BMC Cancer

RESEARCH

Open Access

Automated contouring of gross tumor volume lymph nodes in lung cancer by deep learning

Yuling Huang[1,2], Xingxing Yuan[1,2], Longfei Xu[1,2], Junming Jian[1,2], Changfei Gong[1,2], Yun Zhang[1,2,3] and Wenheng Zheng[1,2*]

Abstract

Purpose The precise contouring of gross tumor volume lymph nodes (GTVnd) is an essential step in clinical target volume delineation. This study aims to propose and evaluate a deep learning model for segmenting GTVnd specifically in lung cancer, representing one of the pioneering investigations into automated segmentation of GTVnd specifically for lung cancer.

Method Ninety computed tomography (CT) scans of patients with stage Ⅲ-Ⅳ small cell lung cancer (SCLC) were collected, of which 75 patients were assembled into a training dataset and 15 were used in a testing dataset. A new segmentation model was constructed to enable the automatic and accurate delineation of the GTVnd in lung cancer. This model integrates a contextual cue enhancement module and an edge-guided feature enhancement decoder. The contextual cues enhancement module was used to enforce the consistency of the contextual cues encoded in the deepest feature, and the edge-guided feature enhancement decoder was used to obtain edge-aware and edge-preserving segmentation predictions. The model was quantitatively evaluated using the three-dimensional Dice Similarity Coefficient (3D DSC) and the 95th Hausdorff Distance (95HD). Additionally, comparative analysis was conducted between predicted treatment plans derived from auto-contouring GTVnd and established clinical plans.

Results The ECENet achieved a mean 3D DSC of 0.72 ± 0.09 and a 95HD of 6.39 ± 4.59 mm, showing significant improvement compared to UNet, with a DSC of 0.46 ± 0.19 and a 95HD of 12.24 ± 13.36 mm, and nnUNet, with a DSC of 0.52 ± 0.18 and a 95HD of 9.92 ± 6.49 mm. Its performance was intermediate, falling between mid-level physicians, with a DSC of 0.81 ± 0.06, and junior physicians, with a DSC of 0.68 ± 0.10. And the clinical and predicted treatment plans were further compared. The dosimetric analysis demonstrated excellent agreement between predicted and clinical plans, with average relative deviation of < 0.17% for PTV D2/D50/D98, < 3.5% for lung V30/V20/V10/V5/Dmean, and < 6.1% for heart V40/V30/Dmean. Furthermore, the TCP (66.99% ± 0.55 vs. 66.88% ± 0.45) and NTCP (3.13% ± 1.33 vs. 3.25% ± 1.42) analyses revealed strong concordance between predicted and clinical outcomes, confirming the clinical applicability of the proposed method.

Conclusion The proposed model could achieve the automatic delineation of the GTVnd in the thoracic region of lung cancer and showed certain advantages, making it a potential choice for the automatic delineation of the GTVnd in lung cancer, particularly for young radiation oncologists.

*Correspondence:
Wenheng Zheng
zhengwh06@foxmail.com

Full list of author information is available at the end of the article

Keywords Automatic segmentation, Gross target volume lymph nodes (GTVnd), Lung cancer, Deep learning

Introduction

In recent years, lung cancer has become the second most common cancer with the highest mortality rate in the world [1]. Radiation therapy is an important treatment for lung cancer and the accurate contouring of the gross target volume (GTV) of the lung is essential for the precise delivery of radiotherapy [2]. Lymph Node (LN) metastasis is common in lung cancer, particularly in inoperable patients. Precision delineation of the GTV lymph nodes (GTVnd) is a necessary step in clinical target delineation for lung cancer radiotherapy and was recommended by the treatment guidelines for lung cancer [3]– [4]. Although positron emission tomography CT (PET-CT) is the most appropriate method by which to identify the GTVnd, not all lung cancer patients undergo PET-CT scanning because of economic and other factors, particularly in underdeveloped regions [5]– [6]. Therefore, it is mainly segmented manually by oncologists in clinical application, highly dependent on manual experience, and results in significant heterogeneity, despite oncologists following the same guidelines [7]– [8]. Moreover, it is particularly difficult for beginners or inexperienced clinical oncologists to delineate the GTVnd. Therefore, it is of great clinical value to achieve automatic and accurate segmentation of the GTVnd in patients with lung cancer with LNs metastasis.

With the development of deep learning, convolutional neural network (CNN)-based methods have been successfully applied to the automatic contouring of lung cancer. Significant progress has been made in the study of automatic organ-at-risk (OAR) segmentation [9–11], the GTV of primary tumors [12], and the clinical target volume (CTV) in postoperative radiation therapy (PORT) [13]– [14]. Ibragimov et al. used CNNs for OARs segmentation in H&N CT images and obtained Dice Similarity Coefficient (DSC) values that varied from 37.4% for the chiasm to 89.5% for the mandible [15]; however, no target was segmented. UNet, proposed by Ronneberger et al. is primarily used for the segmentation of biomedical images [16] and did not specialize in target segmentation in the field of radiotherapy. Recently, Shen et al. proposed the DiUNet model to automatically delineate the CTV of patients with lung cancer with LN metastasis [17]. Although the mean value of the 3D DSC was greater than 0.7, the GTVnd was manually contoured and added to the CT slice as a second input in the CNN architecture. Recent work by Stephen Skett, et al. [18] demonstrating automated contouring of primary lung tumors and nodal disease using solely CT images highlights automated GTVnd segmentation for lung cancer as a promising and pioneering area of investigation.

The autodelineation of the GTVnd is difficult because of its unclear boundaries and variability in the size and shape. Although UNet, which was based on encoder-decoder architecture, has shown effective performance in conducting multilevel feature fusions, the network is insufficient to represent high-level features, particularly structures that are of significant importance for GTVnd recognition. Therefore, we propose a novel edge-guided contextual cue enhancement network (ECENet) to improve the segmentation performance of the GTVnd by introducing a contextual cue module and incorporating edge guidance into the feature learning of the segmentation network. A contextual cue modeling module is implemented to enhance spatial and semantic information at a distance. This study applies a deep learning approach to delineate lung cancer GTVnd, evaluating automated segmentation quality across geometric, dosimetric, and TCP/NTCP computational metrics.

