Original Article

Ensemble Learning for Three-dimensional Medical Image Segmentation of Organ at Risk in Brachytherapy Using Double U-Net, Bi-directional ConvLSTM U-Net, and Transformer Network

Soniya Pal[1,2], Raj Pal Singh[1], Anuj Kumar[3]

[1]Department of Physics, GLA University, Mathura, [3]Department of Radiotherapy, LLRM Medical College, Meerut, Uttar Pradesh, [2]Batra Hospital and Medical Research Center, New Delhi, India

Abstract

Aim: This article presents a novel approach to automate the segmentation of organ at risk (OAR) for high-dose-rate brachytherapy patients using three deep learning models combined with ensemble learning techniques. It aims to improve the accuracy and efficiency of segmentation. **Materials and Methods:** The dataset comprised computed tomography (CT) scans of 60 patients obtained from our own institutional image bank and 10 patients from the other institute, all in Digital Imaging and Communications in Medicine format. Experienced radiation oncologists manually segmented four OARs for each scan. Each scan was preprocessed and three models, Double U-Net (DUN), Bi-directional ConvLSTM U-Net (BCUN), and Transformer Networks (TN), were trained on reduced CT scans (240 × 240 × 128) due to memory limitations. Ensemble learning techniques were employed to enhance accuracy and segmentation metrics. Testing and validation were conducted on 12 patients from our institute (OID) and 10 patients from another institute (DID). **Results:** For DID test dataset, using the ensemble learning technique combining Transformer Network (TN) and BCUN, i.e., TN + BCUN, the average Dice similarity coefficient (DSC) ranged from 0.992 to 0.998, and for DUN and BCUN (DUN + BCUN) combination, the average DSC ranged from 0.990 to 0.993, which reflecting high segmentation accuracy. The 95% Hausdorff distance (HD) ranged from 0.9 to 1.2 mm for TN + BCUN and 1.1 to 1.4 mm for DUN + BCUN, demonstrating precise segmentation boundaries. **Conclusion:** The proposed method leverages the strengths of each network architecture. The DUN setup excels in sequential processing, the BCUN captures spatiotemporal dependencies, and transformer networks provide a robust understanding of global context. This combination enables efficient and accurate segmentation, surpassing human expert performance in both time and accuracy.

Keywords: Brachytherapy, image segmentation, LSTM U-Net, medical image segmentation, organ at risk, transformer, U-Net

Received on: 15-09-2024 **Review completed on:** 10-11-2024 **Accepted on:** 13-11-2024 **Published on:** 18-12-2024

Introduction

Brachytherapy is a form of internal radiation therapy where radioactive sources are placed close to or inside the tumor. This technique is highly effective for treating various cancers, including those of the cervix, prostate, breast, and skin. It offers several advantages over external beam radiation therapy such as high precision, localized treatment, shorter treatment duration, and reduced side effects.[1,2] The success of brachytherapy hinges on the precise delivery of radiation doses, which necessitates meticulous treatment planning. Accurate segmentation of organs at risk (OARs), which are normal tissues or organs sensitive to radiation and located near the treatment area, is a critical component of the planning process. Precise segmentation of these structures is essential for optimizing dose distribution, minimizing radiation-induced toxicity, improving treatment planning and verification, and enabling personalized treatment.[3]

Access this article online	
Quick Response Code:	**Website:** www.jmp.org.in
	DOI: 10.4103/jmp.jmp_160_24

Address for correspondence: Dr. Anuj Kumar, Department of Radiotherapy, LLRM Medical College, Meerut, Uttar Pradesh, India. E-mail: toaktyagi@gmail.com

How to cite this article: Pal S, Singh RP, Kumar A. Ensemble learning for three-dimensional medical image segmentation of organ at risk in brachytherapy using Double U-Net, Bi-directional ConvLSTM U-Net, and Transformer Network. J Med Phys 2024;49:574-82.

Advancements in medical imaging have revolutionized the field of medical diagnosis and treatment planning, particularly in areas such as oncology.[4] The precision in delineating critical structures directly influences the effectiveness of treatment and the minimization of radiation-induced damage to healthy tissues.[5-8] However, Manual segmentation is time-consuming and depends heavily on the expertise of radiation oncologists, often resulting in variable outcomes. Consequently, automated segmentation methods, especially those leveraging deep learning (DL), have garnered significant attention in recent years.[9-11] DL models, particularly convolutional neural networks (CNNs), have demonstrated exceptional capabilities in image analysis and segmentation tasks.[12,13] Among these, the U-Net architecture is a foundational model for medical image analysis, renowned for its ability to accurately localize structures and capture contextual information. However, continuous efforts aim to enhance its performance through advanced techniques and architectures. This study explores the integration of Double U-Net (DUN), Bi-directional Convolutional Long Short-Term Memory (BCUN) U-Net, and Transformer networks (TN) in an ensemble learning framework for three-dimensional (3D) medical image segmentation. The DUN architecture is an innovative extension of the traditional U-Net, designed to overcome the limitations of single-stage networks by employing a two-step process. The first U-Net serves as a coarse structure identifier, generating an initial prediction map. This preliminary output is then fine-tuned by the second U-Net, which corrects and sharpens the boundaries of the identified regions. This setup leverages the strengths of sequential processing, ensuring that the network captures both global context and local details more effectively. Incorporating temporal dependencies and spatial relationships is crucial for accurate 3D medical image segmentation.[14,15] The BCUN introduces recurrent neural network components into the segmentation pipeline, enabling the network to learn from sequential slices of the 3D volumes. Convolutional long short-term memory (ConvLSTM) are particularly suited for this task as they can maintain spatial hierarchies while capturing temporal dependencies. By processing the image slices bi-directionally, the network benefits from contextual information from both past and future slices, leading to more coherent and accurate segmentations. On the other hand, TN has also transformed various fields of machine learning, including natural language processing and image analysis. Transformers, known for their self-attention mechanisms, excel at capturing long-range dependencies and complex relationships within data. Their application to medical image segmentation is relatively novel but promising. Transformer network segment images by treating patches of the image as sequences, similar to words in a sentence, and applying self-attention to capture intricate patterns and features.[16,17] This study integrates transformer network into the segmentation framework, aiming to exploit their powerful feature extraction and contextual learning capabilities.

