Physics and Imaging in Radiation Oncology 32 (2024) 100655

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

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

Original Research Article

Automatic segmentation for magnetic resonance imaging guided individual elective lymph node irradiation in head and neck cancer patients

Floris C.J. Reinders [a,*], Mark H.F. Savenije [a,b], Mischa de Ridder [a], Matteo Maspero [a,b], Patricia A.H. Doornaert [a], Chris H.J. Terhaard [a], Cornelis P.J. Raaijmakers [a], Kaveh Zakeri [c], Nancy Y. Lee [c], Eric Aliotta [c], Aneesh Rangnekar [c], Harini Veeraraghavan [c], Marielle E.P. Philippens [a]

[a] Department of Radiotherapy, University Medical Centre Utrecht, the Netherlands
[b] Computational Imaging Group for MR Therapy and Diagnostics, Cancer and Imaging Division, University Medical Center Utrecht, Utrecht, the Netherlands
[c] Department of Radiotherapy, Memorial Sloan Kettering Cancer Centre, New York, United States

ARTICLE INFO

Keywords:
Deep learning
Artificial intelligence
Lymph nodes
Magnetic resonance imaging
Radiotherapy
Squamous cell carcinoma of head and neck
Elective neck irradiation

ABSTRACT

Background and purpose: In head and neck squamous cell carcinoma (HNSCC) patients, the radiation dose to nearby organs at risk can be reduced by restricting elective neck irradiation from lymph node levels to individual lymph nodes. However, manual delineation of every individual lymph node is time-consuming and error prone. Therefore, automatic magnetic resonance imaging (MRI) segmentation of individual lymph nodes was developed and tested using a convolutional neural network (CNN).

Materials and methods: In 50 HNSCC patients (UMC-Utrecht), individual lymph nodes located in lymph node levels Ib-II-III-IV-V were manually segmented on MRI by consensus of two experts, obtaining ground truth segmentations. A 3D CNN (nnU-Net) was trained on 40 patients and tested on 10. Evaluation metrics were Dice Similarity Coefficient (DSC), recall, precision, and F1-score. The segmentations of the CNN was compared to segmentations of two observers. Transfer learning was used with 20 additional patients to re-train and test the CNN in another medical center.

Results: nnU-Net produced automatic segmentations of elective lymph nodes with median DSC: 0.72, recall: 0.76, precision: 0.78, and F1-score: 0.78. The CNN had higher recall compared to both observers ($p = 0.002$). No difference in evaluation scores of the networks in both medical centers was found after re-training with 5 or 10 patients.

Conclusion: nnU-Net was able to automatically segment individual lymph nodes on MRI. The detection rate of lymph nodes using nnU-Net was higher than manual segmentations. Re-training nnU-Net was required to successfully transfer the network to the other medical center.

1. Introduction

Radiotherapy for patients with head and neck squamous cell carcinoma (HNSCC) is associated with dose-dependent long-term toxicity such as dysphagia [1], xerostomia [2], hypothyroidism [3], and carotid stenosis [4]. Our center aims to use new imaging techniques to redefine target volumes in HNSCC patients, reducing radiotherapy toxicity without compromising oncological outcomes.

Radiotherapy for HNSCC patients generally includes a high dose to the primary tumor and lymph nodes with macroscopic metastases, with elective neck irradiation (ENI) often added as a lower radiation dose to lymph node levels at risk for microscopic metastases. Given that regional recurrence rates are low in literature (1–5 %) [5,6], the ENI dose might be too high or the elective target volumes too large [7]. Studies already have successfully reduced the ENI dose without increasing the regional recurrence rates [8–11].

Abbreviations: HNSCC, head and neck squamous cell carcinoma; ENI, elective neck irradiation; MRI, magnetic resonance imaging; MR-Linac, MRI Linac; CNNs, convolutional neural networks; T2 mDixon TSE, multiple Dixon T2-weighted turbo spin echo; EORTC, European Organization for Research and Treatment of Cancer; GPU, graphical processing unit; SPSS, Statistical Package for Social Sciences.