Methods

Patient datasets

Eligible patients had histologically confirmed stage III-IV small cell lung cancer (SCLC), complete baseline thoracic CT scans (defined as coverage of entire lung parenchyma, primary tumor, and involved lymph nodes without severe artifacts), and receipt of concurrent chemoradiotherapy comprising intensity-modulated radiotherapy (IMRT; 60 Gy/30 fractions at 2 Gy/fraction, 5 fractions/week) with platinum-based chemotherapy. Among 90 patients meeting these criteria treated between January 2021-January 2024, all had documented GTVnd delineation data. The number of slices per patient ranged from 70 to 105, with a total of 8,758 slices. The data were 512×512, with a pixel pitch of 1.27×1.27 mm and a thickness of 5 mm, and they were acquired using a Brilliance CT Big Bore (Siemens Healthcare). The private information of the patients was kept confidential during data collection and processing. GTVnd was defined as the area of the LN stations with a maximum short axis of >1 cm. The GTVnd was delineated in the mediastinal and lung windows according to Radiation Therapy Oncology Group (RTOG) guidelines [19]. GTVnd was manually delineated by senior radiation oncologists and reviewed by two additional experienced radiation oncologists (each with over 10 years of expertise in pulmonary tumor radiotherapy) to ensure contouring quality, following the standard clinical protocol of the hospital's radiotherapy department.

Architecture of ECENet

As shown in Fig. 1, based on the encoder-decoder architecture of UNet [20], the encoder of the proposed

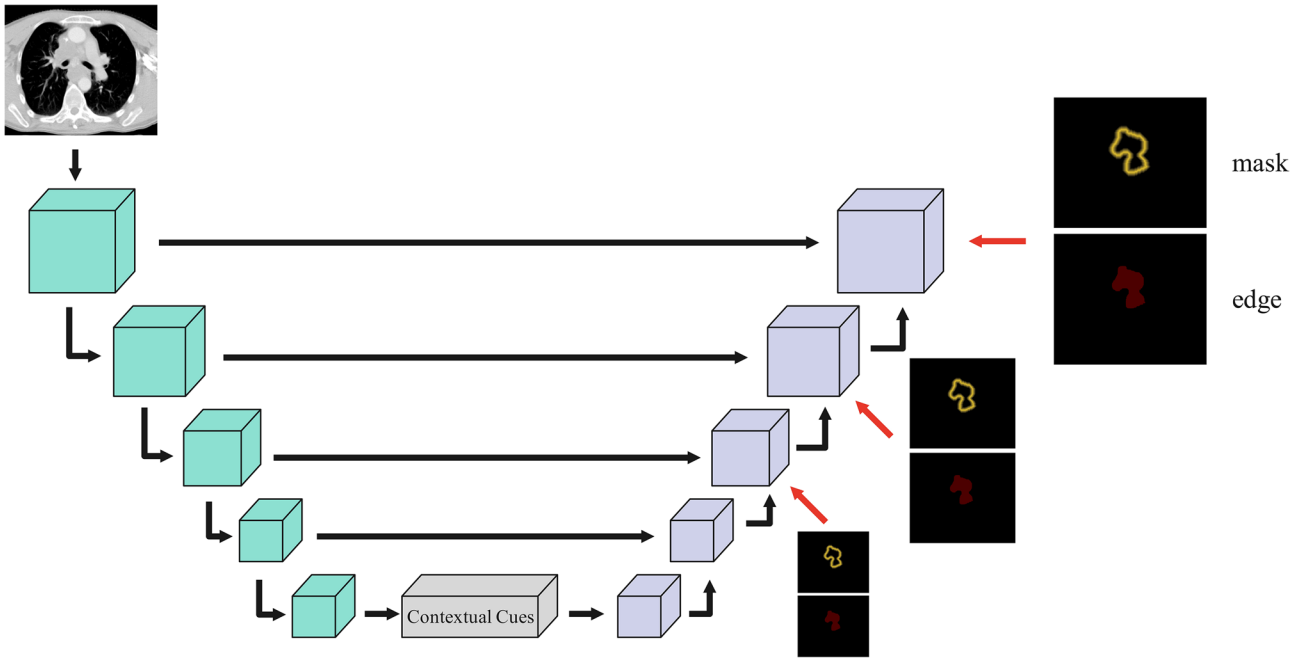

Fig. 1 Overall network structure of the ECENet

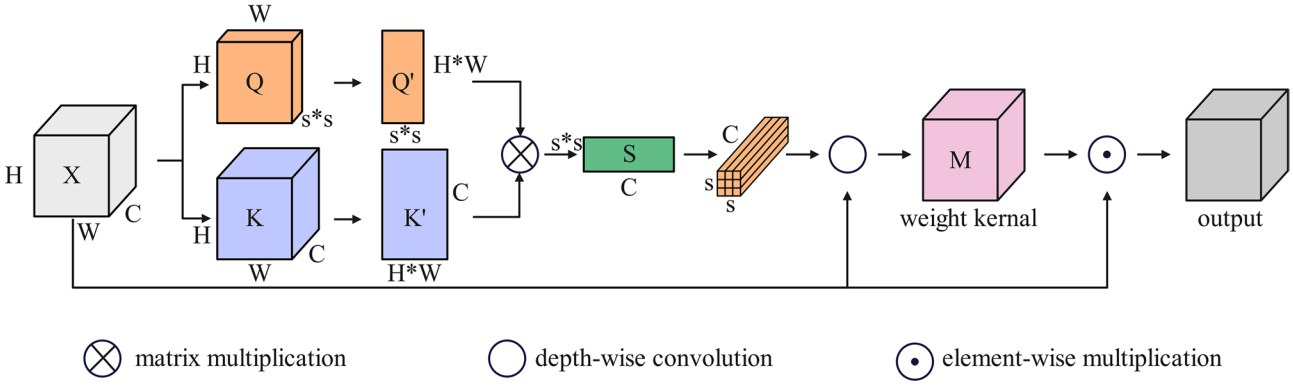

Fig. 2 Details of the proposed contextual cues enhancement module

ECENet extracts the multiscale features from the input image, including low-level spatial details and high-level semantic cues. The proposed contextual cue enhancement module is then used to enforce the consistency of the contextual cues encoded in the deepest feature. The proposed ECENet decoder aims to combine the extracted multiscale features. In this process, for the three lowest layers, the mask of the GTVnd and the edge of the GTVnd area are used to supervise and guide the simultaneous learning of segmentation mask prediction and edge prediction. In this manner, using the final learned edge-aware features, an edge-preserving segmentation prediction can be obtained.

(1) Contextual cues enhancement module.