The Dice similarity coefficient (DSC) uses a complex formula that considers both the areas where the predicted and ground truth segmentations overlap and those where they do not. This formula results in a score ranging from 0 to 1, with 1 indicating a perfect overlap between the predicted and ground truth segmentations. The 95% Hausdorff distance (HD) is another sophisticated metric that measures the maximum distance between corresponding points in the predicted and ground truth segmentations, allowing for a 95% tolerance of the total distance. This metric is highly adaptable, achieved using weighting factors and threshold values that can be adjusted to optimize the metric for specific segmentation tasks.

As automated segmentation in medical image is a complex and critical task that directly impacts treatment efficacy and patient safety, this study proposes an ensemble learning framework that combines the strengths of DUN, BCUN, and TN to enhance performance. By addressing the limitations of individual models and leveraging their complementary strengths, this approach aims to provide more accurate and reliable delineation outcomes. The integration of these advanced architectures represents a significant step forward in the field of medical image analysis, with the potential to improve clinical workflows and patient outcomes in radiation therapy.

Materials and Methods

Data collection and processing

In this study, a set of 3D computed tomography (CT) images was used. The data were collected from our own institute. It included images of 60 patients' CT scans (noncontrast), with their corresponding structure sets, which were taken for the purpose of brachytherapy treatment of cancer patients. For training and testing the model, 80% and 20% of data were used, respectively, accounting for 48 patients for training and 12 patients for testing the model, respectively. All the scans were in Digital Imaging and Communications in Medicine (DICOM) format having a size of 512 × 512 × 128. The scans were done on Siemens go.SIM pro (Siemens Healthcare Private Limited) CT Simulator. The CT scan DICOM files as well as structure set were converted to NIFTI format, so that it can be easily read by the DL models. The structure set includes manually segmented OARs such as bladder, bowel, sigmoid, and rectum using the protocol provided by EMBRACE II[18] (Image guided intensity modulated External beam radiochemotherapy and MRI based adaptive BRAchytherapy in locally advanced CErvical cancer-v. 1.0.) protocol for OAR delineation. All the scans were processed for size reduction due to the limitation of system memory. The reduced size was 240 × 240 × 128 (128 slices were exported for all the patients using treatment planning system), which is the final input for the model, as shown in Figure 1. For testing the model, we had used two different types of datasets. The first testing dataset consisted of 12 patients from our institute (OID), with organs-at-risk

(OAR) manually segmented by experienced oncologists following a standardized delineation protocol. The second testing dataset, comprising 10 patients, was sourced from an external institute for model validation; scans were conducted using the Canon Aquilion Start CT simulator (Canon Medical Systems Corporation). The code was executed on Google Collaboratory Pro, utilizing Python v3.7, with 25 GB of RAM and a Tesla P100 GPU equipped with 16 GB of graphics memory.[19]

Models

Originally transformer networks were built for natural language processing. The capacity to simulate global context and long-range interdependence has made them suitable for medical image analysis. Transformers employ self-attention processes to record links between visual areas independent of spatial distance, unlike CNNs, which concentrate on local characteristics by applying filters to tiny regions. Transformers excel at interpreting the whole image, which is beneficial for segmenting huge anatomical structures or where global properties are critical. Transformers can better capture distant pixel associations than CNN-based models in medical images, notably in global context segmentation tasks. At the same time, transformers need plenty of tagged data to work correctly. Due to inadequate labeled data, medical imaging transformers may not perform well without augmentation or transfer learning. They use attention techniques that demand plenty of memory, making them computationally costly, particularly for high-resolution medical images.

The Bi-directional ConvLSTM U-Net (BCUN) combines U-Net and ConvLSTM networks. LSTMs are commonly used in sequence data because they can capture temporal dependencies and spatial correlations across time and integrate well into U-Net structures. The bi-directional component lets the model evaluate data dependencies forward and backward. BCUN captures medical image spatiotemporal connections well. Dynamic or volumetric imaging (magnetic resonance imaging scans or CT volumes) may benefit from this since image slice relationships can give contextual information. BCUN models may use structural consistency to improve segmentation accuracy in 3D or sequential medical data.

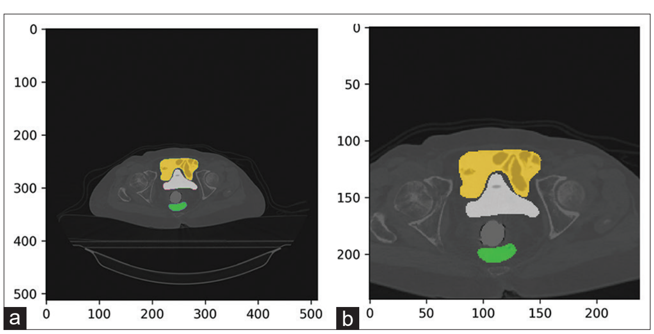

Figure 1: (a) The raw computed tomography (CT) scan image before any size reduction with their corresponding segmented contours and (b) the CT scan image after reducing the size to 240 × 240 × 128

BCUN models may overfit if trained on short datasets because of their intricacy and capacity to collect a lot of data, requiring careful regularization.