* Corresponding author.
E-mail address: f.c.j.reinders-2@umcutrecht.nl (F.C.J. Reinders).

https://doi.org/10.1016/j.phro.2024.100655
Received 1 April 2024; Received in revised form 26 September 2024; Accepted 26 September 2024
Available online 27 September 2024

The radiation dose to the surrounding tissues can be further reduced by converting ENI target volumes from conventional lymph node levels to individual lymph nodes within these levels. These so-called "elective lymph nodes" are not suspected of containing overt metastases based on histology or radiology, but there is a risk of occult metastases, warranting elective treatment [5,7,12]. The MRI Linac (MR-Linac) will be used for this new treatment concept, as elective lymph nodes of HNSCC patients are better visualized with magnetic resonance imaging (MRI) due its superior soft tissue contrast compared to CT [13–15]. Moreover, day-to-day dose delivery can be closely monitored and radiotherapy plans can be adapted if necessary. In a previous planning study comparing the new concept with conventional treatment, significant reductions in the mean dose of >5 Gy were achievable in the submandibular gland, carotid arteries, and thyroid gland [14].

Although individual elective lymph node irradiation is promising, the large number of target structures requiring segmentation poses a practical challenge to its clinical adoption. With approximately 70 elective lymph nodes per patient on MRI, manual segmentation by the radiation oncologist is time-consuming. Manual target and OAR segmentation for HNSCC patients takes 2–3 h [16,17], with an estimated additional hour for elective lymph nodes.. This extra time, along with delineating multiple small structures might increase the risk of errors.

Automatic segmentation of elective lymph nodes could address these issues. Recent developments in model-based [18,19], atlas-based [20–22], and machine learning methods [23] have enabled automatic segmentation of OARs in HNSCC patients. Deep learning networks, including convolutional neural networks, provide the best performance [24–27] and save time in clinical workflows [28]. Unfortunately few MRI-based deep learning models exists for automatically segmenting the primary tumor, elective treatment fields, and OARs in HNSSC patients [29–32], and none are available for elective lymph nodes.

In this study, a state-of-the-art 3D convolutional neural network (CNN) was trained and tested to automatically segment the elective and suspect individual lymph nodes in HNSCC patients on MRI. Different observers' manual lymph node segmentations were compared with the automatic segmentations of the CNN. Additionally, transfer learning was evaluated to assess the CNN's performance on data from another medical center.

2. Materials and methods

2.1. Study design and patient selection

MR images of 50 patients with HNSCC from the larynx, hypopharynx, or oropharynx were retrospectively collected in the UMCU, the Netherlands. For transfer learning purposes, additional MR images of 20 patients were retrospectively collected at Memorial Sloan Kettering Cancer Centre (MSKCC), New York, USA. All patients consented to use their imaging data (UMCU PREDICT = NL57164.041.16, UMCU UPGRADE = NL46354.091.15 or MSKCC = 16–1648). Patient, tumor, and treatment characteristics are summarized in Supplementary Table 1/2.

2.2. Imaging

Individual lymph nodes and lymph node levels Ib-II-III-IV-V were identified using multiple Dixon T2-weighted turbo spin echo (T2 mDixon TSE) MRI without using exogenous contrast [33]. Water and in-phase images of the T2 mDixon TSE were used as input channels CNN training. Both medical centers used different vendors, receiver coils, echo/repetition times, slice thickness, and in-plane resolutions (Supplementary Table 3). No 3D scan was used as the fat suppression of these sequences is not sufficient in the lower neck.

During all imaging procedures, patients in the UMCU were immobilized in a custom-made 5-point thermoplastic mask (MacroMedics) to prevent motion and ensure reproducible patient setup. In the MSKCC, no fixation procedures were applied.

