noise in uncorrected images. However, in a study comparing three commercial metal artifact reduction algorithms, Huang et al. (Huang et al., 2015) found that the algorithms reduced the severity of orthopedic implant artifacts but did not show a significant benefit for dental artifacts. In addition, their results indicated that when the algorithms are applied to slices containing dental artifacts, the algorithms may introduce additional artifacts on images adjacent to those with artifacts present.

In radiation therapy treatment planning, dental artifacts are further managed after image acquisition to address errors in HUs. Such management techniques include not correcting the artifact (i.e., no action), segmenting the volume of the artifact on the CT scan and overriding the corresponding tissue density to water, or segmenting the entire slice with the artifact present and overriding the slice tissue density to water (Kisling et al., 2018). Kisling, et al. (Kisling et al., 2018) conducted a survey of common medical physics practice patterns in 2018 and found that 75 % of surveyed clinics segment the artifact only and perform a manual density override to the segmented volume. To date, managing dental artifacts in a TPS is a manual process that occurs before plan optimization.

Automation techniques to alleviate busy workflows and improve variability within and between physician groups have shown great success in medical imaging and radiation oncology (“Artificial intelligence and medical imaging, 2018: French Radiology Community white paper,” 2018; Suzuki, 2017; Thompson et al., 2018; Weidlich and Weidlich, 2018). For example, the Radiation Planning Assistant (RPA) is a web-based platform designed to fully automate the contouring and radiation therapy planning process for multiple treatment sites (e.g., HN, chest wall, female pelvis) (Court et al., 2018). The plan-generation process begins with a user uploading a CT scan and planning directive (treatment site, prescription,

etc.). Then, the RPA automatically performs dose optimization and dose calculation by using the Eclipse Treatment Planning System (Varian Medical Systems, 2017). In addition, the RPA has advanced artificial intelligence–based tools that perform QA and automatically identify errors, which are flagged for the user throughout the contouring and planning process. Once the RPA has automatically generated a treatment plan, the user must recalculate the treatment plan in their own TPS before using the plan clinically.

As automated treatment planning algorithms such as the RPA continue to advance, it is important to consider how to develop an automated method to mirror the management of dental artifacts in a typical clinical workflow. Recent work by Welch et al. (Welch et al., 2020) detects the presence or absence of dental artifact within a patient's entire CT scan. The tool achieved a PR-AUC of 0.92 ± 0.03. However, localizing the exact CT slices where the artifact occurs is beyond this approaches' capability. Current dental artifact management techniques are applied to each slice of a CT scan where the artifact is present; Therefore, classifying dental artifact on a slice-by-slice basis is critical to designing a dental artifact management technique compatible with an automated treatment planning workflow. To our knowledge, there has yet to be a study to develop and assess a tool capable of automatically classifying dental artifacts on a slice-by-slice basis. In addition, there has not been a study to test the dosimetric effect of using an automatic dental artifact classification tool to guide artifact management techniques in the TPS.

Therefore, the purpose of this study was to (1) develop and investigate the dosimetric impact of a robust slice-by-slice dental artifact classification tool and (2) determine a robust management strategy of dental artifacts within the RPA workflow.

Although this study was designed to specifically address dental artifact management in the RPA, the classification tool may be applied to other studies that would benefit from automatically identifying dental artifacts. An example would be analyzing the effect of metal artifact reduction algorithms on radiomics studies (Ger et al., 2018). In addition, the results of the dosimetry study may be applied to a manual treatment planning workflow.

2. Materials and methods

A convolutional neural network (CNN)-based tool was developed to automatically identify the presence or absence of dental artifacts on each slice of a HN CT scan. The output of the tool was then used in a study to investigate the dosimetric effect of implementing such a tool to guide artifact management techniques in an automated treatment planning workflow.

2.1. CNN-based automatic classification tool

CT scans from 549 patients with HN cancer previously treated with radiation therapy at The University of Texas MD Anderson Cancer Center were collected for this study under an IRB-approved protocol (PA16−0379). The total number of CT slices within the scans of these patients was 83,676. Image details are outlined in Table 1. The image sets were converted from DICOM to NumPy arrays to be compatible with the CNN. This conversion was done by using various python libraries including SimpleITK, Pydicom, cv2, and scikit-image (Bradski, 2000; Lowekamp et al., 2013; "pydicom documentation — pydicom 2.1.2

documentation, 2021," n.d.; Van Der Walt et al., 2014). In this process, patient-specific metadata such as name, gender, and age were removed, whereas scan-specific metadata such as pixel spacing, slice thickness, and image orientation were preserved. To robustly curate the data set, each slice was manually annotated by three reviewers as to whether the artifact was present or not present with the use of an in-house annotation software. The anonymized CT scans were imported into the software, shuffled, and viewed with a window width of 250 HU and window level of 0 HU. For each patient, the software allowed each annotator to visualize individual CT slices and assign "yes" or "no" to indicate the presence or absence of the artifact. The annotations assigned to each slice by each user were saved for each patient. To improve the agreement of individual annotations on the ground truth data set, majority-voting was applied to the individual annotations to define the presence or absence of a dental artifact on each CT slice. Having three independent users collectively annotate each slice of each scan increased the robustness of our dataset.

The 83,676 CT slices obtained from 549 patients were split on a per-patient basis into training, validation, and test data sets at a ratio of 3:1:1. Since artifacts were present in only 1843 of 49,843 (<4%) of the slices in the final training set, a random subset of CT slices without dental artifacts was sampled to balance the classes to a 1:1 ratio, resulting in a more robust training data set of 3686 slices. No such sampling was performed on the test set in order to reflect the actual clinical environments in which most CT slices will not contain dental artifacts.

Before training the network, the data was preprocessed with multiple image augmentations to further improve the robustness of the training process. The image augmentations increased the size of the data set and mimicked common variations found in clinical CT scans. The augmentations performed included ZCA whitening (Bell and Sejnowski, 1997), a rotation of 30°, a horizontal and vertical shift range of 0.2 pixels, a counterclockwise shear angle of 0.3°, a zoom range of 0.3, a horizontal flip, and an intensity rescaling factor such that image intensities were within [0, 1].