The proposed contextual cue enhancement module consists of convolution kernel prediction and contextual cue aggregation submodules. As shown in Fig. 2, the convolution kernel prediction submodule is first used to model both the long-distance and short-distance semantic and spatial relations between any two pixels in the manner of matrix multiplication, thereby generating a spatial-adaptive contextual convolution kernel. The generated contextual convolution kernel is then applied to the original feature, leading to aggregated contextual cue enhanced features.

Concretely, for the deepest feature $X \in R^{H\times W\times C}$, the X is first transformed with two independent convolutional operations, generating two features: the key $K \in R^{H\times W\times C}$ and the query $Q \in R^{H\times W\times s^2}$, where H, W and C denote the height,

width and number of channels of the features, respectively, and s denotes the scale of the contextual convolution kernel. In order to model the relation between any two pixels, the feature of any pixel locations should be interacted in an explicit manner. To this end, the features of key and query are reshaped into a 2D shape: $K' \in \mathrm{R}^{(\mathrm{H}\times \mathrm{W})\times \mathrm{C}}$, $Q' \in \mathrm{R}^{(\mathrm{H}\times \mathrm{W})\times \mathrm{s}^2}$. For the key feature, the length of the feature vector for every pixel location is C. For the query feature, the length of the feature vector for every pixel location is s^2. To explicitly model the interaction between the feature vector of every pixel location of the key feature and the query feature, based on the self-attention mechanism, the element-wise matrix multiplication between the two feature vectors is calculated:

$$\mathrm{S}^{'}(\mathrm{i},\mathrm{j}) = \sum {}^{\mathrm{H}\times \mathrm{W}}_{\mathrm{q}=1} \mathrm{Q}'(\mathrm{q},\mathrm{i}) \times \mathrm{K}'(\mathrm{q},\mathrm{j}) \quad (1)$$

where $\mathrm{i} = 1,2,\ldots,\mathrm{s}^2$ and $\mathrm{j} = 1,2,\ldots,\mathrm{C}$. Thus, the contextual cues that enhance feature S' can be obtained. This operation can be rewritten as follows:

$$\mathrm{S}^{'} = \mathrm{Q}'^{\,\mathrm{T}} \times \mathrm{K}' \quad (2)$$

where $\mathrm{S}' \in \mathrm{R}^{\mathrm{s}^2\times \mathrm{C}}$ and $\mathrm{Q}'^{\,\mathrm{T}}$ is the transpose of Q'. Then, in order to obtain the final contextual convolution kernel, the $\mathrm{S}' \in \mathrm{R}^{\mathrm{s}^2\times \mathrm{C}}$ should be reshaped into $\mathrm{S} \in \mathrm{R}^{\mathrm{s}\times \mathrm{s}\times \mathrm{C}}$.

Furthermore, to generate the spatially adaptive weight factor M, the contextual convolution kernel S and the original input feature X were calculated using depthwise convolution. First, $\mathrm{S} \in \mathrm{R}^{\mathrm{s}\times \mathrm{s}\times \mathrm{C}}$ is split into C kernels, with each kernel being $\mathrm{s}\times \mathrm{s}$ in shape. Then, these kernels are applied for every feature channel of the input feature $\mathrm{X} \in \mathrm{R}^{\mathrm{H}\times \mathrm{W}\times \mathrm{C}}$. Finally, the weight factor M is multiplied with the original input feature X in an element-wise manner to generate the final contextual cue enhanced features.

(2) Edge-guided feature enhancement.

The decoder is designed to effectively combine the multiscale features generated by the encoder in a layer-wise manner. In this process, for the lowest three layers, both the mask of the GTVnd region and the edge of the GTVnd region are used for supervisions. The features of the decoder learn to predict the segmentation mask and edge of the GTVnd simultaneously. In this manner, the learned edge-aware features can generate edge-preserving predictions.

To obtain the ground truth of the edge of the GTVnd, we use the following operations to transform the mask of the GTVnd into the edge of the GTVnd. For each pixel, if the ground truth label of this pixel is different from that of any one of the eight neighboring pixels, then this pixel is an edge pixel.

(3) Loss function.

To supervise the mask learning of the proposed ECENet, the widely used binary segmentation loss function is used, which is formulated as follows:

$$L_{mask} = -\frac{1}{N}\sum {}^{N}_{i=1}[y_i \log(p_i) + (1-y_i)\log(1-p_i)] \quad (3)$$

where y_i and p_i means the ground-truth mask label and the mask prediction for the i-th pixel, N is the total number of pixels.

Moreover, since we introduce the edge guidance into the feature learning of decoder network, the edge loss function is also employed to supervise the edge-aware feature learning, which is formulated as follows:

$$L_{edge} = -\frac{1}{N}\sum {}^{N}_{i=1}[e_i \log(q_i) + (1-e_i)\log(1-q)] \quad (4)$$

where e_i means the edge label of i-th pixel, q_i refers to the edge prediction for the i-th pixel, N is the total number of pixels.

Experiments

Complete CT image data were randomly divided into a training dataset containing 75 cases and a testing dataset containing 15 cases in a ratio of 5:1. ECENet was trained for GTVnd contouring with the 75 cases in the training dataset and it was applied to predict 15 cases in the testing dataset. Stochastic gradient descent was used in the optimization, and the initial learning rate was set to 0.001. An NVIDIA RTX 4090 GPU was used in the experiments. Because ECENet was developed based on UNet, the performance of this model in contouring the GTVnd in lung cancer was comprehensively evaluated by comparing its performance with that of UNet. UNet, proposed by Ronneberger et al., features a symmetric encoder-decoder architecture for medical image segmentation. Its skip connections fuse shallow encoder features with deep decoder outputs, resolving downsampling-induced information loss. UNet was trained separately for GTVnd contouring under the same settings as ECENet. As an established benchmark in medical image segmentation, nnUNet was incorporated into our comparative analysis under identical experimental settings to ensure a rigorous performance evaluation [21]. nnUNet , proposed by Fabian Isensee et al., is a fully automated medical image segmentation framework that autonomously configures preprocessing, network architecture,

training protocols, and postprocessing for any new biomedical segmentation task. By systematizing the traditionally manual method configuration process—previously reliant on empirical tuning—nnU-Net achieves state-of-the-art performance in multiple international segmentation challenges without architectural innovations. Its name ("no new net") explicitly reflects this paradigm: superior outcomes stem from adaptive workflow orchestration rather than novel network design.