2.3. Ground truth segmentations

All individual lymph node segmentations and lymph node level segmentations (to select lymph nodes inside these levels) were performed manually on the water-only image of the T2 mDixon TSE, as with these images individual lymph nodes can be distinguished from the surrounding fatty environment. The in-phase image of the T2 mDixon TSE provided additional information for segmenting the lymph node levels (Fig. 1). All individual lymph nodes visible in at least two sequential transverse slices were identified and delineated. The conventional lymph node levels Ib/II/III/IV/V were segmented according to the guidelines published by the European Organization for Research and Treatment of Cancer (EORTC) [34].

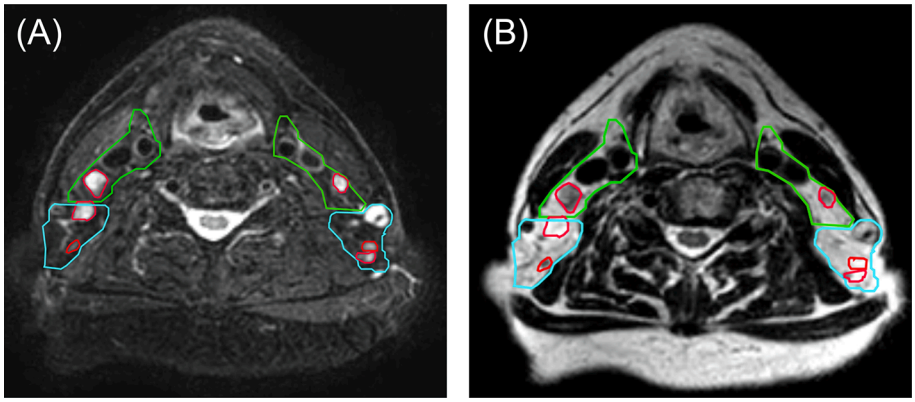

Fig. 1. Transverse water-only image (A) and in-phase image (B) of a T2 mDixon TSE MRI of a HNSCC patient depicting individual lymph nodes (red), lymph node levels III (green), and level V (blue). The visibility of individual lymph nodes is better on the water image, whereas the borders of the lymph node levels are better visible on the in-phase image.

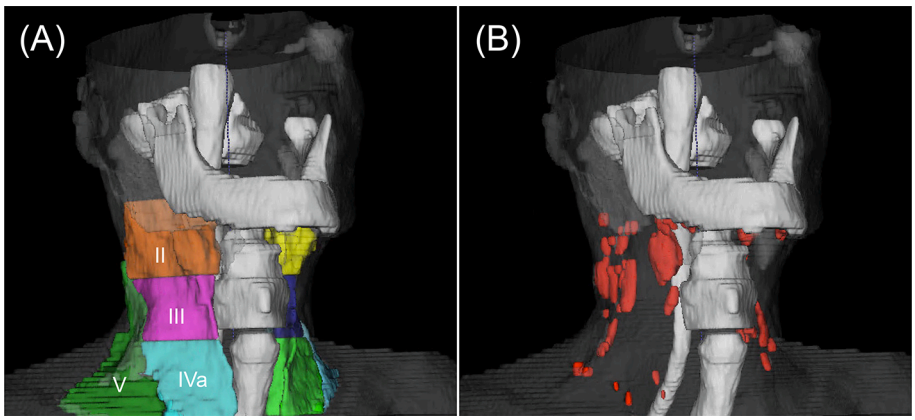

Fig. 2. 3D example of automatic segmentations of lymph node levels II/III/IV/V on the left side (A) and individual lymph nodes on the right side (B) in one HNSCC patient produced by a trained neural network (nnU-Net).

One observer (FR[1]) manually segmented individual lymph nodes and lymph node levels. The segmentations were reviewed with one of the dedicated HNSCC radiation oncologists in both departments (UMCU: MR[1], PD[1], and CT[1]. MSKCC: KZ[2]) and if necessary adaptations were made by consensus. The approved segmentations served as ground truth (GT).

2.4. Training convolutional neural networks

In this study nnU-Net [35] served as state-of-the-art CNN based on deep learning principles. nnU-Net won several segmentation challenges so far [35] and comes with the advantage that it has automatic task-dependent hyper-parameter optimization.

