# From Scan to Plan: Organ-Specific Deep Learning Networks for Brachytherapy

Nirmalya Gayen[1], Aditya Kumar[2][a], Bhuneshwar Singh Netam[3][b], Nithin Shivashankar[3][c], Kirthi Koushik A.S.[4][d] and Vijay Natarajan[1][e]

[1]*Indian Institute of Science, Bangalore, India*

[2]*Uber India Systems Pvt. Ltd., Bangalore, India*

[3]*MIMYK Medical Simulations Pvt. Ltd., Bangalore, India*

[4]*HCG Cancer Centre, Bangalore, India*

*vijayn@iisc.ac.in, nithin@mimyk.com*

Keywords: Brachytherapy, CT scan data, organ segmentation, contouring, clinical validation.

Abstract: The accurate identification of anatomical structures within volumetric data derived from medical scanning devices, such as CT and MRI, is a significant aspect of clinical workflows in radiology and oncology treatment planning. With the advance of AI and high-performance computing, many methods and tools have been developed over the past decade and a half. However, end-to-end integration of solutions with existing workflows and practices remains a challenge. Here, we focus on the need for segmentation of anatomical structures whose delineations are clinically defined by a combination of anatomy, function, and treatment planning. Existing deep learning approaches for segmentation often struggle to effectively differentiate closely placed organs, such as the bladder, rectum, and sigmoid colon. We propose an efficient and robust, organ-specific segmentation pipeline based on tailored 2D U-Net models, coupled with anatomy-guided preprocessing and geometric postprocessing algorithms. We validate our method by a user study involving trained radiation oncologists, demonstrating high segmentation accuracy and significant reductions in contouring time. The results show that our approach produces consistently accurate contours that closely match expert delineations, with minimal corrections needed in clinical practice. This work highlights the benefits of deep learning integration in brachytherapy, enabling quicker planning and improved consistency through clinically validated organ segmentation.

## 1 INTRODUCTION

Organ segmentation from computed tomography (CT) scans is fundamental in radiation therapy procedures such as cervical cancer brachytherapy (Mahantshetty et al., 2019). Accurate delineation of organs-at-risk (OARs), including the bladder, rectum, and sigmoid colon, is critical to ensure effective tumor targeting while minimizing exposure to surrounding healthy tissues (Derks et al., 2018). Despite significant technological advancements, manual segmentation remains prevalent in clinical practice, posing challenges related to inter-observer variability, labor intensity, and prolonged planning time. Organ segmentation is a challenging problem due to the complex and variable appearance of organs, variations in imaging modalities, noise, artifacts, and anatomical differences across individuals (Skowronek, 2017). Generalized multi-organ models often fail in the context of cervical cancer brachytherapy due to inadequate handling of anatomical intricacies, gender-specific differences, and inconsistencies in small-organ segmentation (Bandyk et al., 2021; Perslev et al., 2019).

While numerous segmentation models have been proposed, their adoption in clinical practice remains limited by high computational requirements and long inference times. Our focus is on translating these advances into a deployable tool that runs efficiently on standard hardware while preserving segmentation accuracy. Addressing these limitations, we introduce a clinically validated, organ-specific segmenta-

[a] https://orcid.org/0009-0006-8152-0762
[b] https://orcid.org/0009-0006-6348-4585
[c] https://orcid.org/0009-0004-2344-9510
[d] https://orcid.org/0000-0002-3500-6168
[e] https://orcid.org/0000-0002-7956-1470

tion methodology tailored explicitly for cervical cancer brachytherapy planning. The contributions of this paper include:

1. Training of organ-specific neural networks with optimized multi-stage loss functions,
2. Development of an anatomy-informed preprocessing step that significantly enhances segmentation precision and efficiency,
3. An interactive visualization tool that enables deployment of the segmentation utility in a clinical setting, and
4. Rigorous clinical validation through expert user evaluations that is facilitated by inclusion of the the visualization tool within the treatment planning workflow.

The experimental results and user evaluations highlight the practical clinical utility and substantial workflow enhancements of the proposed method, resulting in time savings of 3-4 minutes per patient. Collectively, these advancements underscore the potential of targeted deep learning solutions to streamline radiation treatment planning, ultimately improving patient care outcomes in cervical cancer brachytherapy.

## 2 RELATED WORK

Automated segmentation of medical images has been widely studied, with deep learning methods becoming the state-of-the-art in recent years (Huang et al., 2024). The U-Net architecture (Ronneberger et al., 2015), with its encoder-decoder design and skip connections, has become a standard for biomedical segmentation tasks. Variants such as 3D U-Net (Çiçek et al., 2016) and U-Net with ConvLSTM (Xu et al., 2019) have been developed to incorporate spatial or temporal context, but often require more data and computational resources.

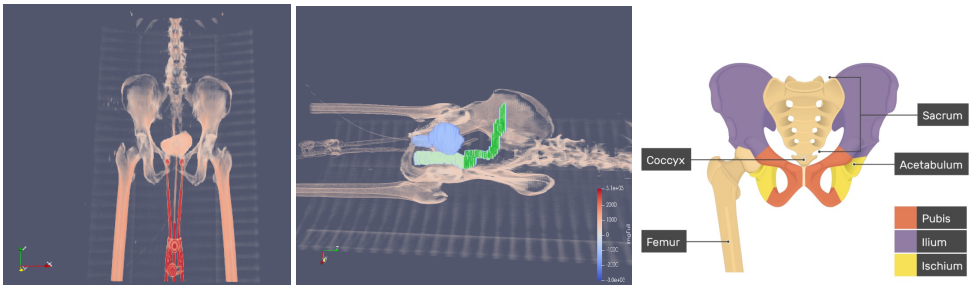
Prior studies (Bandyk et al., 2021; Perslev et al., 2019) show that multi-organ segmentation networks often under-perform on small or tubular organs due to class imbalance and structural similarity. Focal loss (Lin et al., 2017) and Dice loss (Sudre et al., 2017) have been introduced to improve performance on such difficult regions. Recent studies (Krupien et al., 2025) on the application of deep learning models for bladder and rectum segmentation on CT data report that a lighter UNet++ architecture performs as well as a heavier nnU-Net (Isensee et al., 2021). Notably, the study adopted UNet++ (Zhou et al., 2019) in routine practice due to its lower hardware demands, facilitating use on standard clinical workstations. A recent prospective study (Kraus et al., 2024) demonstrated that auto-contouring of OARs can streamline HDR brachytherapy planning, thereby reducing plan approval times by approximately 19% compared to manual contouring workflows. This indicates that automation not only maintains acceptable accuracy but also saves time in real clinical settings.