Quantitative evaluation metrics

To quantify the contouring accuracy, the Dice Similarity Coefficient (DSC) [22] and 95th percentile Hausdorff Distance (95HD) [23] were used. The DSC is defined as follows:

$$DSC\left(A,\ B\right)=\frac{2\left|A\cap B\right|}{\left|A\right|+\left|B\right|} \tag{5}$$

where A represents the predicted mask, B the GT mask, and $|A\cap B|$ the intersection of A and B. The DSC measured the relative volumetric overlap between two segmented masks, with a higher value indicating a higher overlap ratio and a value of one meaning that the two masks are identical.

The 95 HD is defined as:

$$\begin{gathered}95HD\left(A,B\right)=\text{percentile}\left[h\left(A,B\right)\cup\ h\left(B,A\right),95th\right]\\ h\left(A,B\right)=\underset{a\in\ A}{maxmin}\underset{b\in\ B}{\left\|a-b\right\|}\\ h\left(B,\ A\right)=\underset{b\in\ B}{maxmin}\underset{a\in\ A}{\left\|b-a\right\|}\end{gathered} \tag{6}$$

where $\|.\|$ represents the Euclidean norm of points A and B. A= {a_1, a_2, …, a_{n1}} and B= {b_1, b_2, …, b_{n2}} represent two finite point sets. The 95HD value indicates the 95th percentile of mismatches between A and B. The 95HD reflected the alignment between the wo contours, with a higher value indicating a larger difference. The DSC and 95HD were calculated at the 3D level, and the mean standard deviation (SD) values were calculated by averaging the values.

Plan evaluation

A comparative analysis was performed between the predicted treatment plans generated from auto-contoured GTVnd and established clinical plans. The predicted doses were derived from reoptimized plans, which were refined using the auto-contoured GTVnd and the original clinical target volume of the primary tumor (CTV-T). The total dose-volume histogram (DVH) curves of planned target volume (PTV) and distinct OARs were first exhibited between the prediction and clinical truth plans in terms of DVH parameters. Second, the clinically relevant dosimetric indices (DI) were determined, including the mean dose (Dmean), D2, D50, and D98 for PTV (where Di represents the dose received by i percent of PTV volume) and Dmean, V40, V30, V20, V10 and V5 for OARs (where Vi represents the volume fraction of OARs irradiated by i Gy). The following formula was used to calculate the homogeneity index (HI) and conformation index (CI) where V_{ptv} and V_{pres} are the volumes of the PTV and the prescription dose region, respectively, and V_{ref} is the volume of the irradiated PTV of the prescription dose.

$$\begin{aligned}HI&=\frac{D2\text{-}D98}{D50}\\ CI&=\frac{V_{ref}*V_{ref}}{V_{ptv}*V_{pre}}\end{aligned} \tag{7}$$

Equivalent Uniform Dose (EUD)

EUD is defined as the dose absorbed with a biological effect similar to nonhomogeneous irradiation, if applied uniformly to the tumour or normal tissue [24].

$$\begin{aligned}EUD&=\left(\sum\nolimits_{i=1}V_iEQD_i{}^{\text{a}}\right)^{1/\,a}\\ EQD_i&=D_i\frac{\frac{\alpha}{\beta}+\frac{D_i}{n_f}}{\frac{\alpha}{\beta}+2}\end{aligned} \tag{8}$$

where EQD_i is dose delivered to a sub volume V_i; a represents unitless model parameter, which is specific to the tumour of interest or normal structure; D_i is total dose received by the bin; n_f is total number of fraction; D_i/n_f is dose received by the bin at each fraction; α/β is parameters of the linear-quadratic (LQ) model (for a particular tumour or organ that is generally exposed), and the ratio was set at 10.0 Gy for tumor control probability (TCP) calculations and 3.0 Gy for normal tissue complication probability (NTCP) estimations, respectively.

The Poisson distribution is used to derive the Poisson linear quadratic model [25]. from the model of linear quadratic cell survival.

$$TCP=\left[\frac{1}{2}\right]^{\sum_i V_i*\exp\left[\frac{2\gamma\left(1-\frac{EQD2}{D50}\right)}{\ln 2}\right]} \tag{9}$$

where $D50$ is dose yields 50% of tumour control (51.24 Gy in the calculation); γ is normalized dose-response gradient (set at 0.83 for these computations); $EQD2$ is equivalent dose given in 2 Gy fraction.

Lyman-Kutcher-Burman (LKB) is used to estimate the complication probability for the heterogeneously irradiated normal organs [26].