DUN, an expansion of U-Net, is commonly used for medical image segmentation. U-Net is famous for its encoder-decoder structure, which downsamples image features at several scales and upsamples the segmentation mask. DUN improves feature extraction and hierarchical learning by stacking two U-Nets. The first U-Net gathers coarse, high-level characteristics, while the second refines them to segment smaller or more sophisticated structures. It helps segment organs or tumors with complex borders. The second U-Net layer enhances segmentation detail, making DUN good at detecting tiny medical image areas.

Model used combined with ensemble learning technique

Ensemble learning combines the predictions from multiple models to improve the overall performance and robustness of the system. In this study, we employ an ensemble of DUN,[20] BCUN,[21] and TN.[22,23] Each model contributes its unique strengths: DUN provides refined segmentation maps, BCUN captures spatial-temporal dependencies, and TN leverages self-attention for capturing complex patterns. The ensemble approach aims to mitigate the weaknesses of individual models, providing a more accurate and reliable segmentation outcome.[24-26] This study aims to advance the accuracy and reliability of 3D medical image analysis by utilizing the combined strengths of multiple state-of-the-art deep learning architectures within an ensemble framework, striving to enhance automated performance.[27,28]

After training, the trained models were used to evaluate the models' performance on two different datasets. Figure 2 shows the segmentation workflow, which includes the training and testing stages. In this instance, 48 scans were input into the model to train it. Following training, the models were evaluated on two datasets that were manually segmented by experienced oncologists. The training and testing exercise was also conducted using the ensemble learning model also which includes the combination of TN and BCUN (TN + BCUN) and DUN and BCUN (DUN + BCUN).

Performance evaluation metrics

For accessing the model performance, we had used different types of image segmentation metrics such as DSC,[29,30] 95% HD (mm),[31,32] and mean boundary distance.[33] The DSC or f1 score for two given sets X and Y was expressed as given in equation (1) and shown in Figure 3.

$$DSC = 2^* \mid \text{Intersection } (X, Y) \mid / (|X| + |Y|) \quad (1)$$

where X represents ground truth and Y represents segmented mask.

The 95% HD is the maximum distance between the bounds of the expected segmentation and the ground truth. The 95% HD measure is often used in medical imaging to minimize outliers by concentrating on the 95th percentile of distances as

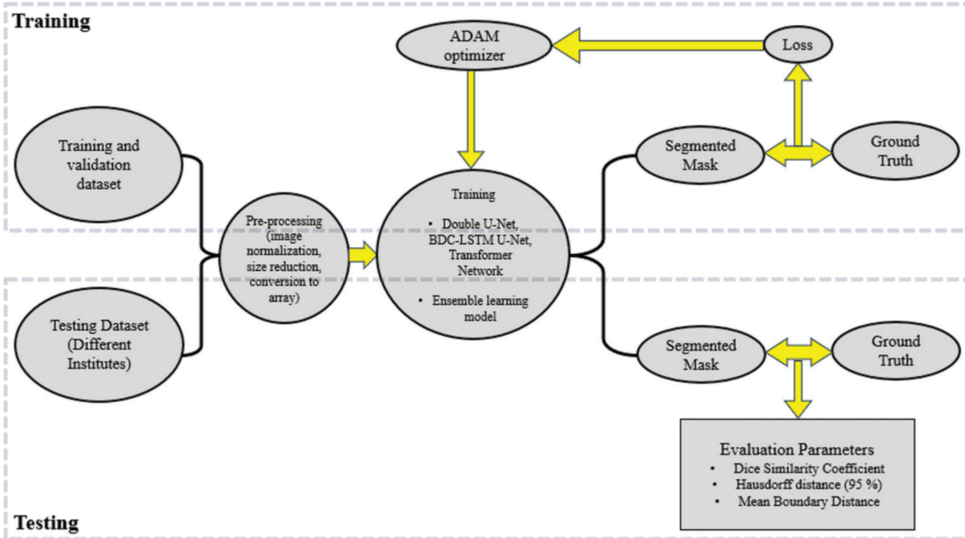

Figure 2: The workflow of segmenting multiple organs. ADAM: Adaptive moment estimation, BDC-LSTM U-Net: Bi-directional ConvLSTMU-Net

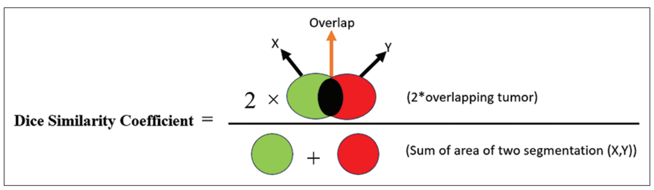

Figure 3: Dice similarity coefficient: 2 × overlap/sum of area

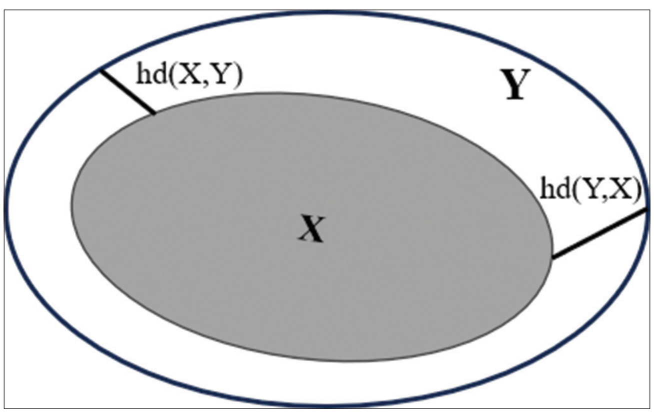

Figure 4: The 95% Hausdorff distance between two-point sets, X and Y

given in equation (2) and shown in Figure 4. In other words, the 95% HD is the greatest distance from one point in a set to the nearest point in another.