Our approach builds on 2D U-Nets by using organ-specific networks and a multi-stage training schedule to improve segmentation accuracy for the bladder, rectum, and sigmoid colon. In comparison to the nnU-Net and UNet++ architectures, our method produces comparable accuracy, especially in organ boundaries, thanks to our proposed staged training loss methodology, while using the simpler and fewer parameter 2D U-Net architecture during inference.

## 3 DATASET

This study utilizes a dataset comprising 150 anonymized pelvic CT volumes, acquired from a single clinical source and annotated for three target organs: the bladder, rectum, and sigmoid colon (see Figure 1). Each CT scan comprises between 80 and 130 2D slices. The dataset was stratified into training (100 cases), validation (20 cases), and test (30 cases) sets to support model development and evaluation. The dataset was acquired over different sessions and annotated by multiple clinicians, thus introducing variability in annotation and imaging conditions. Slices lacking corresponding organ annotations were excluded from all stages to maintain label integrity and reduce class imbalance. Our dataset is comparable in size to many prior brachytherapy segmentation studies (Mohammadi et al., 2021; Lempart et al., 2022; Huang et al., 2024).

The preprocessing pipeline focused on narrowing the input to anatomically relevant regions. Two complementary techniques were employed to identify the region of interest (ROI): a geometric method that used positional heuristics based on pelvic bone landmarks, and a lightweight U-Net model trained to identify probable organ locations in sagittal and coronal views. The outputs of both methods were spatially averaged to produce consistent, organ-centered bounding boxes. Extracted slices were then cropped to the ROI and resized to $128 \times 128$ pixels, creating a standardized input format for all segmentation models. This preprocessing stage not only accelerates training convergence but also improves segmentation quality by reducing the presence of out-of-scope anatomical structures.




(a) Bone Structure (b) Organs of interest (c) Anatomical landmarks

Figure 1: Volume rendering of the organs of interest and the bone structure that provides landmarks.

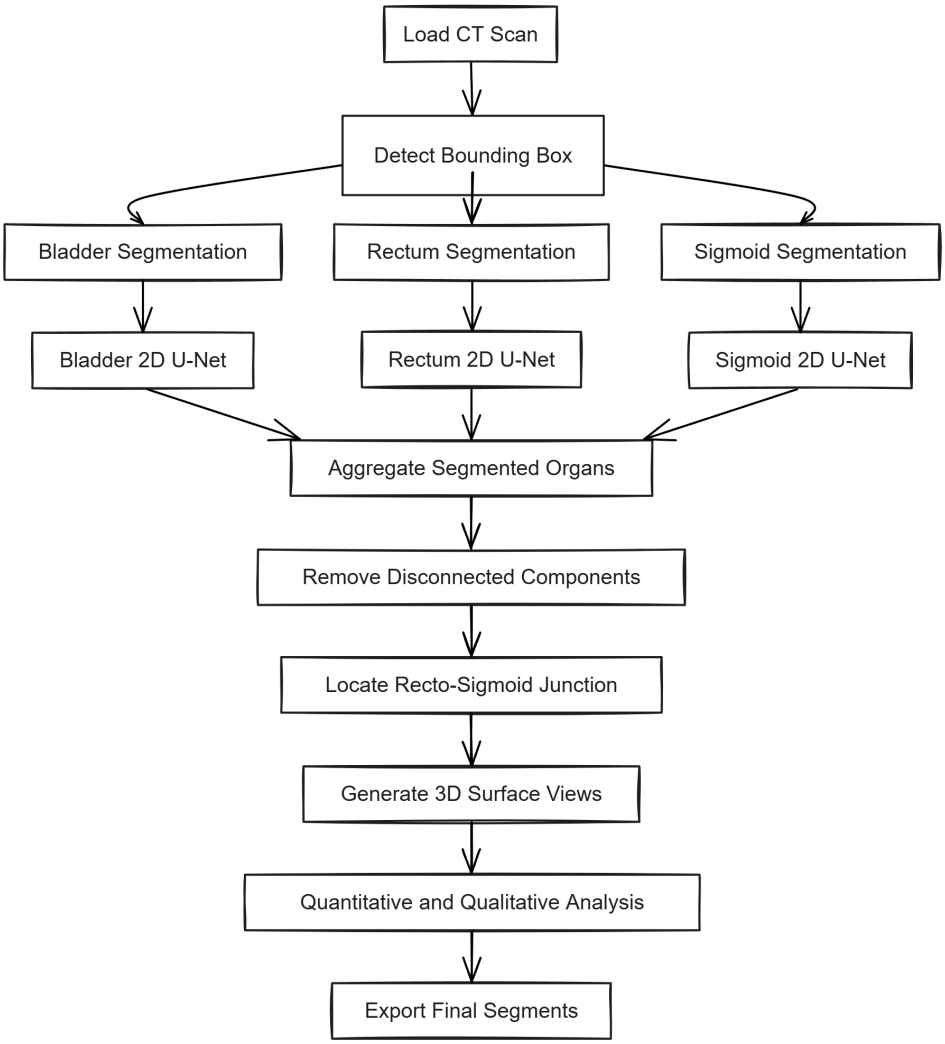



Figure 2: Overall framework of our auto-segmentation tool for cervical cancer brachytherapy planning.

## 4 METHODOLOGY

The entire procedure encompasses the reception of an input CT image, preprocessing of the data, organ prediction, postprocessing of the predicted organs to generate the final image, and generating an output format compatible with the clinical workflow. The step-by-step description of the process is discussed in detail below. Figure 2 presents a schematic of the workflow.

## 4.1 Preprocess

The preprocessing of the input data consists of four steps: reading the input file, obtaining a bounding box via a geometric approach, obtaining a bounding box via a machine learning (ML) approach, and finally rectifying the box with a user adjustment if needed.

### 4.1.1 Reading input file

The input CT volume (in DICOM `.dcm` format) is first converted to the VTK format for visualization and processing (Anderson et al., 2021). During this conversion, an anisotropic scaling is applied to adjust voxel dimensions: we scale the data by a factor of 3 along the $z$-axis (transverse plane) while keeping a scale of 1 along the other axes. This ensures that the data has uniform aspect ratio for subsequent processing and visualization.

### 4.1.2 Computing bounding box (geometric)

As discussed earlier, our approach uses a coarse bounding box to isolate the organs of interest. The geometric method leverages anatomic structures of the bones as fixed reference points to approximate the location of the organs of interest.