The Inception-V3 (Szegedy et al., 2016) image classification architecture was trained from scratch to automatically classify the presence or absence of dental artifacts on a slice-by-slice basis. The Inception-V3 architecture consists of 311 layers and 23,885,392 parameters. The trained size of the architecture is 97 MB. The architecture differs from its inception predecessors in that it uses label smoothing, factorized convolutions, and batch normalization. Factorized convolutions allow the Inception-V3 architecture to be 33 % less expensive computationally than the previous iterations of the architecture.

An NVIDIA station with V100 GPUs was used to train the Inception-V3 model. The batch size of the model consisted of 24 slices each with an image size of 512 × 512. The binary cross-entropy loss function and Adam optimizer were implemented during training. The learning rate, $\beta 1$, and $\beta 2$ of the Adam optimizer were set to 0.001, 0.900, and 0.999, respectively.

The performance of the model was evaluated on a per-slice basis on the final test set of 16,700 CT slices with use of receiver-operator characteristic analysis (accuracy, sensitivity,

and specificity). The efficiency of the model was also assessed by calculating the average classification time per CT slice and per patient.

2.2. Artifact management: dosimetric evaluation

In the current version of the RPA, dental artifacts are not accounted for automatically. A user may manage dental artifact after the optimized plan has been generated by the RPA. Users import the automatically generated plan into their TPS, manage the dental artifacts according to their clinics' practice, and then recalculate the post-optimized plan. However, translating the typical clinical artifact management workflow into an automated one would require managing artifacts before plan optimization.

To assess how the classification tool could best guide artifact management in an automated HN treatment planning workflow, we applied three common artifact management techniques before and after plan optimization. We quantified the dosimetric impact of each management technique and at each stage of potential implementation. Our dosimetry study design is visually outlined in Fig. 1.

2.2.1. Automated plan generation—The data for the dosimetry study consisted of 15 VMAT HN patients whose treatment plans were automatically generated by using the RPA. The 15 patients used in the dosimetry study were not used to train the artifact classification model. These plans were intended to treat the following disease sites: nasopharynx, hypopharynx, base of tongue, and tonsil. Each plan had up to 3 planned target volumes (PTVs). The doses prescribed ranged from 60− 70 Gy for PTV1, 56− 63 Gy for PTV2, and 54− 57 Gy for PTV3 (if any). The prescription dose delivered for each plan ranged from 60 to 70 Gy in 33 fractions (1.82–2.12 Gy/fx). Each plan consisted of 3 treatment arcs with collimator rotations of 90, 345, and 15 degrees. In the current workflow of the RPA, dental artifacts are not accounted for automatically (i.e., "no action" approach).

2.2.2. Study design—To perform the dosimetry study, the dental artifact classification tool was used to predict the presence or absence of artifacts on each slice from all 15 study patients. These results were then used to quantify the dosimetric impact of the tool by using the 5 scenarios previously described.

2.2.2.1. Dosimetric impact of applying artifact management after plan optimization.: We assessed the dosimetric impact of applying artifact management techniques after plans had been optimized to simulate potential implementation of this tool into the current workflow of the RPA. In this implementation, the user is given an optimized treatment plan so any dental artifact management would occur after plan optimization. To simulate this workflow, two new segmentation volumes, segment (Fig. 2A) or slice (Fig. 2B), were contoured on the slices of the plan where the classification tool predicted the presence of an artifact. After the segmentation of each respective volume, two plans were replicated from the original, resulting in a total of three post-optimized plans with two additional segmentation volumes on each plan. In plans two and three, the density of the segmented artifact volume (Fig. 2A) and entire patient anatomy on a slice (Fig. 2B), respectively, were overridden to that of water. This workflow is outlined in Fig. 1.

To compare the dosimetric impact of applying each artifact management method, the dose was recalculated with the same monitoring units for both plans. This experiment assessed applying two different artifact management techniques after plan optimization, which mirrors incorporating dental artifact management into the current version of the RPA.

2.2.2.2. Dosimetric impact of applying artifact management before plan optimization.: We assessed the dosimetric effect of applying artifact management techniques before plan optimization. This implementation is intended to assess the feasibility of translating the typical clinical workflow (management of the dental artifact before plan optimization) into an automated workflow. In an automated workflow, the system would automatically perform a dental artifact management technique. Next, an oncologist would approve the dental artifact management method performed and send the plan back into the RPA for optimization. The plan would then be reviewed by the dosimetrist before delivery. To simulate this workflow, two artifact management techniques were applied to plans that had not been optimized by the RPA. To define the segmentation volumes of the artifact alone and the slice, the structure set from the original plan was duplicated onto the new, non-optimized plan. The non-optimized plan was then replicated twice to generate a total of three non-optimized plans with the segmented artifact volumes. In plans two and three, the density of the segmented artifact and slice, respectively, were overridden to that of water. After the density override, plans two and three were re-optimized and compared with the original plan to quantify the effects of applying dental artifact management techniques before plan optimization.

2.2.3. Quantification of dosimetric impact—To assess the dosimetric effect of using this tool before and after plan optimization, we quantified the percent difference in maximum dose resulting from two artifact management methods relative to making no changes. To assess coverage, the percent difference in V95 of PTV1 was quantified for both management techniques applied before and after plan optimization. The percent max dose difference delivered to the brainstem, optic chiasm, and spinal cord was quantified for both management techniques applied before and after plan optimization. The percent mean dose difference was quantified for the left and right parotids for both management techniques applied before and after plan optimization. These dose metrics were used for the qualitative study since they are the common dose constraints used for VMAT optimization.

3. Results

3.1. CNN-based automatic classification tool performance

The overall accuracy, sensitivity, and specificity achieved on the final test data set were 99 %, 91 %, and 99 %, respectively, on a slice-by-slice basis. Moreover, on a patient-by-patient basis (111 total patients), the model correctly identified every patient for whom an artifact was present (89 patients). In addition, the model did not assign a dental artifact to any patient who did not have one (22 patients). On average, the model incorrectly classified the presence or absence of an artifact on 1 slice per patient. Although being trained on scans that had not been preprocessed to localize this region, the model never assigned a dental artifact outside of the oropharyngeal region. The model never misclassified a large artifact.