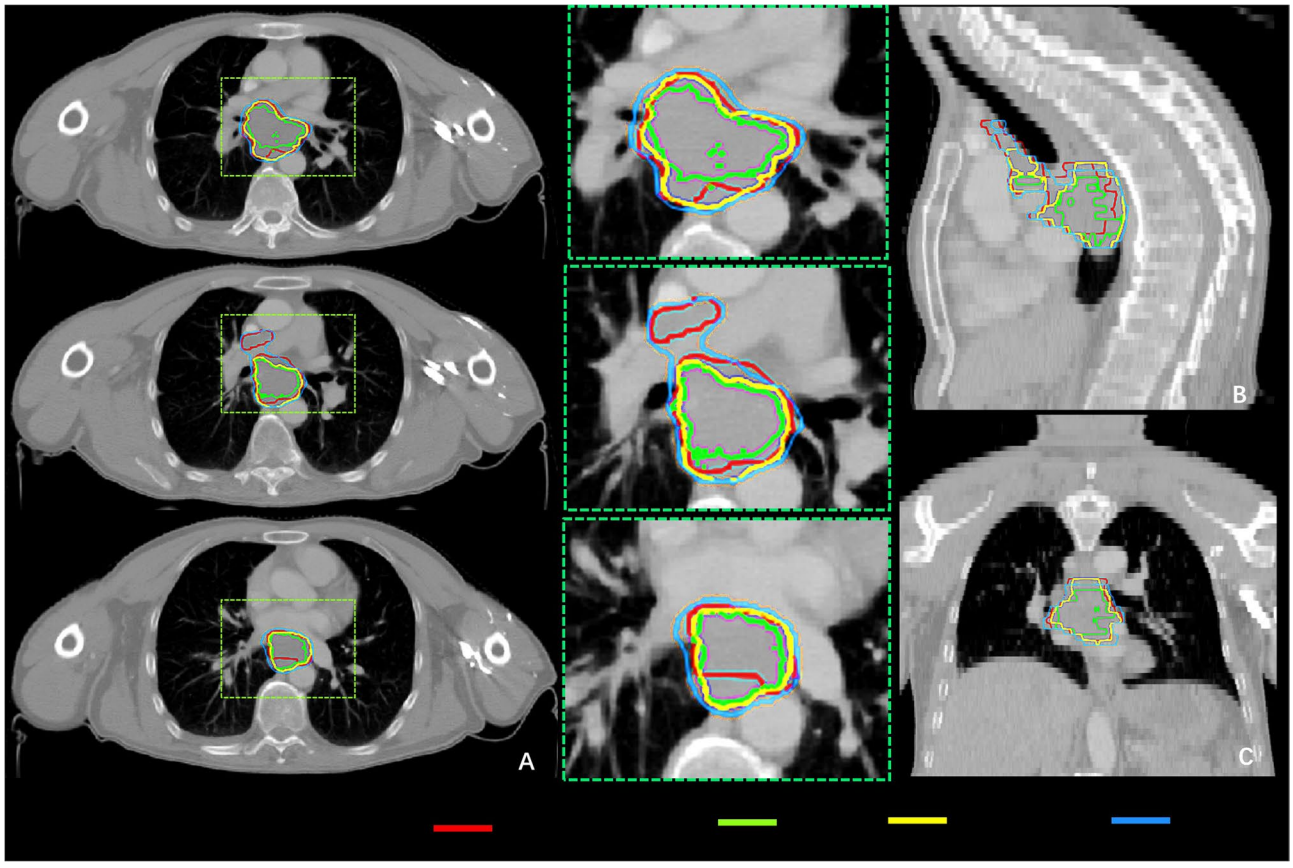

Fig. 3 Example of GTVnd segmentation slice for a patient predicted from UNet, nnUNet and ECENet. **A** GTVnd on three transversal planes, (**B**) GTVnd on a coronal plane, (**C**) GTVnd on a sagittal plane

$$
\begin{aligned}
NTCP &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx \\
t &= \frac{D_{eff} - TD_{50}}{m * TD_{50}} \\
D_{eff} &= \left(\sum\nolimits_i V_i * D_i^{\frac{1}{n}} \right)^n
\end{aligned}
\tag{10}
$$

where *TD50* (set at 30.8 Gy for these computations) is tolerance dose of OARs to produce 50% of complication (if uniformly irradiated); V_i is volume in a specific dose bin *i*; D_i is dose given to each bin; *m* is dimensionless parameter for determining the slope of complication probability versus dose curve (set at 0.37 for these computations); *n* is volume dependence of the complication's probability (set at 0.98 for these computations).

Results

Fig. 3 shown the GTVnd segmentation slice of one patient, including the transversal, coronal, and sagittal sections. The GTVnd was located in zones 2 and 4 of the lung cancer LN station, with a volume of 137.22 cc. The 3D DSC of this case from UNet, nnUNet and ECENet were 0.40, 0.74 and 0.83, respectively, and the 95HD values were 6.93 mm, 4.12 mm and 2.24 mm, respectively. Compared to the UNet models and the nnUNet, ECENet performed much better in the segmentation of the GTVnd, which was highly consistent with the oncologist's manual segmentation contours.

Table 1 The statistics of 3D DSC and 95HD with the volume of the GTVnd for unet, NnUnet and ECENet in the testing dataset

	DSC(mean ± SD)	95HD ± SD (mm)
UNet	0.46 ± 0.19	12.24 ± 13.36
nnUnet	0.52 ± 0.18	9.92 ± 6.49
ECENet	0.72 ± 0.09	6.39 ± 4.59

The 3D DSC and 95HD were calculated at a 3D level for the testing dataset and the mean ± standard deviation (SD) values for UNet, nnUNet and ECENet were calculated by averaging the values. The quantitative evaluation results were summarized in Table 1. The mean 3D DSCs for UNet, nnUNet and ECENet were 0.46 ± 0.19, 0.52 ± 0.18 and 0.72 ± 0.09, respectively, and the mean 95HDs were 12.24 ± 13.36 mm, 9.92 ± 6.64 mm and 6.39 ± 4.59 mm, respectively. The proposed ECENet tended to achieve higher accuracy than UNet and nnUNet by achieving greater 3D DSC values and lower 95HD values. Moreover, the proposed model performed

more robustly on the testing dataset, displaying a small average standard deviation. There was a statistically significant difference between the different models for both 3D DSC ($P < 0.001$) and 95HD ($P < 0.001$). In addition, the mean volume of GTVnd was 46.06 ± 42.33 cc. The standard deviation values are large, owing to the large differences in the GTVnd volume for each case in the testing dataset.

The contouring accuracy of ECENet was benchmarked against radiation oncologists stratified by clinical experience: senior physicians (> 10 years, $n = 2$), mid-level physicians (5–10 years, $n = 2$), and junior physicians (3–5 years, $n = 3$). All observers held specialized expertise in lung cancer radiotherapy and met institutional credentialing standards. The mean 3D DSC values were 0.90 ± 0.02 for senior physicians, 0.81 ± 0.06 for mid-level physicians, and 0.68 ± 0.10 for junior physicians. The performance of the auto-segmentation model was intermediate, achieving accuracy levels between those of mid-level and junior physicians.

The clinical DVH of the PTVs and OARs were compared to the predicted DVHs for the testing dataset. All treatment plans were optimized using the Eclipse™ treatment planning system (Varian Medical Systems) with inverse planning methodology. A prescription dose of 60 Gy in 30 fractions (2 Gy per fraction) was delivered over 6 weeks using intensity-modulated radiotherapy (IMRT), with 95% of the planning target volume (PTV) receiving 100% of the prescribed dose. Dose constraints adhered to RTOG 0617 protocols for normal tissues. The results revealed that there was an acceptable agreement between the clinical and predicted DVHs of the PTV and OARs. The mean and standard deviation of the clinically relevant DI for PTV and OARs in the testing dataset are presented in Table 2. The average relative dose differences in mean value of D2, D50 and D98 for the PTVs are within 0.17%, and the average relative volume differences in mean value of V30, V20, V10, V5 and Dmean for lung are within 3.5%. Meanwhile, those the average relative volume differences in mean value of V40, V30 and Dmean for Heart are within 6.1%, while that of V20 for Heart are 10.5%. The TCP analysis revealed close agreement between clinical and predicted outcomes, with mean values of 66.99% ± 0.55 and 66.88% ± 0.45, respectively. Regarding NTCP for lung, the clinical and predicted values also showed strong concordance, measuring 3.13% ± 1.33 and 3.25% ± 1.42, respectively. These results demonstrate excellent consistency between clinical observations and model predictions for both TCP and pulmonary NTCP endpoints. The projected findings are comparable to the clinical truth without any statistical significance ($P > 0.05$). Fig. 4 compares the dose-volume histograms (DVHs) of clinical and predicted plans for PTV, lung, and heart in a representative case. Close agreement is observed between the DVH curves. The Dmean difference for GTVnd is clinically negligible (Δ = 0.2%, 62.63 Gy vs. 62.48 Gy). The GTVnd-to-PTV volume ratio increased from 35.8% (60.3 cc vs. 168.4 cc) in the clinical plan to 40.7% (79.4 cc vs. 195.1 cc) in the predicted plan. Fig. 5 shows a comparison between the clinical and predicted isodoses of the case.