$$\text{HD (mm)} = \max(\text{hd}[X, Y], \text{hd}[Y, X]) \tag{2}$$

Mean boundary distance is crucial in medical image segmentation, particularly for OAR segmentation, since it offers an accurate assessment of the alignment between projected segmentation boundaries and the ground truth (manual expert segmentation). It quantifies the mean distance between the borders of the segmented item and the true organ boundary, providing critical information into the model's proficiency in delineating organ edges.

Clinical relevance of metrics used for accessing the model

The DSC is a sophisticated metric and this advanced measure evaluates the similarity between the predicted segmentation and the ground truth segmentation, enabling the assessment of the accuracy and robustness of the segmentation process.

In medical imaging, accurate border delineation is essential for ensuring that treatments, such as radiation therapy, precisely target the intended locations while protecting adjacent healthy tissue. Reduced 95% HD values indicate more clinically acceptable segmentation borders.

Medical procedures, such as radiation therapy and surgical planning, rely significantly on the precise delineation of anatomical boundaries. A high mean boundary distance indicates a substantial discrepancy between the anticipated boundary and the actual border, which may result in erroneous treatment zones. A low value indicates accurate border alignment, resulting in safer and more efficient treatments. In brachytherapy, a little inaccuracy in border definition may result in excessive radiation exposure to adjacent important tissues, such as the bladder or rectum. It quantifies and mitigates boundary errors, hence enhancing the protection of delicate organs during treatment.

Results

In this study, we evaluated the proposed ensemble learning technique using two datasets of brachytherapy CT images, each containing four structures: rectum, bladder, bowel, and sigmoid. The CT images had a 3 mm interslice thickness. This technique harnessed the combined strengths of different learning algorithms to improve the overall performance and reliability of the segmentation results. Our approach aimed to enhance precision and efficiency in cervix brachytherapy planning, addressing the challenges posed by these anatomical structures and metal artifacts from the applicators. The results, discussed below, provide valuable insights into the feasibility and effectiveness of the proposed ensemble learning technique, contributing to advancements in treatment planning and improving patient outcomes in cervix brachytherapy.

Segmentation results using Double U-Net, Bi-directional ConvLSTM U-Net, Transformer Network, and ensemble learning models

For OID dataset, using DUN, for all the OARs, the mean and standard deviation of DSC was 0.961 ± 0.001, 95% HD (mm) was 13.0 ± 0.2, and mean boundary distance was 8.0 ± 0.1. Using the BCUN, DSC was 0.972 ± 0.001, 95% HD was 12.9 ± 0.3, and mean boundary distance was 4.3 ± 0.1. For TN, DSC was 0.980 ± 0.001, 95% HD was 4.5 ± 0.1, and mean boundary distance was 3.1 ± 0.1. By employing combination of DUN + BCUN ensemble learning model, the mean and standard deviation of DSC was 0.992 ± 0.002, 95% HD was 1.9 ± 0.1, and mean boundary distance was 1.8 ± 0.1, and for BCUN + TN ensemble learning model, the mean and standard deviation of DSC was found to be 0.996 ± 0.001, 95% HD was 1.2 ± 0.1, and mean boundary distance was 1.3 ± 0.1. Table 1 shows the values of segmentation parameters for OID and DID datasets.

The locality and spread using DUN + BCUN ensemble learning model for DSC, 95% HD, and the mean boundary distance are

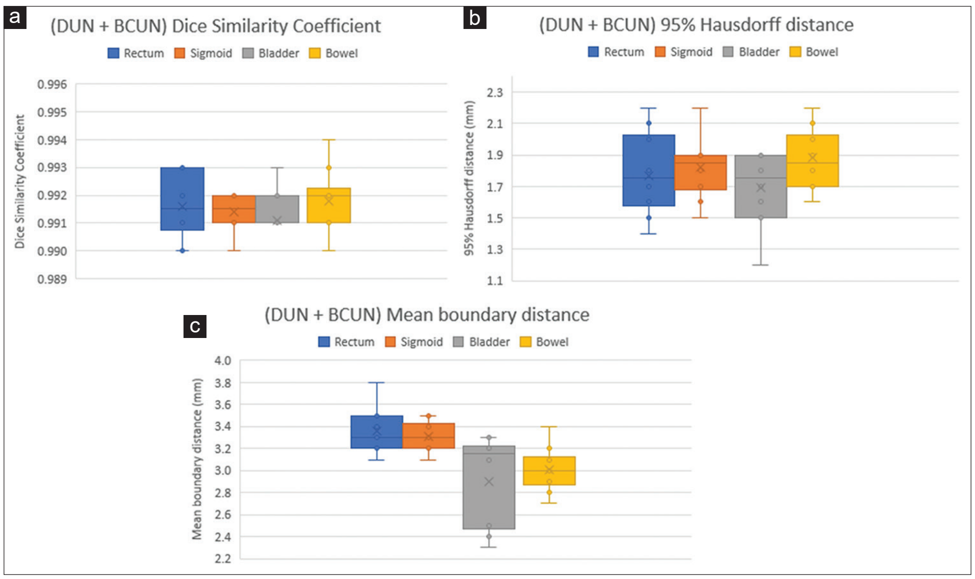

Figure 5: The Dice similarity coefficient (a), 95% Hausdorff distance (b), Mean boundary distance (c) for Double U-Net and Bi-directional ConvLSTM U-Net ensemble learning models. DUN: Double U-Net, BCUN: Bi-directional ConvLSTM U-Net