**Bone using isovalue**: We extract the bony structures by thresholding the CT image at an intensity (HU) range of 150 to 2000. All voxels within this "bone" range are set to 1 (foreground) while others are set to 0. This also captures any densely radio-opaque devices (e.g., brachytherapy applicator needles) present in the scan. However, the algorithm is not designed to handle large foreign objects other than bones; the presence of such objects could introduce confusion since their positions are not fixed such as the skeletal anatomy.

**Components**: For every slice along the z-axis, we identify all sizable disconnected components that exhibit sufficient density. These heuristic criteria ensure that we focus on anatomical structures as opposed to elements such as the scanning table and surgical instruments. The criteria employed to determine the inclusion of a component is as follows: $20 < h, w < 140$ and $\frac{totalArea}{h*w} > 0.15$ where $h, w$ is the height and width of the bounding box of the component, and $totalArea$ is the area of all components.

**Starting axial extent** ($z_{\min}$): When the number of bones in a slice is more than 4, we can consider taking slices (transverse plane) as we will have six disconnected bone segments of the femur and os-coxa. Figure 3 illustrates the identification of the starting slice.

**Sagittal extents** ($x_{\min}, x_{\max}$) **and coronal extents** ($y_{\min}, y_{\max}$)**.** : Next, we determine the ROI extents in the other two planes (sagittal and coronal ). We use the os coxae, bladder, and coccyx as anatomical landmarks for this purpose (Figure 3). Notably, although the bladder is not a bone, its thick muscular wall often produces intensities within the bone isovalue range, causing it to appear in the thresholded image.

In the sagittal view (X-axis), the vertical span of the bladder provides the approximate anterior-posterior range of the ROI; in the coronal view (Y-axis), the ROI extends from the top of the bladder down to the top of the coccyx. To identify the bladder in the thresholded volume, we first locate the left and right os coxae (pelvic hip bones) as the components to the extreme left and extreme right, respectively. We then locate the connected component whose upper region lies above the midpoint between these two bones (excluding the bones' components themselves); this component is imputed as the bladder. Among all slices containing the bladder, the slice with the highest superior boundary (the topmost bladder slice) is denoted as LB and marks the superior extent of the bladder. The coccyx is identified on this LB slice as a small, low-lying bone component inferior to the bladder and near the midline. For this detection, we temporarily lower the minimum component size threshold to 3 pixels in height/width, since the coccyx is much smaller than other bones. The superior surface of the coccyx on the LB slice provides the inferior (caudal) bound ($y_{\max}$) for the ROI in the coronal plane.

**Ending axial extent** ($z_{\max}$): Moving onward from the starting axial extent ($z_{\min}$), we mark the slice at which the number of detected bone components drops below two. At that point, the pelvic bones have largely disappeared (only the lumbar spine remains visible), indicating that we have passed the region containing the organs of interest. We consider that slice to be the ending (top) slice of the ROI.

### 4.1.3 Computing bounding box (ML)

We create our dataset using the labeled data for the ML approach. In the sagittal plane, we mark rectangles bounding the organs, obtain the ranges from the labels provided and convert the layers to 128 $\times$ 128 image. We used U-Net to predict the boundary from layers and resized the image to its original size, as shown in Figure 4. The predicted boundary is not

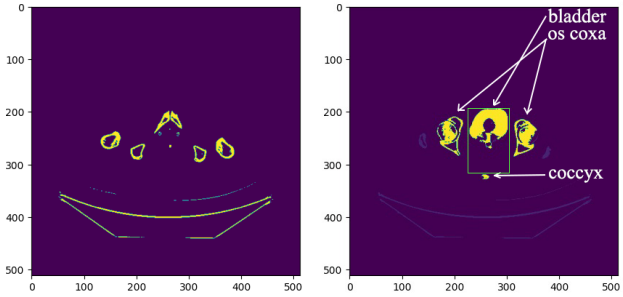



Figure 3: Organs that specify the (left) starting axial extent and the (right) sagittal and coronal extents.

always perfect. We compute the rectangle using an implementation of the *Ramer-Douglas-Peucker* algorithm (Ramer, 1972; Douglas and Peucker, 1973; Visvalingam and Whyatt, 1990) from OpenCV (*approxPolyDP*).

Next, we get the coordinates of the Y-axis and Z-axis from the rectangles as we are performing the process on X-axis (saggital plane). We consider a plane with a bounding box if we get four points from the *Douglas-Peucker* algorithm. We can get y-min, y-max, z-min and z-max from the four points for the particular slice. If we store all valid Y-Z ranges, we will have many outliers, we can remove the outliers using the median absolute distance to the median. After removing outliers we compute the average of the remaining points.

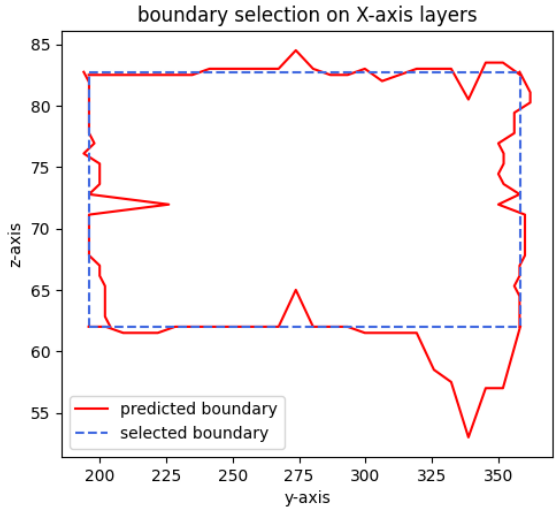



Figure 4: Computing bounding box using predicted boundary.

We use the same approach for predicting X-axis or sagittal plane range by taking slices in the coronal plane or Y-axis. The mean absolute error for the ML and the geometry integrated approaches are as shown in Table 1.

## 4.2 Training of organ-specific models

The core segmentation models are trained in an organ-specific manner. As discussed in the earlier sections, we constrain the training data to each organ's own ROI (in the transverse plane for rectum and sigmoid, focusing as closely as possible on each organ to improve predictions). For each organ, we compute its bounding box from the ground truth labels (using a similar approach as the preprocessing above) and extract only the slices within that range for training data.