Two different CNNs were developed, one for segmenting the lymph node levels and one for the individual lymph nodes. For each patient lymph node levels were predicted by the CNN and manually checked by the radiation oncologist. The confirmed level segmentations were used as boundaries during post-processing to select the required lymph nodes. Both elective and metastatic individual lymph nodes were included as GT segmentations.

The UMCU images and GT segmentations of 40 patients were used as training cases (UMCU train set). A graphical processing unit Tesla P100 (NVIDIA Corporation) with 16 GB of memory was used for training. All pre-processing, training, and post-processing steps are described in Supplementary file 1. To evaluate the potential of transfer learning for reducing the anticipated performance decline when applying the network in a different medical center, we re-trained the CNN with a limited dataset comprising images and ground truth segmentations of ten MSKCC patients (MSKCC training set). The pre-trained weights of the networks based on the UMCU dataset were transferred to the new model prior to retraining with MSKCC data. The learning rate of the networks was reduced to 1×10^{-5} to prevent the pre-trained weights from being lost in the first epochs. This process was repeated using five and ten MSKCC image sets to assess the impact of retraining dataset size.

2.5. Testing networks

2.5.1. Evaluation metrics to compare network performances

The UMCU images and GT segmentations of the ten remaining patients (UMCU test set) were used to evaluate both networks. Evaluation metrics for the lymph node levels included the Dice Similarity Coefficient (DSC) and the 95th percentile Hausdorff distance in mm (HD95) [36]. For individual lymph nodes, additional evaluation metrics were recall (i.e., sensitivity or true positive rate), precision (i.e., positive predictive value), and F1-score (combination of recall and precision), which all were measured per lymph node [37]. HD95 was omitted to evaluate individual lymph node segmentations since false positive and false negative predictions disproportionally increased this metric.

True positive lymph nodes were defined as isolated predictions that intersected with GT segmentations, while false positive lymph nodes were defined as isolated predictions that did not cross GT segmentations. In the case of individual elective lymph node irradiation, obtaining a network with a high recall score is especially important so that no lymph nodes will be missed during treatment. DSC for individual lymph nodes was only evaluated on true positive predictions, as false positive lymph nodes disproportionally increased this metric.

To evaluate the performance of the CNN in MSKCC, baseline results were first produced by applying the UMCU CNN directly to ten MSKCC images (MSKCC test set) without applying transfer learning. After the CNNs were re-trained with five and ten MSKCC patients, the models were tested on the MSKCC test set.

2.5.2. Observer performance of individual lymph nodes segmentations

To determine the additional value of the CNN in the clinic, the performance of UMCU observers with manual segmentations was compared to the output of the CNN. To examine the performance of different observers, the individual lymph nodes on one side of the neck of ten patients (UMCU test set) were independently re-segmented by two observers (MR[1] and PD[1]). The segmentations of the two observers and the CNN predictions were compared with the UMCU GT segmentations using the same evaluation metrics described in 2.6.1.

Additionally, an interobserver analysis was performed by comparing the DSC scores from segmentations of observer 1 and 2 with the DSC scores from nnU-Net and the ground truth using a Wilcoxon matched-pair signed rank test.

2.6. Statistical analysis

Due to the small sample size, non-parametric testing was applied. Therefore, descriptive evaluation variables were reported as median with an interquartile range (IQR). Interobserver differences were analyzed with the Friedman test. The UMCU and MSKCC evaluation scores were compared with the Mann–Whitney *U* test. All statistical testing was performed with Statistical Package for Social Sciences (SPSS) version 25, IBM, New York, USA. The alpha level of statistical significance was set at 0.05. The Bonferroni correction was applied for multiple tests on the two metric scores across all five lymph node levels (alpha level = 0.05/10 = 0.005) and the four metric scores for lymph nodes (alpha level = 0.05/4 = 0.0125).