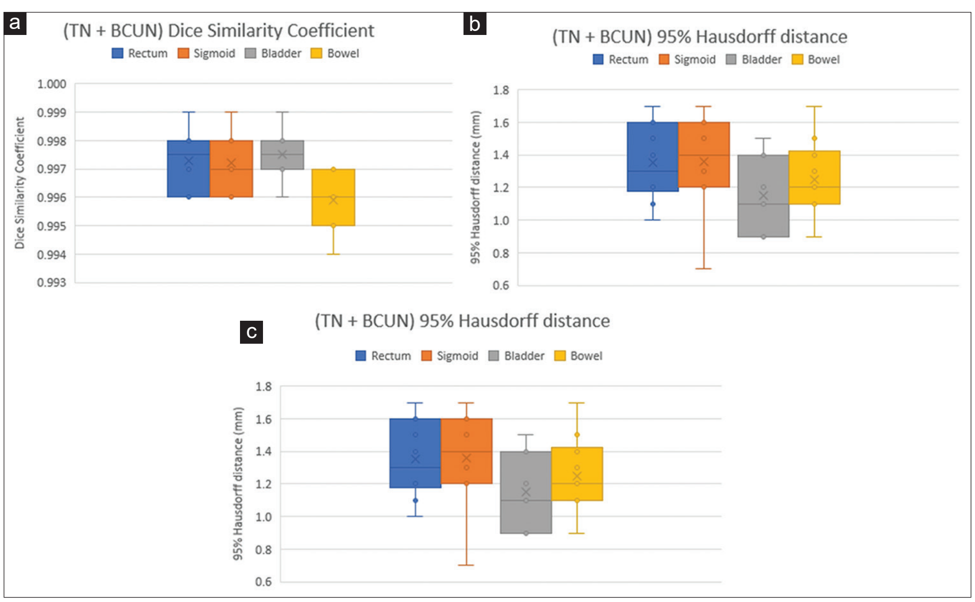

Figure 6: The Dice similarity coefficient (a), 95% Hausdorff distance (b), Mean boundary distance (c) for Bi-directional ConvLSTM U-Net and Transformer Network ensemble learning models. TN: Transformer Network, BCUN: Bi-directional ConvLSTM U-Net

shown in Figure 5a-c, respectively. Figure 6a-c shows the DSC, 95% HD, and the mean boundary distance using BCUN + TN ensemble learning model. Figure 7 represents the different test images with their corresponding mask and segmented mask for different OARs.

Comparison with existing work

We also evaluated our methods against other existing approaches. As illustrated in Table 2, the proposed method consistently outperforms the others across all parameters. This comparison clearly demonstrates the superior performance of our method in the segmentation task.

Discussion

The article aims to automate the registration-free segmentation process by leveraging three distinct models – DUN, BCUN, and TN – combined using an ensemble learning technique. These models were trained on cervix CT scans to segment four different OARs: rectum, bladder, bowel, and sigmoid. To enhance model accuracy and other segmentation metrics, the ensemble learning approach was employed. For testing, CT scans from 12 patients, manually segmented by oncologists, were used. Additionally, the model was evaluated on a dataset from a different institute, also consisting of scans from 10 patients. The results highlight the effectiveness of the proposed method. The average DSC, which measures the overlap between predicted and manual segmentations, was consistently high for all models across all tested OARs. The use of the ensemble learning technique further improved the results, indicating enhanced accuracy. The 95% HD, which measures the maximum distance between two sets, also showed satisfactory segmentation results. Furthermore, the proposed method was compared with other existing methods, and the findings suggest that it achieves high accuracy and maintains excellent segmentation metrics. By reducing the dependence on manual segmentation by human experts, this approach offers significant time savings. Automating this process can potentially expedite planning for brachytherapy

Table 1: The dice similarity coefficient, 95% Hausdorff distance (mm), and mean boundary distance for all the models for our institute as well as different institute dataset

Parameters	Method	Bladder		Rectum		Bowel		Sigmoid	
		OID	DID	OID	DID	OID	DID	OID	DID
DSC	DUN	0.967	0.966	0.959	0.96	0.961	0.963	0.958	0.957
	BCUN	0.975	0.975	0.973	0.973	0.972	0.973	0.975	0.976
	TN	0.987	0.988	0.985	0.986	0.978	0.977	0.987	0.988
	Ensemble of DUN + BCUN	0.991	0.992	0.992	0.992	0.992	0.991	0.991	0.992
	Ensemble of TN + BCUN	0.998	0.998	0.997	0.997	0.996	0.996	0.998	0.997
95% HD (mm)	DUN	12	12	12	12.1	12.8	12.7	13.7	13.8
	BCUN	12.6	12.5	13.4	13.5	11.6	11.5	13.6	13.5
	TN	4.6	4.5	4.8	4.7	4.9	4.7	5.2	5.3
	Ensemble of DUN + BCUN	1.7	1.7	1.8	1.8	1.9	1.8	1.7	1.6
	Ensemble of TN + BCUN	1.2	1.3	1.4	1.5	1.3	1.4	1.2	1.1
Mean boundary distance (mm)	DUN	7.7	7.9	8.2	8.1	8	8.1	8.3	8
	BCUN	3.9	4.1	4.5	4.7	4.3	4.4	4.6	4.5
	TN	3	3.2	3.4	3.5	3	3.1	3.3	3.2
	Ensemble of DUN + BCUN	1.5	1.6	1.8	1.7	1.8	1.7	1.7	1.6
	Ensemble of TN + BCUN	1.2	1.1	1.2	1.1	1.3	1.2	1.4	1.2