We train four separate 2D U-Net models for segmenting the target structures, each focused on a specific organ or orientation: a *bladder* model (trained on all transverse slices containing the bladder, within its ROI), a *rectum* model (trained on the transverse slices spanning the rectum from its start to end), a *sigmoid colon* model (trained on the transverse slices covering the sigmoid colon's extent), and a combined *rectum+sigmoid* model (trained on sagittal-plane slices capturing the entire continuous rectum-sigmoid structure). All training images and corresponding label masks were cropped and resized to $128 \times 128$, and all models use a U-Net architecture operating on this resolution.

To further improve performance, we employed a multi-stage training schedule with progressively specialized loss functions. Training was carried out in four sequential stages: first using binary cross-entropy (BCE) on a dataset of 4000 sample slices for 37 epochs, then BCE on 1000 samples for 100 epochs, followed by binary focal cross-entropy on 2000 samples for 50 epochs, and finally using a Dice loss on 400 samples for 10 epochs. This staged approach starts with a broad learning phase (many samples, simpler loss) and gradually fine-tunes the model with more focused losses on curated subsets, which we found helped refine the segmentation boundaries.

After the initial BCE-only training stages, the predicted organ contours tended to have somewhat blurred boundaries, making it sometimes unclear which parts were inside the organ region of inter-

Table 1: Mean absolute error for the ML and geometry integrated approaches.

| MAE | xmin | xmax | ymin | ymax | zmin | zmax |
|---|---|---|---|---|---|---|
| ML | 7.32 | 4.97 | 4.23 | 2.83 | 2.36 | 1.55 |
| ML+Geom | 3.61 | 3.17 | 4.23 | 2.83 | 2.36 | 1.55 |

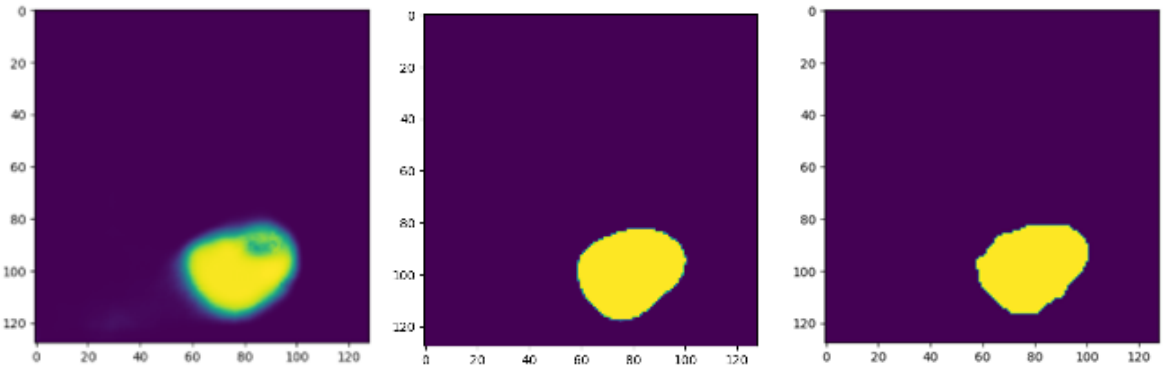


Figure 5: Prediction using BCE-only training (left), prediction after final stage with Dice loss (centre), and ground truth (right).

est (Figure 5). However, by the final stage of training (with Dice loss) the predictions sharpened significantly, yielding contours much closer to the ground truth.

## 4.3 Prediction of the organs

During inference, we first obtain the ROI bounding box for each organ using the preprocessing pipeline (with manual fine-tuning of the box if necessary). All three organs share the same in-slice cropping dimensions (height and width) in the transverse plane. However, we allow separate starting and ending slice indices for the rectum and sigmoid in the axial (transverse) direction to help focus each prediction. Notably, if the rectum and sigmoid are given identical axial slice ranges (i.e., $z_{\min}$ and $z_{\max}$ are the same for both), the model predictions for rectum and sigmoid will overlap throughout that region—since the two structures are contiguous, a CNN that does not inherently distinguish them will label the entire shared region for both organs. In practice, providing distinct axial slice bounds for rectum and sigmoid (based on anatomy, or by splitting at the recto-sigmoid junction in a postprocessing step) ensures that the model focuses on the appropriate portion of each organ.

For generating the final segmentation, we run all four trained models on their respective target regions (one for bladder, one for rectum, one for sigmoid, and one for rectum+sigmoid sagittal). In our experiments we utilized all these models to validate the approach, though in a practical deployment one could use only the sagittal-plane rectum+sigmoid model to directly produce a combined mask for those two organs.

## 4.4 Postprocessing

For further improvement of the result, we provide two geometric algorithms: removing disconnected components, and prediction of the recto-sigmoid junction.

### 4.4.1 Removing disconnected components

The raw predictions for each organ can sometimes contain small spurious regions or isolated blobs that are not part of the main organ structure. We eliminated these by performing a 3D connected-component analysis on each predicted organ mask and filtering out small disconnected components. In our implementation, all connected components for each organ are identified and their voxel counts computed. A clinician can then interactively review and discard components below a certain size threshold (via our visualization software), or this step can be automated by removing all components smaller than a fixed percentage of the largest component. After filtering, we reconstruct the final organ volume by combining only the retained components. This step effectively removes prediction noise while preserving the structure of the organ.

### 4.4.2 Prediction of a Recto-Sigmoid junction slice

As discussed, the segmentation model does not inherently delineate the junction between the rectum and sigmoid colon, often producing a continuous con-

nected mask for these two organs. We therefore apply a postprocessing algorithm to estimate the recto-sigmoid junction as a slice and split the segmentation accordingly. During our development, we considered three different anatomical formations to estimate a recto-sigmoid junction slice: (1) using the distance of the recto-sigmoid prediction from the sacrum bone, (2) using the cross-sectional radius of the recto-sigmoid prediction, and (3) using the angle of the centerline of the recto-sigmoid prediction relative to the axial plane. Based on discussions with experts and early experiments, we narrowed down on the third method. (Figure 6 illustrates these anatomical considerations for ).

In our chosen method, we first compute the center of mass of the segmented rectum-sigmoid region on each axial slice. Connecting these center points, we form an approximate centerline curve of the organ. We then calculate the angle that this centerline makes with respect to the axial plane (essentially, the tilt of the organ's trajectory between consecutive slices). The recto-sigmoid junction manifests as a noticeable change (drop) in this angle as we move along the slices. To reduce noise, we apply a Savitzky–Golay filter (Savitzky and Golay, 1964) to smooth the angle-vs-slice curve. We provide 5 options to the user on this curve to select the junction slice (as detailed in Algorithm 1 above): for example, detecting the largest instantaneous angle drop, the largest drop in the smoothed curve, the largest change in the discrete derivative of the smoothed curve, the first major bend (local maximum of the slope curve), and a fixed angle threshold (e.g., first slice where the angle falls below $76^\circ$, based on empirical observation from ground truth data).