Table 2 The statistics of dosimetric parameters and TCP/NTCP in testing dataset

	Clinical	Prediction	*P* Value
PTV			
D98(Gy)	59.22 ± 0.22	59.20 ± 0.48	0.875
D50(Gy)	61.94 ± 0.41	61.91 ± 0.34	0.925
D2(Gy)	63.32 ± 0.54	63.21 ± 0.30	0.440
HI	0.066 ± 0.011	0.065 ± 0.010	0.700
CI	0.851 ± 0.017	0.830 ± 0.069	0.280
TCP	66.99 ± 0.55	66.88 ± 0.45	0.776
Lung			
V30(%)	43.14 ± 10.09	43.74 ± 9.56	0.125
V20(%)	32.54 ± 8.45	33.19 ± 8.07	0.078
V10(%)	22.16 ± 6.82	22.48 ± 6.57	0.394
V5(%)	12.68 ± 5.99	13.13 ± 6.06	0.100
Dmean(Gy)	11.10 ± 3.00	11.28 ± 3.05	0.061
NTCP	3.13 ± 1.33	3.25 ± 1.42	0.105
Heart			
V40(%)	13.17 ± 8.97	13.32 ± 8.89	0.510
V30(%)	7.35 ± 6.55	7.80 ± 6.55	0.151
V20(%)	4.46 ± 5.01	4.93 ± 5.18	0.133
Dmean(Gy)	6.92 ± 4.02	7.15 ± 4.15	0.173

Abbreviations:DI Dosimetric indices, *Dmean* Mean dose, *HI* Homogeneity index, *CI* Conformation index, *TCP* Tumour control probability, *NTCP* Normal tissue complication probability

Discussion

Few studies have been conducted on the automatic delineation of the GTVnd for LN metastasis in lung cancer, and oncologists mainly perform manual contours based on their clinical experience, which leads to heterogeneity between different oncologists and intra-observers at different times [27, 28]. However, consistent target delineation is essential for improving radiotherapy outcomes. Since the artificial intelligence successfully performed automatic target delineation, particularly in improving the accuracy of contours and reducing heterogeneities [29], an ECENet based on UNet was constructed to automatically delineate the contours of the GTVnd in lung cancer accurately and consistently. The segmentation performance of the model was also evaluated.

The DSC value was commonly used to assess the accuracy of auto-segmentation during radiation treatment planning. Because ECENet was developed based on UNet, its performance in contouring the GTVnd in lung cancer was comprehensively evaluated by comparing it with UNet. Furthermore, to ensure a comprehensive

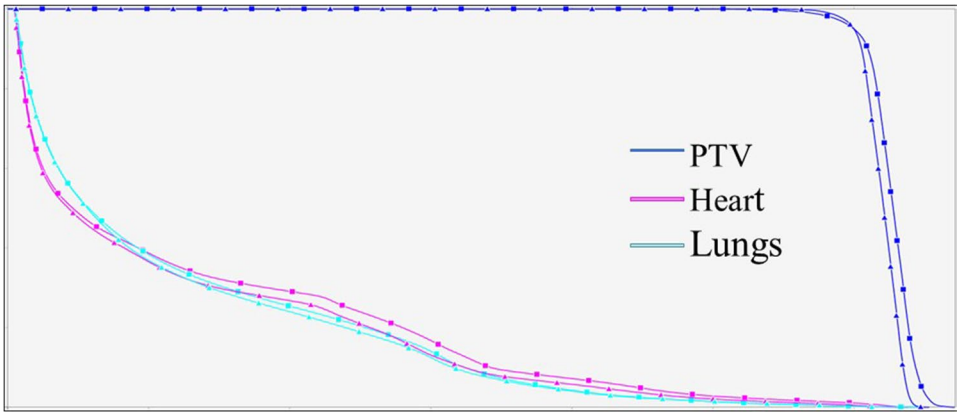

Fig. 4 Comparison of DVHs for the clinical plan (solid triangle) and predicted plan (solid square) for a representative case

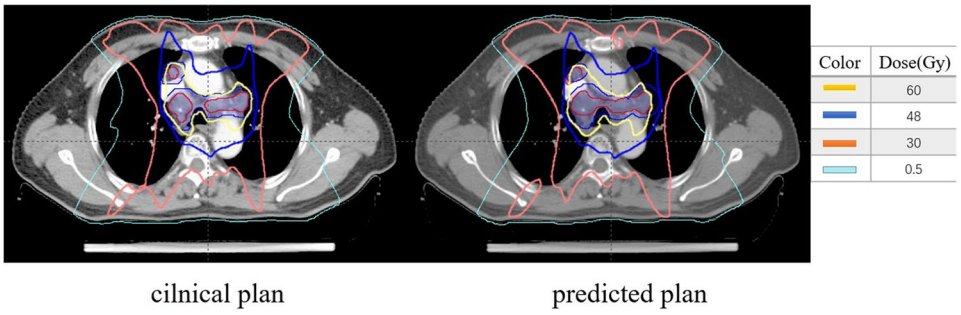

Fig. 5 Clinical versus predicted treatment plans for a representative case. The predicted plans based on the autocontoured GTVnd. PTV is displayed as a blue-shaded region, GTVnd is delineated by red contours within the PTV

evaluation, the state-of-the-art nnUNet framework was included in the comparison under identical experimental conditions. As shown in Table 1, the experimental results demonstrate that the proposed ECENet model achieves superior performance in pulmonary GTVnd delineation compared to both UNet and nnUNet, with a DSC value of 0.72 versus 0.46 and 0.52 for the two models, respectively. While these results were obtained by specifically optimizing and validating for pulmonary GTVnd segmentation in this particular clinical scenario, and thus cannot be generalized to claim the universal superiority of ECENet over other approaches, it is reasonable to conclude that it is feasible and effective to develop ECENet for segmenting GTVnd contours in lung cancer by implementing a contextual cue enhancement module.