Misclassified artifacts occurred in cases in which there were sharp differences in density values such as in the oral cavity, with the use of tongue compression, or with the presence of non-dental artifacts in the oropharyngeal region (Fig. 3).

The classification tool proved to be very efficient. The tool was able to classify artifacts in 111 patients in <10 min by using a single GPU. The classification tool averaged 0.03 s (range: 0.02− 0.06 s) to classify each image and 5.6 s to classify all slices per patient (range: 1.84–29.0 s). The large range for the time per patient was expected, given the range of slices per patient (51− 825 slices).

3.2. Artifact management dosimetry study

The number of slices classified as having artifacts present across the 15 sampled patients ranged from 2 to 12, with an average of 6 ± 3 slices.

3.2.1. Managing dental artifact after plan optimization—Fig. 4 displays the results of the dosimetry study designed to mirror implementation of an artifact management tool in the current RPA workflow. When comparing the three artifact management techniques, our results showed that the slice artifact management method resulted in more appreciable changes in the maximum dose and the dose delivered to the brainstem and spinal cord. Specifically, applying the artifact management techniques after plan optimization resulted in an overall increase in maximum dose. The average increase in maximum dose delivered to the plan was 0.02 Gy for the segment method and 0.25 Gy for the slice method. The slice management method applied after optimization resulted in the largest increase and decrease in maximum dose delivered to the plan (+1.53 Gy and −0.097 Gy, respectively).

The percent change to PTV coverage was negligible for both artifact management techniques (0.006 % and 0.013 % for the segment and slice methods, respectively). The maximum change in PTV V95 was 0.24 % for the slice management method.

The average increase in maximum dose to the oropharyngeal structures (brainstem, spinal cord, and chiasm) was 0.012 Gy for the segment method and 0.124 Gy for the slice method. The average increase in mean dose to the parotids was 0.002 Gy for the segment method and 0.040 Gy for the slice method. The largest increase in maximum dose resulting from artifact management applied after optimization was 1.033 Gy to the brainstem for the slice method. Similarly, the largest decrease in maximum dose resulting from the slice post-optimized management method was −0.258 Gy to the chiasm.

Applying dental artifact management after plan optimization resulted in adequate coverage of the PTVs and normal structures. In addition, for each structure, the calculated dose remained within each respective OAR constraint.

3.2.2. Managing dental artifact before plan optimization—Fig. 5 displays the results of the dosimetry study designed to mirror implementation of a typical clinical workflow (management of dental artifact before plan optimization) into an automated workflow. Applying dental artifact management techniques before plan optimization resulted in minor reductions to the maximum dose delivered to the plan. The segmented

artifact method reduced the maximum dose by −0.230 Gy on average, and the segmented slice method reduced the maximum dose by −0.060 Gy on average.

The two artifact management techniques had very little influence on the coverage of the PTV. The segmented artifact method decreased coverage by 0.04 %, and the segmented slice method decreased coverage by 0.03 %. For both artifact management techniques applied before plan optimization, the difference in dose to the oropharyngeal structures remained within ±2.0 Gy. In addition, the doses delivered to all structures remained within the dose tolerances for each OAR for both artifact management techniques. Applying both techniques before plan optimization resulted in an overall increase to the doses delivered to each structure. The average increase in maximum dose delivered to the chiasm, brainstem, and spinal cord was 0.140 Gy for the segmented artifact method and 0.160 Gy for the segmented slice method. The average increase in mean dose delivered to the parotids was 0.200 Gy for the segmented artifact method and 0.020 Gy for the segmented slice method.

Overall, when comparing the dosimetric differences between pre- and post-optimized plans, 90 % and 99 %, respectively, of dose comparisons resulted in normal structure dose differences of ±1 Gy. Differences in V95 [%] to PTV1 were <0.25 % for 100 % of plans.

4. Discussion

We successfully developed a tool capable of automatically classifying the presence or absence of dental artifacts on HN CT scans on a slice-by-slice basis with 99 % accuracy. We used this tool to automatically identify the presence of artifacts on a separate cohort of patients with automatically generated treatment plans. Where the tool predicted the presence of an artifact, we contoured the artifact alone (segmented artifact) or the patient body volume (segmented slice) and then performed density overrides on the volumes to mimic manual management techniques commonly used in the clinic. We quantified the dosimetric impact of applying these techniques before and after plan optimization to simulate implementation of the tool into an automated treatment planning workflow.

Our results show that the dosimetric impact of applying dental artifact management techniques is small for both pre- and post-optimized plans. In the current RPA workflow, the user receives a post-optimized plan that has not accounted for dental artifacts (i.e., the “no action” option). Our results suggest that this approach will not have a large dosimetric impact on the automatically generated plans. However, if the user opts to override CT densities as a dental artifact management technique in the current RPA workflow, we recommend the segmented artifact method since it has a smaller dosimetric impact relative to the slice method. Our results suggest that translating the typical clinical workflow (management of artifact before plan optimization) into an automated one will result in minor differences in the dose delivered to OARs and PTV coverage.

4.1. CNN-based automatic classification tool

We manually reviewed cases in which the model incorrectly predicted the presence of dental artifacts. Some examples of this were on slices containing the oral cavity, slices containing a tongue compressor, and occasionally slices that had a non-dental artifact present in the

oropharyngeal region. We believe that in these cases, the misclassifications may be attributed to areas of sharp differences in density between two interfaces.

Overall, the classification tool was able to identify all of the patients for whom dental artifacts were present. The model never incorrectly identified an artifact in a patient without an artifact.

4.2. Artifact management dosimetry study

The dose differences calculated between the three most common dental artifact management techniques were small. We suspect that this was due to the imbalance in the number of slices containing artifacts relative to the total number of CT slices in a patient's image.

We reported a more appreciable dose difference to the oropharyngeal structures for which dental artifact management techniques were applied before plan optimization compared with management techniques applied after plan optimization. This was expected since the current clinical workflow consists of managing dental artifacts with the segmented density override method before plan optimization. Applying density overrides presumably resulted in more correct intensity values for those structures adjacent to dental artifacts. When corrections were made before plan optimization, the TPS had more freedom to achieve the desired dose distribution than when management techniques were applied after the plan had already been optimized.