In our experiments, at least one of these options typically produced junction estimates within a few slices of the true junction. We found that the combination of all five approaches yielded a mean absolute error of about 1.95 slices in locating the junction. To decide on a final junction, we evaluated a small set of sample cases using all five criteria and selected the criterion that performed best (closest to the expert-defined junction) for each case. The frequency with which each method was selected as best is given in Table 2. In our sample of cases, the centerline angle-based criterion (method 3) most often provided the most accurate junction slice.

# 5 EVALUATION

For evaluating the different models, we used the 30 test CT images and used DICE score to evaluate.

**Input:** Segmented rectum+sigmoid volume $I$
**Output:** $\{z_1, z_2, z_3, z_4, z_5\}$ (junction slice candidates)
$I_c \leftarrow removeSmallComponents(I)$
$Slopes \leftarrow slopeOfMidpoints(I_c)$ {centerline angle per slice}
$Slopes_{smooth} \leftarrow savgolFilter(Slopes)$
$dSlope \leftarrow discreteSlope(Slopes_{smooth})$
$dSlope_{smooth} \leftarrow savgolFilter(dSlope)$
$z_1 \leftarrow maxDrop(Slopes)$
$z_2 \leftarrow maxDrop(Slopes_{smooth})$
$z_3 \leftarrow maxDrop(dSlope_{smooth})$
$z_4 \leftarrow maxima(dSlope_{smooth})[0]$ {first local maximum}
$z_5 \leftarrow indexOf(Slopes_{smooth} < 76^\circ)$
**return** $\{z_1, z_2, z_3, z_4, z_5\}$

Algorithm 1: Recto-Sigmoid Junction Slice Predictions

| Approach | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Frequency selected | 0.16 | 0.16 | 0.30 | 0.11 | 0.25 |

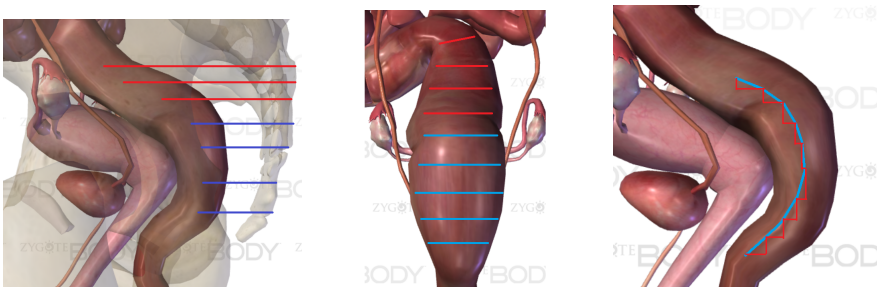
Table 2: Frequency at which an approach (1–5) is selected as most accurate for recto-sigmoid junction localization in sample tests. The angle-based approach (3) was most frequently identified as optimal.

Table 3 summarizes a comparative analysis of different training methods. Specifically, we study the effect of including vs. removing unlabeled slices on Binary Cross Entropy (BCE). We also observe an improvement upon including dice loss. In all methods, sharpness of the boundary is enhanced by setting values 0.4 and above to 1. Note that DICE score ranges between 0 and 1, where higher scores corresponds to better results. In Table 4, we compare the individual training stages for different organs. Stage 1 and stage 2 uses BinaryCrossentropy, stage 3 uses BinaryFocalCrossentropy, stage 4 uses relaxed Dice Loss. We also compute the Dice scores after postprocessing for comparison.

Figure 7 shows some 2D output of interest for the bladder, rectum, and sigmoid. Figure 8 shows results in 3D, comparing the ground truth vs. the predicted segmentation.

# 6 IMPLEMENTATION AND VISUALIZATION

We have developed a user-friendly software interface to deploy the segmentation tool in a clinical setting. The application is implemented in Python and C++, and it integrates with the 3D Slicer (Kikinis et al.,



(a) distance from sacrum (b) cross-sectional radius (c) centerline angle

Figure 6: Anatomical formations considered for developing an algorithm to determine recto-sigmoid junction slice.

| **Model / Training Strategy** | **Loss** | **Accuracy** | **Dice** |
|---|---|---|---|
| Single U-Net (no organ-specific ROI), all slices | Cat. CrossEntropy | 0.9512 | 0.7900 |
| Organ-specific U-Nets (no unlabeled slices removed) | Binary CrossEntropy | 0.9724 | 0.8900 |
| Organ-specific U-Nets (unlabeled slices removed) | Binary CrossEntropy | 0.9919 | 0.9882 |
| Organ-specific U-Nets (unlabeled removed, multi-stage losses) | BCE + Focal + Dice | 0.9937 | 0.9898 |

Table 3: Comparing training approaches. Removal of slices without organ labels and using specialized per-organ models greatly improved segmentation performance. Multi-stage training (with focal and Dice losses) provided a further gain in Dice.

| **Organ (plane)** | Stage1 | Stage2 | Stage3 | Stage4 | Postproc |
|---|---|---|---|---|---|
| Bladder (axial) | 0.9826 | 0.9929 | 0.9822 | 0.9813 | 0.9882 |
| Rectum (axial) | 0.9826 | 0.9919 | 0.9897 | 0.9886 | 0.9891 |
| Sigmoid (axial) | 0.9904 | 0.9968 | 0.9927 | 0.9899 | 0.9886 |
| Rectum+Sigmoid (sagittal) | 0.9882 | 0.9932 | 0.9911 | 0.9897 | 0.9915 |

Table 4: Dice scores across training stages for each organ. Stage1/Stage2: Binary CrossEntropy; Stage3: Focal loss; Stage4: Dice loss. Postproc indicates after applying the geometric postprocessing steps.

2014) platform for visualization. Key features of the interface include real-time rendering of the segmented organs as smooth 3D surfaces (Figure 9), interactive adjustment of the ROI bounding box, and options to include or exclude specific segmented components.