The performance of the ECENet was intermediate, achieving accuracy levels between those of mid-level and junior physicians. This model can serve as a reliable tool assist junior physicians, particularly in reducing inter-observer variability and improving workflow efficiency. However, there remains room for further improvement when compared to the performance of senior physicians, highlighting the need for continued refinement to bridge this gap and enhance the model's ability to capture complex anatomical details. Due to clinical constraints, intra-observer variability was not assessed.

Although discrepancies were observed in GTVnd contouring between the auto-segmentation model and manual clinical delineation, dosimetric comparisons demonstrated clinically negligible differences between predicted and clinical plans for PTV coverage. This consistency may be attributed to the dominant volumetric contribution of the primary tumor clinical target volume (CTV-T), which diluted the impact of GTVnd variations from either auto-segmentation or manual delineation. For the illustrative case in Fig. 4, the GTVnd volume discrepancy amounted to 24.1% (60.3 cc vs. 79.4 cc), whereas the GTVnd-to-PTV ratio demonstrated a substantially smaller difference of 12% (40.7% vs. 35.8%). Consequently, the dosimetric sensitivity of clinical radiotherapy plans to GTVnd contour variations continues to mandate rigorous quantitative validation prior to clinical deployment.

In the context of GTVnd segmentation, the DSC values have been reported to be 0.46 [30] and 0.62 [31], whereas our proposed model demonstrated a high DSC value of 0.72. These models were not tested using the same dataset; therefore, it would be unfair to say that our proposed model was superior. However, it was reasonable to conclude that ECENet yields good results.

Although the segmentation accuracy of GTVnd is better than the previously reported, it still displays inferior

performance compared with the published GTV or OAR contouring models [32–34], in which the mean DSC values tend to be greater than 0.8. This reason was mainly because the GTV and OAR were characterized differently from the GTVnd. While most GTV or OAR regions had well-defined boundaries, the GTVnd did not, and the contour boundaries of GTVnd could easily be confused with the various complex anatomical structures around it. Moreover, there were considerable differences in the shape, volume, and location of the GTVnd among patients. As a result, the more complex GTVnd contouring performed inferiorly compared to the GTV or OAR contouring models.

Although the proposed ECENet model in this study demonstrates relatively high accuracy and potential in the automatic segmentation of lung cancer GTVnd, its primary limitation lies in the small scale of the dataset. The specific manifestations are as follows: Firstly, this was a single-center study with a relatively small sample size, and the segmentation accuracy of ECENet for GTVnd was inferior compared to that of the reported GTV or OAR contouring. Nevertheless, the proposed model achieves a higher DSC value in GTVnd segmentation tasks than previously reported models. Since Zijdenbos et al. suggested that a DSC value of >0.7 represents good overlap [35], this model with a DSC value of 0.72 can accurately achieve the automatic contouring of GTVnd in lung cancer. However, the segmentation accuracy of the GTVnd remains unsatisfactory for clinical application. Another limitation of our study is that it focuses solely on the automatic contouring of lymph nodes (GTVnd) and does not simultaneously address the automatic segmentation of primary tumor lesions. This narrow scope may limit the clinical applicability of our method, as radiotherapy planning typically requires precise delineation of both primary tumors and metastatic lymph nodes. Future research should aim to develop a comprehensive framework capable of jointly segmenting both primary tumor lesions and GTVnd to better support clinical workflows.

In general, this model can provide a potential benefit for the automatic contouring of GTVnd in lung cancer. In the future, the dataset size could be expanded to further optimize the model and improve its applicability and overall accuracy. Simultaneously, this technology can be explored for the automatic segmentation of GTVnd in other types of tumors.

Conclusion

The automatic and accurate segmentation of the GTVnd is key to achieving the automatic delineation of the CTV for lung cancer with LN metastasis. In this study ECENet was proposed for the fully automatic and accurate contouring of GTVnd in lung cancer. The model was based on UNet with edge guidance and spatial context enhancement. The quantitative evaluation results demonstrated that the proposed method achieved acceptable accuracy for GTVnd contour delineation, with predicted dose distributions showing strong agreement with clinical ground truth data. These findings suggest that the automated GTVnd contouring approach exhibits significant potential for clinical implementation in radiotherapy treatment planning. However, a multicenter clinical evaluation with more cases is needed to fully apply the model in clinical practice. In conclusion, this study explored the application of deep learning to the automatic delineation of the GTVnd in lung cancer. ECENet has shown promising potential in terms of contouring accuracy, offering a robust framework that could significantly contribute to the automation of GTVnd delineation.

Authors' contributions

Huang, Zheng and Gong wrote the main manuscript text and Jian prepared figures. Yuan, Zhang and Xu collection and processing data. All authors reviewed the manuscript.

Funding

This work was supported by Science and Technology Research Project of Jiangxi Provincial Department of Education (No.GJJ2203558), Science and Technology Project of Jiangxi Provincial Health Commission of China (No.202310054), Jiangxi Cancer Hospital Doctoral Research Start-up Fund (No. BSQDJ202301) and the Jiangxi Cancer Hospital Scientific Research Open Fund Project (No. KFJJ2023YB19).

Data availability

The data are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This study was approved by the ethics committee of the Jiangxi Cancer Hospital (2023ky192) and individual informed consent for this study was waived for this retrospective analysis. All procedures were conducted in accordance with the principles outlined in the Declaration of Helsinki.

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Radiation Oncology, Jiangxi Cancer Hospital, Nanchang, Jiangxi 330029, PR China
[2]The Second Affiliated Hospital of Nanchang Medical College, Nanchang, Jiangxi 330029, PR China
[3]JXHC Key Laboratory of Tumor Microenvironment and Immunoregulation, Jiangxi Cancer Hospital, Nanchang, Jiangxi 330031, PR China

Received: 29 February 2024 / Accepted: 1 August 2025
Published online: 30 September 2025

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.