Applying the slice artifact management technique before plan optimization mirrors previously published work by Kim et al. (Kim et al., 2006). In their study, a machine learning algorithm was used to classify each structure within a CT image as soft tissue, bone, or metal. In one of four tested artifact management techniques, all pixels that were identified as belonging to these classes were set to have a HU value of water. This technique resulted in an increase to the maximum dose to the spinal cord by 0.60 ± 0.30 Gy. Our work resulted in a 0.35 ± 0.50 Gy increase in maximum dose delivered to the spinal cord, which is within 1 standard deviation of the previously published literature.

In the current version of the RPA, the user receives a plan that has already been optimized. Our results suggest that not accounting for dental artifacts in the current RPA workflow may result in only minor dosimetric differences when considering target volume and OAR doses. However, if RPA users choose to perform CT density overrides to manage dental artifacts, we suggest implementing the segmented artifact technique.

If a typical clinical workflow were translated into an automated workflow, the user would have the ability to apply dental artifact management before plan optimization. Our results suggest that not accounting for artifacts or performing the segmented slice technique would also be safe options for artifact management since neither method caused an appreciable dosimetric impact to OARs or PTV coverage.

Structures such as the spinal cord and brainstem proved to be the most susceptible to dosimetric changes caused by dental artifact management techniques applied before plan optimization. We presume this is because both structures are present on slices that may contain the oral cavity. Therefore, applying a density override to either a segmented region

or the entire slice may change attenuation estimates, which results in an altered dose distribution.

While this study was designed to cater to an automated treatment planning workflow, the classification tool may be applied to other radiation oncology research that may benefit from automatically classifying artifacts on CT images. An example of this would be automatically identifying artifact streaking in a CT image to assess the impact of dental artifacts on radiomics studies (Ger et al., 2018). Another example would be automatically identifying external fiducials in a CT image by using a similar approach. Similarly, whereas the methods of the dosimetry study were specifically designed to cater to an automated treatment planning workflow, the study results may be extended to manual workflows as well. Currently, dental artifacts are most commonly managed in the clinic by contouring the segment of artifact in a CT image and performing a manual density override (Kisling et al., 2018). Therefore, the results of the comparison of various artifact management methods applied before plan optimization are applicable to a manual treatment planning workflow.

Our assessment of overall classification model performance is limited by the fact that our model was only trained, validated, and tested on an internal HN data set. In addition, the plans generated to perform the dosimetry study were created by the RPA, which is an internal interface that our group developed. Further, the results of the dosimetry study were limited to assess the dosimetric impact of three common artifact management methods on five oropharyngeal structures.

5. Conclusions

We successfully developed an automatic dental artifact classification tool and assessed its potential implementation into an automated treatment planning workflow. Dose differences across multiple density-override methods were small. Minor dosimetric differences to the selected OARs and PTV coverage were observed when applying dental artifact management before and after plan optimization. Therefore, in the current RPA workflow, our results suggest that not accounting for dental artifacts would have a minor dosimetric impact on the generated plans. However, if the RPA user elects to implement an artifact management technique, we suggest applying the segmented artifact method. Moreover, translating a typical clinical workflow (managing the dental artifact before plan optimization) into an automated workflow would have a minor dosimetric impact to the OARs and PTV coverage. Finally, the dental artifact classification tool may be applied to other studies in radiation oncology, and the results of the dosimetry study may be applied to manual treatment planning workflows.

Acknowledgements

We acknowledge the support of the High-Performance Computing for research facility at The University of Texas MD Anderson Cancer Center for providing computational resources that have contributed to the research results reported in this article. We also thank Tamara Locke, Scientific Editor, Research Medical Library, for editing this article.

Funding

This work was supported by NIH/NCI grants UH2/UH3 CA202665 (to B.M.B. and L.E.C.) and P30CA016672(L.E.C.).

References

American College of Prosthodontists, 2020. Facts and Figures [WWW Document]. URL (Accessed 3.8.20). https://www.gotoapro.org/facts-figures/.

Artificial intelligence and medical imaging, 2018. French radiology community white paper., 2018. Diagn. Interv. Imaging 99, 727–742. 10.1016/j.diii.2018.10.003. [PubMed: 30470627]

Bell AJ, Sejnowski TJ, 1997. The “independent components” of natural scenes are edge filters. Vision Res. 37, 3327–3338. 10.1016/S0042-6989(97)00121-1. [PubMed: 9425547]

Bradski G, 2000. The OpenCV library. Dr. Dobb’s J. Softw. Tools.

Branco D, Kry S, Taylor P, Rong J, Zhang X, Peterson C, Frank S, Followill D, 2020. Development of a stereoscopic CT metal artifact management algorithm using gantry angle tilts for head and neck patients. J. Appl. Clin. Med. Phys. 21, 120–130. 10.1002/acm2.12922.

Cha J, Kim H-J, Kim ST, Kim YK, Kim HY, Park GM, 2017. Dual-energy CT with virtual monochromatic images and metal artifact reduction software for reducing metallic dental artifacts. Acta radiol. 58, 1312–1319. 10.1177/0284185117692174. [PubMed: 28273739]

Chu JCH, Ni B, Kriz R, Amod Saxena V, 2000. Applications of simulator computed tomography number for photon dose calculations during radiotherapy treatment planning. Radiother. Oncol. 55, 65–73. 10.1016/S0167-8140(00)00159-6. [PubMed: 10788690]

Court LE, Kisling K, McCarroll R, Zhang L, Yang J, Simonds H, du Toit M, Trauernicht C, Burger H, Parkes J, Mejia M, Bojador M, Balter P, Branco D, Steinmann A, Baltz G, Gay S, Anderson B, Cardenas C, Jhingran A, Shaitelman S, Bogler O, Schmeller K, Followill D, Howell R, Nelson C, Peterson C, Beadle B, 2018. Radiation planning assistant - a streamlined, fully automated radiotherapy treatment planning system. J. Vis. Exp. 10.3791/57411.