## 6.1 Software workflow

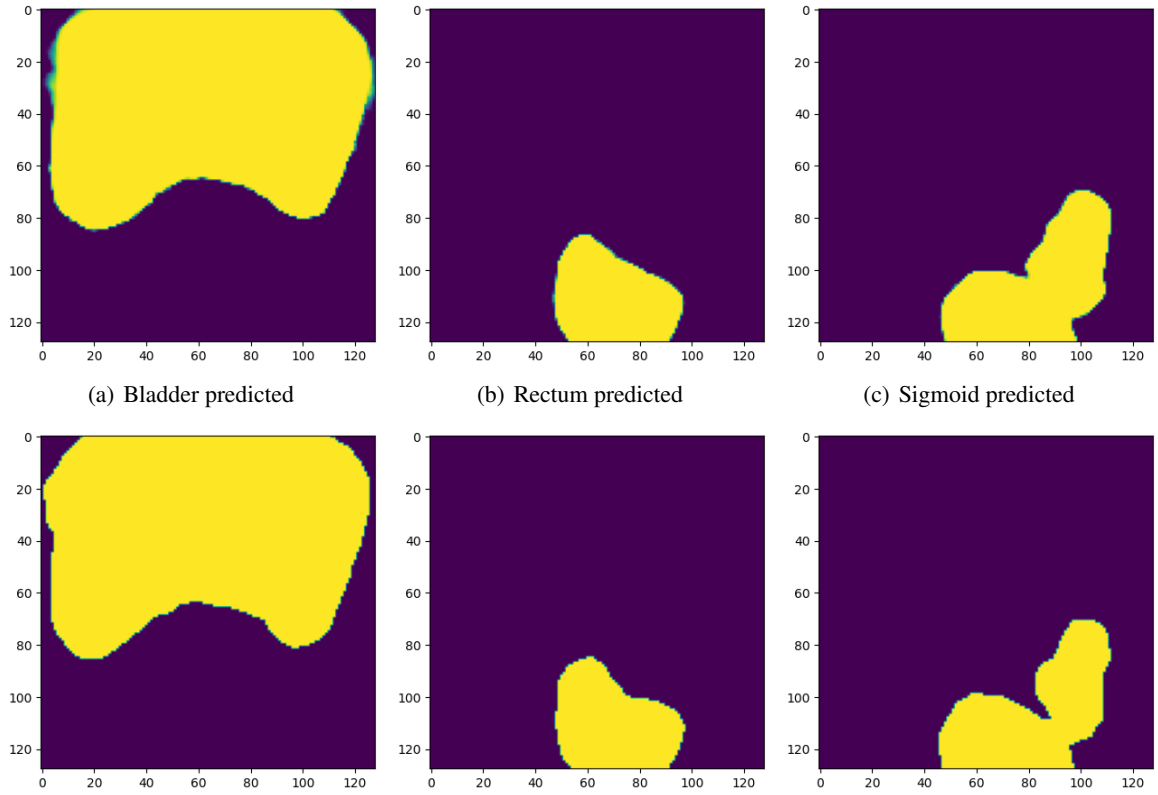
Upon loading a patient's CT scan (DICOM format), the software automatically performs the preprocessing steps to compute the ROI. Next, it applies each organ-specific model to generate the bladder, rectum, and sigmoid segmentation masks. The results are displayed to the user with support to inspect and adjust the results as necessary. The user may leverage the disconnected component analysis to identify and remove small stray segments using a simple click interface. The segmented organs can be exported as standard DICOM RT-Structure files for use in treatment planning systems.

## 6.2 Surface visualization

The segmented organs are visualized as smooth triangulated surfaces as compared to raw voxel grids, which makes it easier for clinicians to interpret the results. This surface is computed using an accelerated marching cubes algorithm (Lorensen and Cline, 1987) followed by a surface smoothing operation, which enhances the clarity of organ boundaries and their spatial relationships. The surface rendering is updated live in response to user adjustments.

## 6.3 Computation and runtime

The segmentation models were run on a system with an NVIDIA GTX 1060 GPU with 6GB VRAM and 32GB RAM. Typically, the segmentation for all three organs is obtained in 3–4 seconds per patient scan. The efficient runtime is achieved due to the relatively small ROI computed in the pre-processing step and the efficiency of the 2D U-Net models. Even on a modest workstation without a powerful GPU, the seg-



(a) Bladder predicted (b) Rectum predicted (c) Sigmoid predicted

(d) Bladder true (e) Rectum true (f) Sigmoid true

Figure 7: Predicted vs. ground truth for 2D slices.

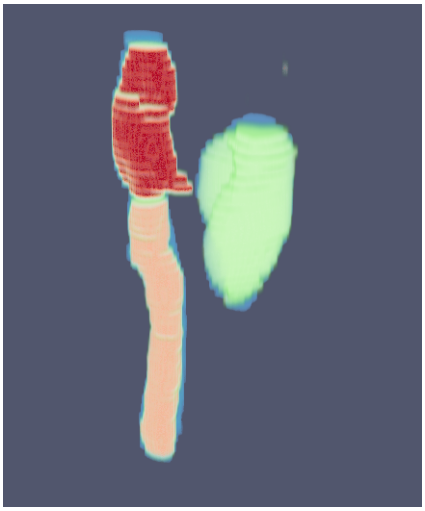


Figure 8: Predicted vs. ground truth (blue) 3D segmentation. The predicted organs are in green (bladder), light red (rectum), and dark red (sigmoid).