2. Van Rikxoort EM, de Hoop B, Viergever MA, et al. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. Med Phys. 2009;36(7):2934–47.
3. Ettinger DS, Wood DE, Akerley W, Bazhenova LA, Borghaei H, Camidge DR, et al. Non-small cell lung cancer, version 6.2015. J Natl Compr Cancer Netw. 2015;13(5):515–24.
4. De Ruysscher D, Faivre-Finn C, Nestle U, Hurkmans CW, Le Péchoux C, Price A, et al. European organisation for research and treatment of cancer recommendations for planning and delivery of high-dose, high-precision radiotherapy for lung cancer. J Clin Oncol. 2010;28(36):5301–10.
5. Petersen H, Holdgaard PC, Madsen PH, et al. FDG PET/CT in cancer: comparison of actual use with literature-based recommendations. Eur J Nucl Med Mol Imaging. 2016;43:695–706.
6. Fletcher JW, Djulbegovic B, Soares HP, et al. Recommendations on the use of 18F-FDG PET in oncology. J Nucl Med. 2008;49:480–508.
7. Benjamin E, Nelms WA, Tomé G, Robinson James wheeler. Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer. Int J Radiat Oncol Biol Phys. 2010;82(1):368–78.
8. Coen Rasch A, Eisbruch, et al. Irradiation of paranasal sinus tumors, a delineation and dose comparison study[J]. Int J Radiat Oncol Biol Phys. 2002;52(1):120–7.
9. Uzelaltinbulat S, Ugur B. Lung tumor segmentation algorithm. Procedia Comput Sci. 2017;120:140–7.
10. Soltani-nabipour J, Khorshidi A, Noorian B. Lung tumor segmentation using improved region growing algorithm. Nucl Eng Technol. 2020;52(10):2313–9.
11. Bolan Yang, Dehui Xiang, Feihong Yu, et al. Lung tumor segmentation based on the multi-scale template matching and region growing, Proc. SPIE 10578, Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging. 2018;105782Q. https://doi.org/10.1117/12.2293065.
12. H. Mi, C. Petitjean, B. Dubray, P. Vera and S. Ruan. Automatic lung tumor segmentation on PET images based on random walks and tumor growth model. Beijing, China: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI); 2014. pp. 1385–8. https://doi.org/10.1109/ISBI.2014.6868136.
13. Y. Nian et al. Graph-based unsupervised segmentation for lung tumor CT images. Chengdu, China: 2017 3rd IEEE International Conference on Computer and Communications (ICCC); 2017. pp. 1884–8. https://doi.org/10.1109/CompComm.2017.8322865.
14. Oda H, Roth HR, Bhatia KK, et al. Dense volumetric detection and segmentation of mediastinal lymph nodes in chest CT images. In K Mori, & N Petrick (Eds.), Medical Imaging 2018: Computer-Aided Diagnosis Article 1057502 (Progress in Biomedical Optics and Imaging - Proceedings of SPIE; Vol. 10575). SPIE. https://doi.org/10.1117/12.2287066.
15. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. Med Phys. 2017;44(2):547–57. https://doi.org/10.1002/mp.12045.
16. Ronneberger O, Fischer P, Brox T, UNet. Convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer Assisted Intervention, 2015, 9351: 234–241.
17. Shen J, Zhang F, Di M, et al. Clinical target volume automatic segmentation based on lymph node stations for lung cancer with bulky lump lymph nodes. Thorac Cancer. 2022;13(20):2897–903.
18. Skett S, et al. Autocontouring of primary lung lesions and nodal disease for radiotherapy based only on computed tomography images. Phys Imaging Radiat Oncol. 2024;31: 100637.
19. Bradley JD, Paulus R, Komaki R, et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): a randomised, two-by-two factorial phase 3 study. Lancet Oncol. 2015;16(2):187–99.
20. Ronneberger O, Fischer P, Brox T, UNet. Convolutional networks for biomedical image segmentation. arXiv preprint arXiv:1505.04597, 2015.
21. Isensee F, Jaeger PF, Kohl SAA, et al. NnU-net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18:203–11. https://doi.org/10.1038/s41592-020-01008-z.
22. Dice LR. Measures of the amount of ecologic association between species. Ecology. 1945;26(3):297–302.
23. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the hausdorff distance. IEEE Trans Pattern Anal Mach Intell. 1993;15(9):850–63.
24. Mercieca S, Belderbos JSA, van Baardwijk A, et al. The impact of training and professional collaboration on the interobserver variation of lung cancer delineations: a multi-institutional study. Acta Oncol. 2019;58(2):200–8.
25. Niemierko A. A generalized concept of equivalent uniform dose (EUD). Med Phys. 1999;26:1100. https://doi.org/10.1118/1.598063.
26. Wang H, Cooper BT, Schiff P, et al. Dosimetric assessment of tumor control probability in intensity and volumetric modulated radiotherapy plans. Br J Radiol. 2018;92:1094. https://doi.org/10.1259/bjr.20180471.
27. Kutcher G, Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: the effective volume method. Int J Radiat Oncol Biol Phys. 1989;16:1623–30. https://doi.org/10.1016/0360-3016(89)90972-3.
28. Guangying Z, Tingyi X, Lvhua W, et al. Consensus and controversies on delineation of radiotherapy target volume for patients with non-small cell lung cancer. Chin J Rasiat Oncol. 2008;17(6):432–6.
29. Barkati M, Simard D, Taussky D, et al. Magnetic resonance imaging for prostate bed radiotherapy planning: an inter- and intra-observer variability study. J Med Imaging Radiat Oncol. 2016;60:255–9. https://doi.org/10.1111/1754-9485.12416.
30. Yang J, Beadle BM, Garden AS, et al. Auto-segmentation of low-risk clinical target volume for head and neck radiation therapy. Pract Radiat Oncol. 2014;4(1):e31–7. https://doi.org/10.1016/j.prro.2013.03.003.
31. Kuo Men X, Chen Y, Zhang, et al. Deep deconvolutional neural network for target segmentation of nasopharyngeal cancer in planning computed tomography images. Front Oncol. 2017;7:315. https://doi.org/10.3389/fonc.2017.00315. PMID: 29376025; PMCID: PMC5770734.
32. Thibault Marin Y, Zhuo RM, Lahoud et al. Deep learning-based GTV contouring modeling inter- and intra- observer variability in sarcomas, radiotherapy and oncology, 167, 2022, Pages 269–276, ISSN 0167–8140, https://doi.org/10.1016/j.radonc.2021.09.034
33. Lang DM, Peeken JC, Combs SE et al. Deep learning based GTV delineation and progression free survival risk score prediction for head and neck cancer patients. Lecture Notes in Computer Science, vol 13209. Springer, Cham. https://doi.org/10.1007/978-3-030-98253-9_14
34. Shi F, Hu W, Wu J, et al. Deep learning empowered volume delineation of whole-body organs-at-risk for accelerated radiotherapy. Nat Commun. 2022;13:6566. https://doi.org/10.1038/s41467-022-34257-x.
35. Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans Med. 1994 13(4):716–24. https://doi.org/10.1109/42.363096

Publisher's Note