Ger RB, Craft DF, Mackin DS, Zhou S, Layman RR, Jones AK, Elhalawani H, Fuller CD, Howell RM, Li H, Stafford RJ, Court LE, 2018. Practical guidelines for handling head and neck computed tomography artifacts for quantitative image analysis. Comput. Med. imaging Graph. Off. J. Comput. Med. Imaging Soc. 69, 134–139. 10.1016/j.compmedimag.2018.09.002.

Hansen CR, Christiansen RL, Lorenzen EL, Bertelsen AS, Asmussen JT, Gyldenkerne N, Eriksen JG, Johansen J, Brink C, 2017. Contouring and dose calculation in head and neck cancer radiotherapy after reduction of metal artifacts in CT images. Acta Oncol. (Madr). 56, 874–878. 10.1080/0284186X.2017.1287427.

Hegazy MAA, Eldib ME, Hernandez D, Cho Myung Hye, Cho, Hyoung Min, Lee SY, 2018. Dual-energy-based metal segmentation for metal artifact reduction in dental computed tomography. Med. Phys. 45, 714–724. 10.1002/mp.12719. [PubMed: 29220087]

Huang JY, Kerns JR, Nute JL, Liu X, Balter PA, Stingo FC, Followill DS, Mirkovic D, Howell RM, Kry SF, 2015. An evaluation of three commercially available metal artifact reduction methods for CT imaging. Phys. Med. Biol. 60, 1047–1067. 10.1088/0031-9155/60/3/1047. [PubMed: 25585685]

Huang JY, Followill DS, Howell RM, Liu X, Mirkovic D, Stingo FC, Kry SF, 2016. Approaches to reducing photon dose calculation errors near metal implants. Med. Phys. 43, 5117–5130. 10.1118/1.4960632. [PubMed: 27587042]

Kidoh M, Nakaura T, Nakamura S, Tokuyasu S, Osakabe H, Harada K, Yamashita Y, 2014. Reduction of dental metallic artefacts in CT: value of a newly developed algorithm for metal artefact reduction (O-MAR). Clin. Radiol. 69, e11–e16. 10.1016/j.crad.2013.08.008. [PubMed: 24156796]

Kilby W, Sage J, Rabett V, 2002. Tolerance levels for quality assurance of electron density values generated from CT in radiotherapy treatment planning. Phys. Med. Biol. 47, 1485–1492. 10.1088/0031-9155/47/9/304. [PubMed: 12043814]

Kim Y, Tomé WA, Bal M, McNutt TR, Spies L, 2006. The impact of dental metal artifacts on head and neck IMRT dose distributions. Radiother. Oncol. 79, 198–202. 10.1016/j.radonc.2006.03.022. [PubMed: 16677729]

Kim C, Pua R, Lee CH, Choi Din, Cho B, Lee Swook, Cho S, Kwak J, 2019. An additional tilted-scan-based CT metal-artifact-reduction method for radiation therapy planning. J. Appl. Clin. Med. Phys. 20, 237–249. 10.1002/acm2.12523. [PubMed: 30597725]

Kisling KD, Ger RB, Netherton TJ, Cardenas CE, Owens CA, Anderson BM, Lee J, Rhee DJ, Edward SS, Gay SS, He Y, David SD, Yang J, Nitsch PL, Balter PA, Urbauer DL, Peterson CB, Court LE, Dube S, 2018. A snapshot of medical physics practice patterns. J. Appl. Clin. Med. Phys. 10.1002/acm2.12464.

Lowekamp BC, Chen DT, Ibáñez L, Blezek D, 2013. The design of SimpleITK. Front. ˜ Neuroinform. 7, 45. 10.3389/fninf.2013.00045.

pydicom documentation — pydicom 2.1.2 documentation [WWW Document], n.d. URL https://pydicom.github.io/pydicom/stable/ (accessed 1.11.21).

Richard P, Sandison G, Dang Q, Johnson B, Wong T, Parvathaneni U, 2015. Dental amalgam artifact: adverse impact on tumor visualization and proton beam treatment planning in oral and oropharyngeal cancers. Pract. Radiat. Oncol. 5, e583–e588. 10.1016/j.prro.2015.04.007. [PubMed: 26419441]

Suzuki K, 2017. Overview of deep learning in medical imaging. Radiol. Phys. Technol. 10, 257–273. 10.1007/s12194-017-0406-5. [PubMed: 28689314]

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna ZB, 2016. Rethinking the Inception Architecture for Computer Vision. 10.1109/CVPR.2016.308.

Thompson RF, Valdes G, Fuller CD, Carpenter CM, Morin O, Aneja S, Lindsay WD, Aerts HJWL, Agrimson B, Deville CJ, Rosenthal SA, Yu JB, Thomas CRJ, 2018. Artificial intelligence in radiation oncology: a specialty-wide disruptive transformation? Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol. 129, 421–426. 10.1016/j.radonc.2018.05.030.

Van Der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T, 2014. Scikit-image: image processing in python. PeerJ 2014. 10.7717/peerj.453.

Varian Medical Systems, 2017. Eclipse Treatment Planning System [WWW Document]. URL. https://www.varian.com/products/radiosurgery/treatment-planning/eclipse.

Weidlich V, Weidlich GA, 2018. Artificial intelligence in medicine and radiation oncology. Cureus 10, e2475. 10.7759/cureus.2475.

Welch ML, McIntosh C, Purdie TG, Wee L, Traverso A, Dekker A, Haibe-Kains B, Jaffray DA, 2020. Automatic classification of dental artifact status for efficient image veracity checks: effects of image resolution and convolutional neural network depth. Phys. Med. Biol. 65, 15005. 10.1088/1361-6560/ab5427.