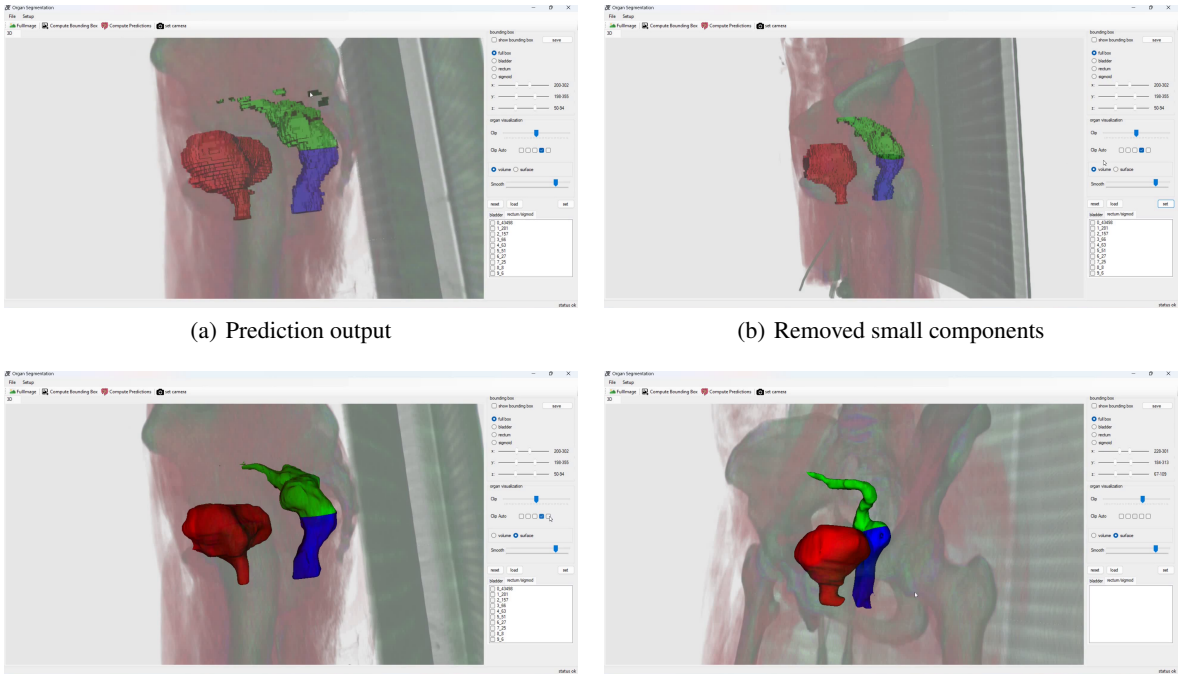
mentation is computed within tens of seconds, making it feasible for clinical use.

### 6.4 Export and integration

The software includes functionality to export the final segmentation in DICOM RT-Struct format, which can be directly imported into brachytherapy planning systems. This allows the auto-segmented contours to be used in dose calculations and treatment planning immediately. The goal is to integrate this tool as a plugin or a standalone module in the clinical workflow so that oncologists can generate organ contours at the time of planning with minimal overhead.

## 7 CLINICAL FEEDBACK

To evaluate the clinical relevance and real-world usability of our segmentation framework, we conducted a structured user study involving five clinicians—two senior consultants and three resident doctors—focused on organ contouring for cervical cancer brachytherapy planning. Each participant segmented the bladder, rectum, and sigmoid colon across a dataset of 20 CT scans, both manually and using our prototype auto-segmentation tool. Figure 10



(a) Prediction output

(b) Removed small components

(c) Surface visualization-1

(d) Surface visualization-2

Figure 9: Interactive visualization of the 3D organ segmentation supports user feedback and proofreading.

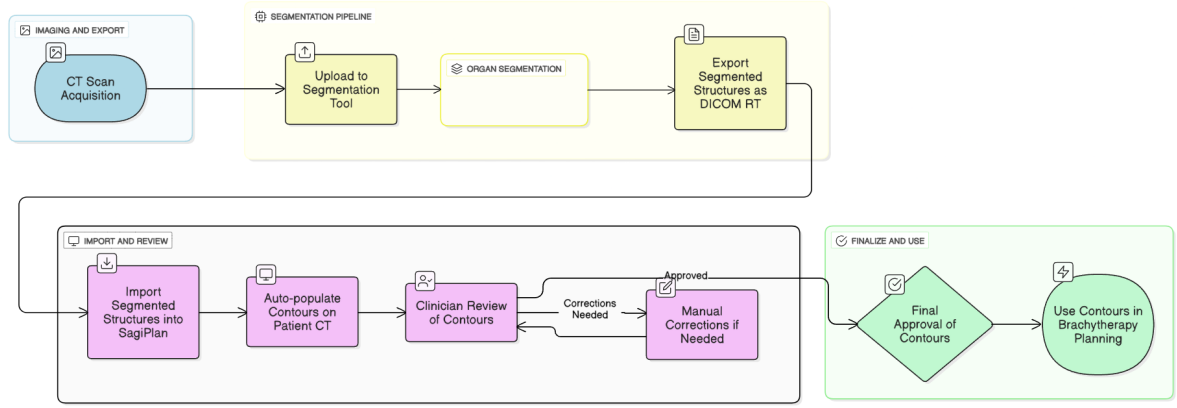



Figure 10: Clinical workflow for evaluating the proposed auto-segmentation tool

shows the workflow given to all participants of the user study. On average, manual delineation by the doctors took about 7 minutes and 51 seconds per case, (ranging between 4 minutes and 14 minutes and 30 seconds) with junior doctors generally taking longer than seniors. In contrast, our software produced contours and were integrated into a treatment planning software to produce a treatment plan in 3 minutes and 38 seconds on average per case (ranging between 3 minutes to 5 minutes and 30 seconds). This translates to approximately a two-fold increase in speed, saving around 3–4 minutes per patient compared to manual work. A key contributor to this time reduction is the integrated interactive visualization interface, which allows clinicians to quickly review and correct specific regions directly within the tool. This ensures that any segmentation inaccuracies can be rapidly resolved before export, without requiring a complete manual re-contouring.

Overall, these findings indicate that the auto-segmentation software can generate usable organ contours in roughly half the time of manual segmentation. This improvement in efficiency—achieved without sacrificing contour accuracy or consistency—could

free up several minutes per patient for clinicians and help standardize the contouring process in practice.

One of the senior consultants provided qualitative feedback stating: *"This software significantly speeds up the contouring process compared to manual methods, which is highly beneficial in improving workflow efficiency. However, it currently generates the same volume for both the rectum and sigmoid, which is not accurate. In manual contouring, the volumes of these two structures typically differ due to their distinct anatomical shapes and positions on CT images. The software may require some improvements to address this issue. The volume similarity might be related to how the algorithm processes and differentiates the structures. Additionally, the output may vary depending on the computer's processing speed, so optimizing the software for consistent and accurate performance across different systems would be advantageous."* Reflecting on the above observation, we note that the geometric postprocessing incorporated into the method does ensure accurate location of the recto-sigmoid junction. In practice, applying this step resolves the volume similarity, ensuring that the final contours align with expected clinical anatomy. This step remains a major contributor to the time required for contouring.