DUN: Double U-Net, BCUN: Bi-directional ConvLSTM U-Net, TN: Transformer Network, DSC: Dice similarity coefficient, HD: Hausdorff distance, OID: Our institute dataset, DID: Different institute dataset

Table 2: Comparison of the proposed method with existing methods

Study	Modal	Dice coefficient	95% HD (mm)	Mean boundary distance (mm)
Zhang *et al*.[34]	3D DSD U-Net	0.780	15.7	*
Jiang *et al*.[35]	Refine Net	0.720	35.8	*
Mohammadi *et al*.[36]	2D ResU-Net	0.950	1.9	*
Zabihollahy *et al*.[37]	3D Dense U-Net (coarse-to-fine)	0.870	2.8	*
Li *et al*.[38]	nnUnet (2DU-Net, 3DU-Net and 3D-Cascade U-Net)	0.860	5.8	*
Yoganathan *et al*.[39]	2.5D ResNet, 2.5D IRN	0.700	13.1	*
Wang *et al*.[40]	Modified CNN	0.880	5.8	*
Xue *et al*.[41]	Prompt-nnUnet	0.960	2.0	*
Kraus *et al*.[42]	Ensemble learning U-Net	0.982	1.5	*
Our method	Ensemble learning	0.998	1.3	1.2

CNN: Convolutional neural network, HD: Hausdorff distance, IRN: Iterative refinenet, DSD: Deep supervision discriminator, *Not Reported

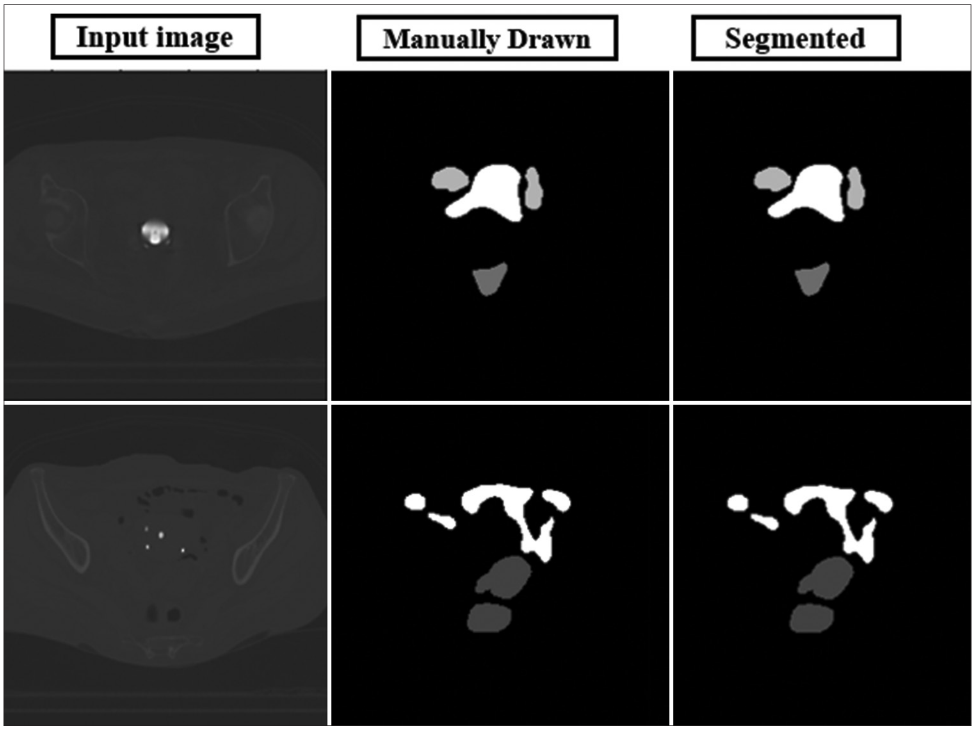

Figure 7: The input computed tomography scan image (left), manually drawn organ at risk (OAR) by an oncologist (middle), and segmented OAR by model (right)

cancer patients, leading to improved efficiency and precision in radiotherapy.

While the study presents promising results, certain considerations must be addressed. The evaluation was conducted on a relatively small dataset, and it would be beneficial to validate the approach on larger cohorts to establish its generalizability.[43-45] Future work could also explore integrating additional clinical features or imaging modalities to further enhance the accuracy and robustness of the segmentation models. Furthermore, the validation of the model on the bigger dataset collected from different regions of the country or throughout the world can enhance the reliability and accuracy of the model. This can be done through multicenter studies as we can increase the generalizability of the models across various clinical settings and demographic groups using data harmonization. Another method is augmenting data through synthetic and semi-synthetic approaches (like use of generative models). Another approach is transfer learning and domain adaptation where we can transfer the learning of one model and adapt trained models on one dataset to perform well on different datasets. Cross-modality data integration will also be helpful to overcome the problem of data scarcity. Moreover, employing a metal artifact reduction method could improve the accuracy of the segmentation.