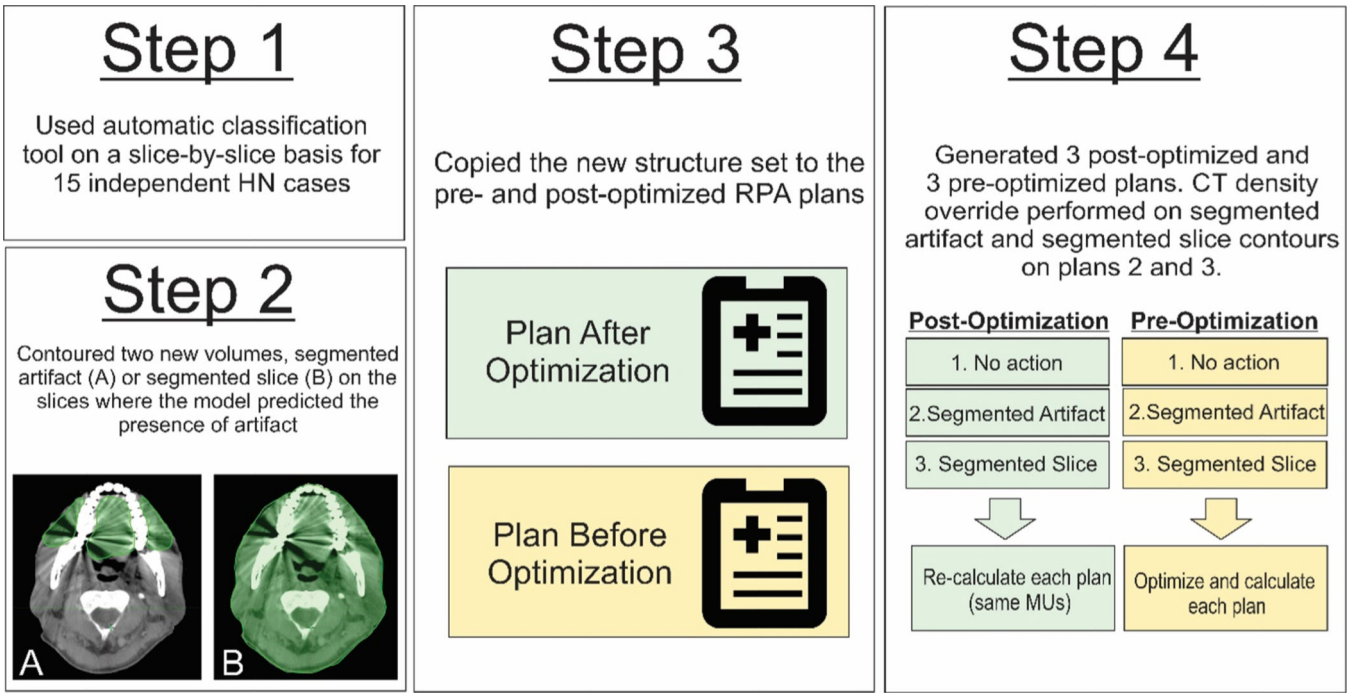

Fig. 1.

Outline of the dosimetry study workflow. First, the contouring tool was used to automatically classify the presence or absence of dental artifacts in 15 patients with HN cancer whose treatment plans were automatically generated. Where the tool predicted the presence of an artifact, two new contour volumes were created on each slice. The segmented artifact alone (Step 2A) and the segmented external patient volume, slice (Step 2B) were contoured and added to the structure set. The structure set was copied onto plans before and after plan optimization. The pre- and post-optimized plans were duplicated to create 3 post-optimized and 3 pre-optimized plans. On the second of the three plans, the density of the segmented artifact was overridden to water. On the third of the three plans, the density of the slice was overridden to water. Each of the six plans were recalculated.

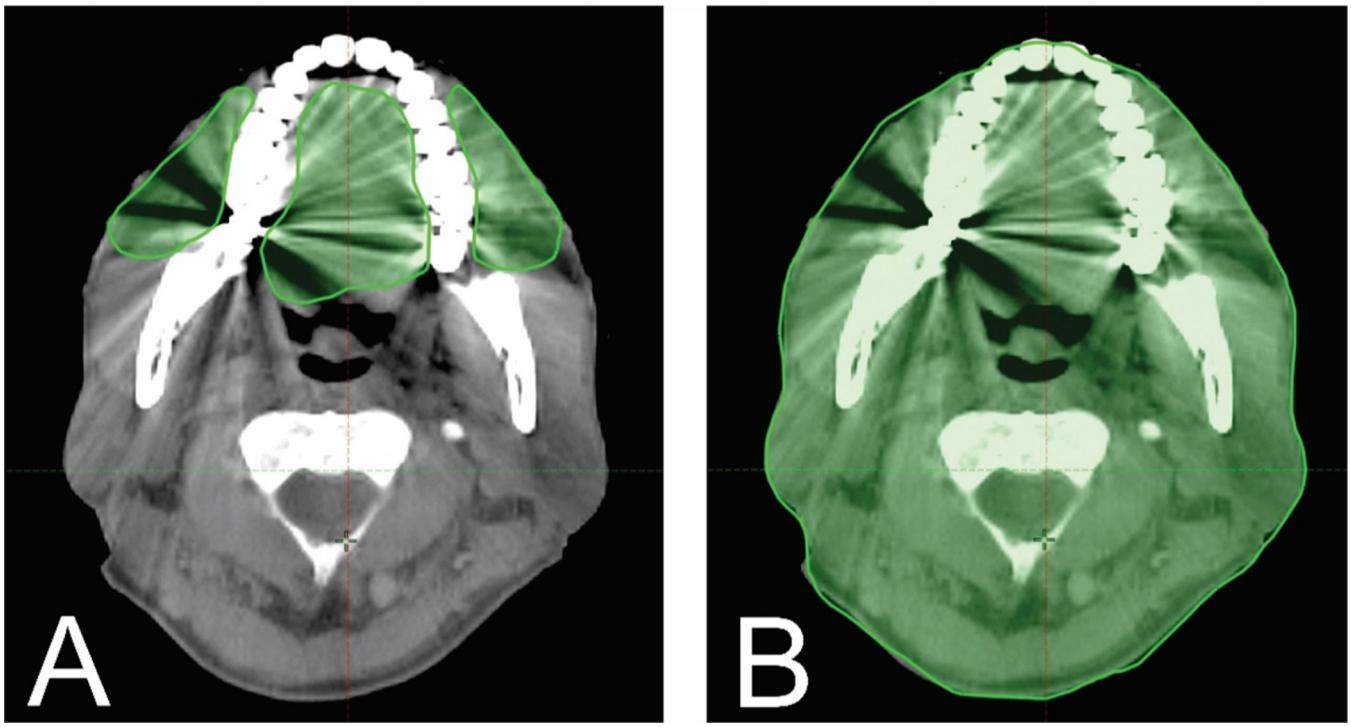

Fig. 2.