# 8 CONCLUSIONS AND FUTURE WORK

We presented an effective and clinically validated tool for fast and accurate segmentation of pelvic organs in cervical cancer brachytherapy. Our organ-specific approach achieves high accuracy while significantly reducing segmentation time. Specifically, the use of tailored 2D U-Net models for each organ enabled rapid predictions (3–4 seconds per scan) suitable for clinical workflows. By focusing the segmentation on anatomy-specific regions through preprocessing and ensuring that each model operates within the region of interest, we improved both accuracy and efficiency. We also addressed the challenge of continuous anatomical structures (like the rectum and sigmoid colon) by applying geometric postprocessing to successfully separate and delineate them. A preliminary user study with radiation oncologists confirmed the clinical acceptability of the segmentation. The clinical evaluation demonstrated substantial time savings (approximately 50% reduction in contouring time) and improved consistency of organ contours among different clinicians. Moreover, the system runs on modest hardware, facilitating easy integration into existing clinical setups.

For future work, we plan to further enhance the ability of the model to differentiate the rectum and sigmoid, possibly by integrating the junction detection directly into the learning process or by training specialized models for different portions of the anatomy. We also aim to expand and validate our approach on larger datasets from multiple institutions to ensure its robustness and generalizability. Finally, incorporating additional organs or extending the framework to other treatment sites could broaden the applicability of the tool.

Overall, our results highlight that organ-specific deep learning models, combined with intelligent preprocessing and postprocessing, can streamline the treatment planning workflow in brachytherapy, providing faster and more consistent organ delineations to ultimately improve patient care.

# ACKNOWLEDGEMENTS

This work was partially supported by scholarships from the MoE Govt. of India and a CSR grant from Ittiam Systems for the Equitable AI Lab at CSA, IISc Bangalore. Part of this work constitutes the Masters project work of Nirmalya Gayen and Aditya Kumar at IISc Bangalore. We thank all doctors from the Oncology team at Ramaiah Medical College who participated in the user study and Dr. Lithika Lavanya M. and Ms. Revati for coordinating the clinical validation exercise and providing valuable inputs.

# REFERENCES

Anderson, B. M., Wahid, K. A., and Brock, K. K. (2021). Simple python module for conversions between dicom images and radiation therapy structures, masks, and prediction arrays. *Practical Radiation Oncology*, 11:226–229.

Bandyk, M. G., Gopireddy, D. R., Lall, C., Balaji, K., and Dolz, J. (2021). Mri and ct bladder segmentation from classical to deep learning based approaches: Current limitations and lessons. *Computers in Biology and Medicine*, 134:104472.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer.

Derks, K., Steenhuijsen, J. L., van den Berg, H. A., Houterman, S., Cnossen, J., van Haaren, P., and Jaeger, K. D. (2018). Impact of brachytherapy technique (2d versus 3d) on outcome following radiotherapy of cervi-

cal cancer. *Journal of Contemporary Brachytherapy*, 10:17–25.

Douglas, D. H. and Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 10(2):112–122.

Huang, L., Miron, A., Hone, K., and Li, Y. (2024). Segmenting medical images: from unet to res-unet and nnunet. In *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 483–489. IEEE.

Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., and Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2):203–211.

Kikinis, R., Pieper, S. D., and Vosburgh, K. G. (2014). 3d slicer: A platform for subject-specific image analysis, visualization, and clinical support. In Jolesz, F. A., editor, *Intraoperative Imaging and Image-Guided Therapy*, pages 277–289. Springer New York, New York, NY.

Kraus, A. C., Iqbal, Z., Cardan, R. A., Popple, R. A., Stanley, D. N., Shen, S., Pogue, J. A., Marcrom, S., and Cardenas, C. E. (2024). Prospective evaluation of automated contouring for ct-based brachytherapy for gynecologic malignancies. *Advances in Radiation Oncology*, 9(4):101417.

Krupien, A. J., Abdulkadir, Y., Luximon, D. C., John, C., Dong, H., Pham, J., O'Connell, D., Neylon, J., and Lamb, J. M. (2025). Open-source deep-learning models for segmentation of normal structures for prostatic and gynecological high-dose-rate brachytherapy: Comparison of architectures. *Journal of Applied Clinical Medical Physics*, 26(1):e70089.

Lempart, M., Nilsson, M. P., Scherman, J., Gustafsson, C. J., Nilsson, M., Alkner, S., Engleson, J., Adrian, G., Munck af Rosenschöld, P., and Olsson, L. E. (2022). Pelvic u-net: multi-label semantic segmentation of pelvic organs at risk for radiation therapy anal cancer patients using a deeply supervised shuffle attention convolutional neural network. *Radiation oncology*, 17(1):114.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988. IEEE.

Lorensen, W. E. and Cline, H. E. (1987). Marching cubes: A high resolution 3d surface construction algorithm. In *SIGGRAPH '87*, pages 163–169. ACM.

Mahantshetty, U., Gudi, S., Singh, R., Sasidharan, A., Sastri, S. C., Gurram, L., Sharma, D., Ganeshrajah, S., MG, J., Badakh, D., Basu, A., James, F., Swamidas, J., Kuppuswamy, T., and Bhalavat, R. (2019). Indian brachytherapy society guidelines for radiotherapeutic management of cervical cancer with special emphasis on high-dose-rate brachytherapy. *Journal of Contemporary Brachytherapy*, 11:293–306.

Mohammadi, R., Shokatian, I., Salehi, M., Arabi, H., Shiri, I., and Zaidi, H. (2021). Deep learning-based auto-segmentation of organs at risk in high-dose rate brachytherapy of cervical cancer. *Radiotherapy and Oncology*, 159:231–240.

Perslev, M., Dam, E. B., Pai, A., and Igel, C. (2019). One network to segment them all: A general, lightweight system for accurate 3d medical image segmentation. In Shen, D., Liu, T., Peters, T. M., Staib, L. H., Essert, C., Zhou, S., Yap, P.-T., and Khan, A., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, pages 30–38. Springer International Publishing.

Ramer, U. (1972). An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, 1(3):244–256.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.

Savitzky, A. and Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36:1627–1639.

Skowronek, J. (2017). Current status of brachytherapy in cancer treatment – short overview. *Journal of Contemporary Brachytherapy*, 9:581–589.

Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., and Jorge Cardoso, M. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *International Workshop on Deep Learning in Medical Image Analysis*, pages 240–248. Springer.

Visvalingam, M. and Whyatt, J. D. (1990). The Douglas-Peucker algorithm for line simplification: Re-evaluation through visualization. *Computer Graphics Forum*, 9:213–225.

Xu, F., Ma, H., Sun, J., Wu, R., Liu, X., and Kong, Y. (2019). Lstm multi-modal unet for brain tumor segmentation. In *4th IEEE International Conference on Image, Vision and Computing (ICIVC)*, pages 236–240. IEEE.

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., and Liang, J. (2019). Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 39(6):1856–1867.