Conclusion

This study successfully demonstrates the feasibility of federated automated multi-organ segmentation in cervical brachytherapy CT images by integrating advanced CNN architectures, namely DUN, BCUN, and Transformer Networks (TN), within an ensemble learning framework. This approach achieved remarkable segmentation accuracy, particularly for the four OARs in cervical cancer patients, reflected by a mean DSC of 0.998. These results signify substantial progress in automating medical image segmentation, potentially revolutionizing clinical workflows by improving precision and efficiency in treatment planning. Despite these promising outcomes, challenges remain. One of the primary limitations is the limited availability of annotated medical image datasets, which hampers further development and optimization of segmentation techniques. Future research should focus on validating the approach on larger and more diverse datasets to establish its generalizability. Multicenter studies, cross-modality data integration, and the use of synthetic or semi-synthetic data could address these limitations and further enhance the model's robustness. Additionally, incorporating metal artifact reduction methods could improve segmentation accuracy in real-world clinical applications.

In conclusion, this study highlights the powerful potential of combining DL models with ensemble learning techniques to automate the segmentation of OARs in cervical brachytherapy. This not only enhances precision in treatment planning but also holds promise for improving patient outcomes. Future work should explore broader applications and further refinement of these models to ensure their effective clinical deployment.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Lim YK, Kim D. Brachytherapy: A comprehensive review. Prog Med Phys 2021;32:25-39. Available from: https://synapse.koreamed.org/articles/1147261. [Last accessed on 2024 Jul 25].
2. Skowronek J. Current status of brachytherapy in cancer treatment – Short overview. J Contemp Brachyther 2017;9:581-9.
3. CancerSurvivorship-GoogleBooks.Availablefrom:https://books.google.co.in/books?hl=en&lr=and id=tVmRDwAAQBAJ&oi=fnd&pg=PA1&dq=importance+of+OAR+delineation +for+head+neck+cancers+for+radiotherapy+treatment&ots=fIfOqRX8Ha&sig=UcTTBsRV1T6OgKOBZX93J1q9xw&redir_esc=y#v=onepage&q=importance of OAR delineation. [Last accessed on 2023 May 02].
4. SENTHILKUMARAN N. Evolution In Medical Image Processing For Quick Diagnosis And Effective Treatment Planning. Webology 2021;18:2166–2175. Available from: http://www.webology.org/abstract.php?id=4840.
5. Annede P, Mailleux H, Sfumato P, Ferré M, Autret A, Varela Cagetti L, *et al.* Multivariate normal tissue complication probability modeling of vaginal late toxicity after brachytherapy for cervical cancer. Brachytherapy 2018;17:922-8.
6. Romano KD, Hill C, Trifiletti DM, Peach MS, Horton BJ, Shah N, *et al.* High dose-rate tandem and ovoid brachytherapy in cervical cancer: Dosimetric predictors of adverse events. Radiat Oncol 2018;13:129.
7. Kim Y, Kim YJ, Kim JY, Lim YK, Jeong C, Jeong J, *et al.* Toxicities and dose-volume histogram parameters of MRI-based brachytherapy for cervical cancer. Brachytherapy 2017;16:116-25.
8. Simpson DR, Scanderbeg DJ, Carmona R, McMurtrie RM, Einck J, Mell LK, *et al.* Clinical outcomes of computed tomography-based volumetric brachytherapy planning for cervical cancer. Int J Radiat Oncol Biol Phys 2015;93:150-7.
9. Heimann T, Meinzer HP. Statistical shape models for 3D medical image segmentation: A review. Med Image Anal 2009;13:543-63.
10. Haskins G, Kruger U, Yan P. Deep Learning in Medical Image Registration: A Survey. Available from: https://doi.org/10.1007/s00138-020-01060-x. [Last accessed on 2023 May 02].
11. Isselmou AE, Zhang S, Xu G. A novel approach for brain tumor detection using MRI images. J Biomed Sci Eng 2016;9:44-52.
12. Singh S, Singh BK, Kumar A. Multi-organ segmentation of organ-at-risk (OAR's) of head and neck site using ensemble learning technique. Radiography (Lond) 2024;30:673-80.
13. Singh S, Singh BK, Kumar A. Magnetic resonance imaging image-based segmentation of brain tumor using the modified transfer learning method. J Med Phys 2022;47:315-21.
14. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics); 2016. Available from: https://arxiv.org/abs/1606.06650v1. [Last accessed on 2023 May 02].
15. Dolz J, Desrosiers C, Ben Ayed I. 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. Neuroimage 2018;170:456-70.
16. Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, *et al*. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation; 2021. Available from: https://arxiv.org/abs/2102.04306v1. [Last accessed on 2024 Jul 25].
17. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, *et al*. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. ICLR 2021 – 9th Int Conf Learn Represent; 2020. Available from: https://arxiv.org/abs/2010.11929v2. [Last accessed on 2024 Jul 25].
18. Tanderup K, Pötter R, Lindegaard J, Kirisits C, Juergenliemk-Schulz I, de Leeuw A, *et al*. EMBRACE-II Protocol. EMBRACE II Study Protocol v.1.0, 2015. p. 0–132. Available from: https://www.embracestudy.dk/UserUpload/PublicDocuments/EMBRACE II Protocol.pdf. [Last accessed on 2024 Jul 25].
19. Bisong E. Building Machine Learning and Deep Learning Models on Google Cloud Platform; 2019. Available from: https://link.springer.com/10.1007/978-1-4842-4470-8. [Last accessed on 2023 May 02].
20. Jha D, Riegler MA, Johansen D, Halvorsen P, Johansen HD. DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation. Proc IEEE Symp Comput Med Syst; 2020. p. 558-64.
21. Azad R, Asadi-Aghbolaghi M, Fathy M, Escalera S. Bi-directional ConvLSTM U-net with densley connected convolutions. In: Proceedings-2019 International Conference on Computer Vision Workshop, ICCVW; 2019. p. 406-15. Available from: https://github.com/rezazad68/BCDU-Net. [Last accessed on 2024 Jul 25].
22. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, *et al*. Attention Is All You Need. Adv Neural Inf Process Syst 2017. p. 5999-6009. Available from: https://arxiv.org/abs/1706.03762v7. [Last accessed on 2024 Jul 25].
23. Karimi D, Dou H, Gholipour A, Senior Medical Image Segmentation Using Transformer Networks HHS Public Access. Available from: https://orcid.org/. [Last accessed on 2024 Jul 25].
24. Liu S, Yuan X, Hu R, Liang S, Feng S, Ai Y, *et al.* Automatic pancreas segmentation via coarse location and ensemble learning. IEEE Access 2020;8:2906-14.
25. Wang S, Yin Y, Cao G, Wei B, Zheng Y, Yang G. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. Neurocomputing 2015;149:708-17.
26. Hu J, Gu X, Gu X. Mutual ensemble learning for brain tumor segmentation. Neurocomputing 2022;504:68-81.
27. Sagi O, Rokach L. Ensemble learning: A survey. Wiley Interdiscip Rev Data Min Knowl Discov 2018;8:e1249. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/widm.1249. [Last accessed on 2023 May 02]
28. Ardabili S, Mosavi A, Várkonyi-Kóczy AR. Advances in machine learning modeling reviewing hybrid and ensemble methods. Lect Notes Networks Syst 2020;101:215-27.
29. Bell D, Moore C. Dice Similarity Coefficient. Radiopaedia.Org; 2020. Available from: https://radiopaedia.org/articles/75056. [Last accessed on 2023 May 03].
30. Zou KH, Warfield SK, Bharatha A, Tempany CM, Kaus MR, Haker SJ, *et al.* Statistical validation of image segmentation quality based on a spatial overlap index. Acad Radiol 2004;11:178-89.
31. Karimi D, Salcudean SE. Reducing the hausdorff distance in medical image segmentation with convolutional neural networks. IEEE Trans Med Imaging 2020;39:499-513.
32. Aydin OU, Taha AA, Hilbert A, Khalil AA, Galinovic I, Fiebach JB, *et al.* On the usage of average hausdorff distance for segmentation performance assessment: Hidden error when used for ranking. Eur Radiol Exp 2021;5:4.
33. Mackay K, Bernstein D, Glocker B, Kamnitsas K, Taylor A. A review of the metrics used to assess auto-contouring systems in radiotherapy. Clin Oncol (R Coll Radiol) 2023;35:354-69.
34. Zhang D, Yang Z, Jiang S, Zhou Z, Meng M, Wang W. Automatic segmentation and applicator reconstruction for CT-based brachytherapy of cervical cancer using 3D convolutional neural networks. J Appl Clin Med Phys 2020;21:158-69.
35. Jiang X, Wang F, Chen Y, Yan S. RefineNet-based automatic delineation of the clinical target volume and organs at risk for three-dimensional brachytherapy for cervical cancer. Ann Transl Med 2021;9:1721.
36. Mohammadi R, Shokatian I, Salehi M, Arabi H, Shiri I, Zaidi H. Deep learning-based auto-segmentation of organs at risk in high-dose rate brachytherapy of cervical cancer. Radiother Oncol 2021;159:231-40.
37. Zabihollahy F, Viswanathan AN, Schmidt EJ, Morcos M, Lee J. Fully automated multiorgan segmentation of female pelvic magnetic resonance images with coarse-to-fine convolutional neural network. Med Phys 2021;48:7028-42.
38. Li Z, Zhu Q, Zhang L, Yang X, Li Z, Fu J. A deep learning-based self-adapting ensemble method for segmentation in gynecological