Visualization of artifact management techniques. Techniques two and three of the three types of artifact management techniques used in the dosimetry study (the first technique is the “no action” approach). Contouring the artifact only (A) or the entire slice (B) and overriding the density of the contoured volume to water addresses the incorrect HU values in the area of artifact.

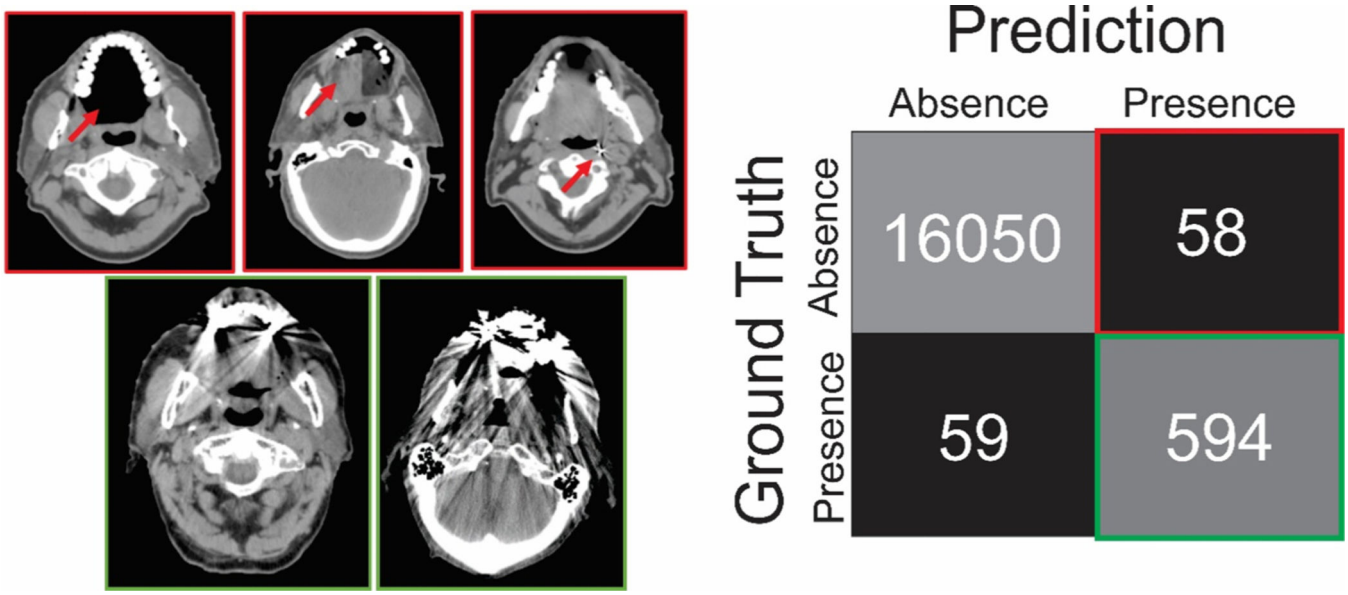

Fig. 3.

Confusion matrix displaying model performance on the final test data set. Red outlines highlight examples of cases in which the model incorrectly predicted the presence of artifacts. From left to right, the cases shown: oral cavity space, tongue compression, and non-dental artifact in the oropharyngeal region. The green outlines highlight examples of cases of large artifacts, which the model successfully classified.

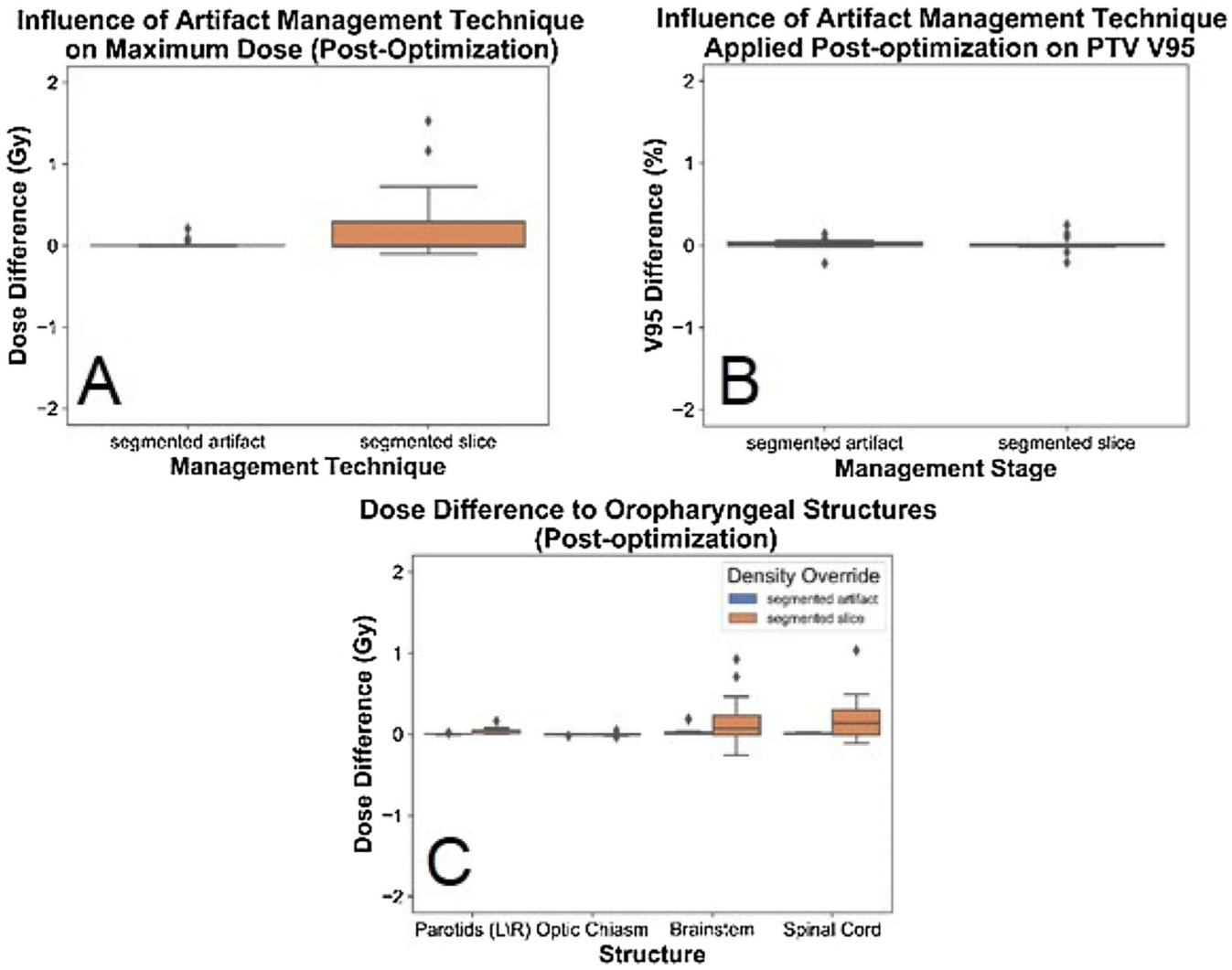

Fig. 4.

Dosimetry study results for artifact management techniques applied after plan optimization. A: Change in maximum dose delivered to the plan for both techniques. B: Changes in PTV V95 for both management techniques. C: Dose difference to oropharyngeal structures for both techniques (mean dose for parotids, maximum dose for all other structures). All calculations are relative to the no action approach.

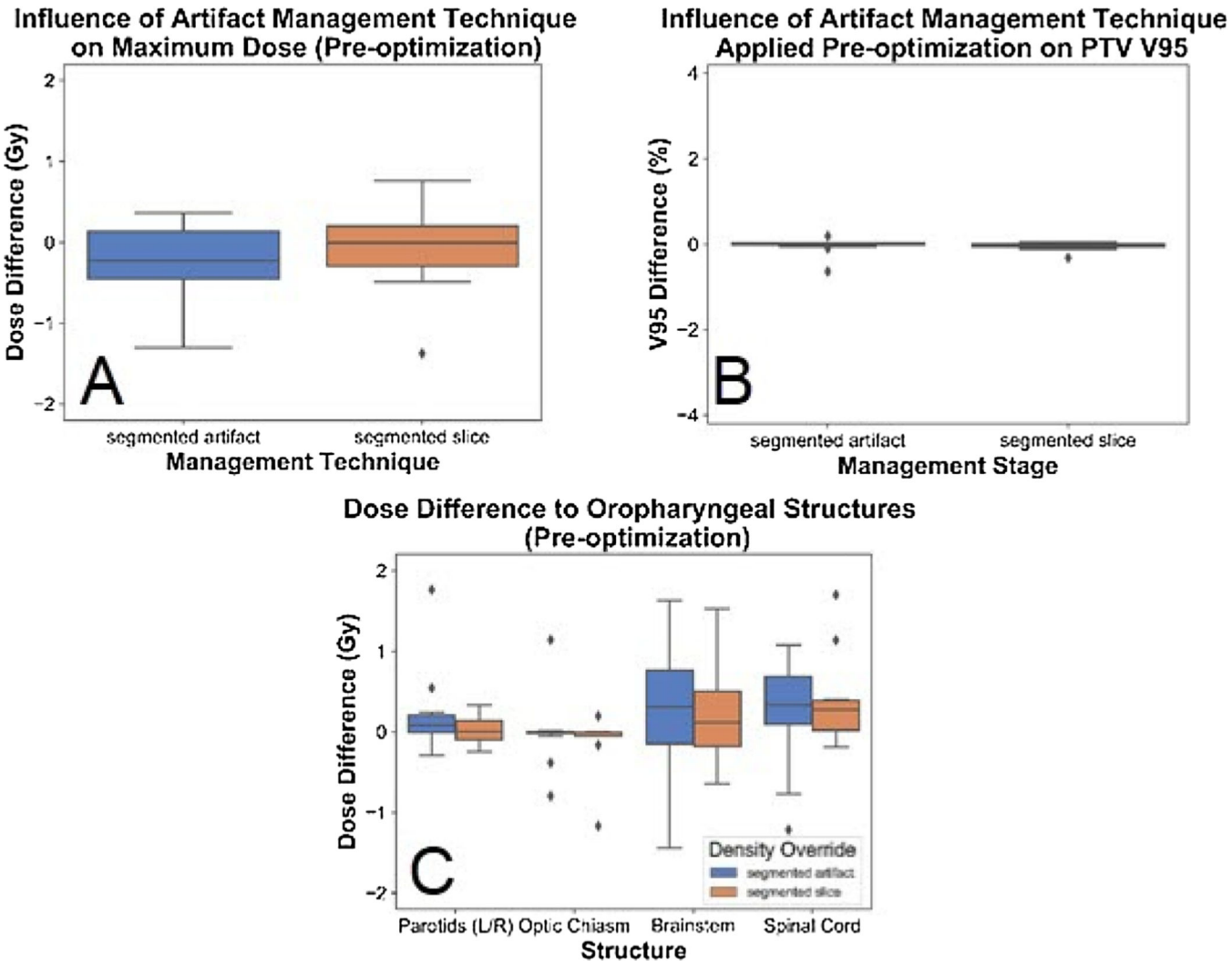

Fig. 5.

Dosimetry study results for artifact management techniques applied before plan optimization A: Change in maximum dose delivered to the plan for both techniques. B: Changes in PTV V95 for both management techniques. C: Dose difference to oropharyngeal structures for both techniques (mean dose for parotids, maximum dose for all other structures). All calculations are relative to the no action approach.

Table 1

Details of the 549 images used to train the classification tool.

Median number of slices and range	Median slice thickness and range (mm)	Median tube voltage peak and range (kVp)	Median imaging year and range	Scanner manufacturers	
157	3	120	2010	Philips GE	514
(51–825)	(1–3.75)	(100–140)	(2004–2015)	Medical Systems	36