brachytherapy. Radiat Oncol 2022;17:152.
39. Yoganathan SA, Paul SN, Paloor S, Torfeh T, Chandramouli SH, Hammoud R, *et al.* Automatic segmentation of magnetic resonance images for high-dose-rate cervical cancer brachytherapy using deep learning. Med Phys 2022;49:1571-84.
40. Wang J, Chen Y, Tu Y, Xie H, Chen Y, Luo L, *et al*. Evaluation of auto-segmentation for brachytherapy of postoperative cervical cancer using deep learning-based workflow. Phys Med Biol 2023;68:1250-62.
41. Xue X, Liang D, Wang K, Gao J, Ding J, Zhou F, *et al*. A deep learning-based 3D Prompt-nnUnet model for automatic segmentation in brachytherapy of postoperative endometrial carcinoma. J Appl Clin Med Phys 2024;25:e14371.
42. Kraus AC, Iqbal Z, Cardan RA, Popple RA, Stanley DN, Shen S, *et al.* Prospective evaluation of automated contouring for CT-based brachytherapy for gynecologic malignancies. Adv Radiat Oncol 2024;9:101417.
43. Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D. Image segmentation using deep learning: A survey. IEEE Trans Pattern Anal Mach Intell 2022;44:3523-42.
44. Hao S, Zhou Y, Guo Y. A brief survey on semantic segmentation with deep learning. Neurocomputing 2020;406:302-21.
45. Guo Y, Liu Y, Georgiou T, Lew MS. Review of semantic segmentation using deep neural networks. Int J Multimed Inf Retr 2018;7:87-93. Available from: https://link.springer.com/article/100.1007/s13735-017-0141-z. [Last accessed on 2024 Jul 25].