[1] Observers in the UMC-Utrecht: FR= Floris C.J. Reinders, MR= Mischa de Ridder, PD= Patricia A.H. Doornaert, CT= Chris H.J. Terhaard.

[2] Observer in MSKCC: KZ = Kaveh Zakeri.

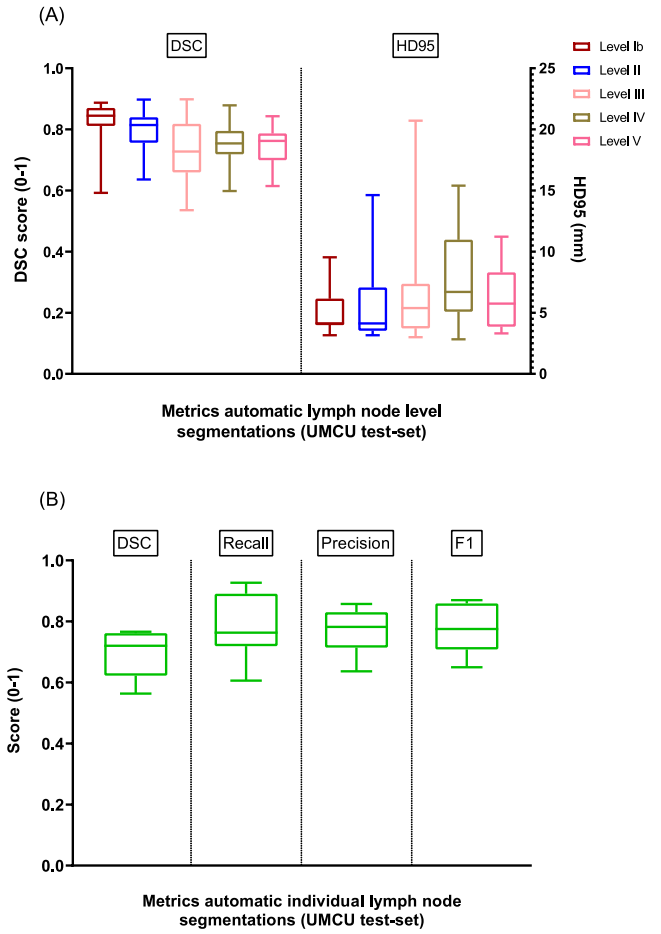

Fig. 3. Performance of automatic segmentations of and lymph node levels Ib-V (A) and individual lymph nodes (B) produced by a trained neural network (nnU-Net) for 10 HNSCC patients (UMCU test-set). The performance metrics for the lymph node levels included DSC and Hausdorff Distance 95th percentile (HD95), for the individual lymph nodes this was dice (DSC), Recall, Precision, and F1-score. Boxes represent the first, second, and third quartile whereas the whiskers represent the 95 percentile of all values. The dots represent outliers above or below the 2.5–97.5 percentile.

3. Results

3.1. Ground truth segmentations UMCU and MSKCC database

In total, 2851 and 890 lymph nodes were segmented in UMCU and MSKCC database images. The median number of segmented individual lymph nodes per patient was 56 (IQR = 46–68) in the UMCU database and 43 (IQR = 32–53) in the MSKCC database. After post-processing, lymph nodes or groups of lymph nodes had a mean volume of 0.16 cm^3 (IQR = 0.09–0.33) in the UMCU database and 0.17 cm^3 (IQR = 0.01–0.32) in the MSKCC database.

3.2. Automatic lymph node level segmentations UMCU

Lymph node levels Ib-II-III-IV-V were predicted by nnU-Net with a median DSC score of 0.84/0.81/0.73/0.75/0.76 and HD95 of 4.1/4.1/5.4/6.7/5.7 mm (Fig. 2/3A and Supplementary Table 4).

3.3. Automatic individual lymph node segmentation UMCU

Individual lymph node segmentations in the UMCU dataset using nnU-Net resulted a median DSC score of 0.72, recall of 0.76, precision of 0.78, and F1-Score of 0.78 (Fig. 2/3B and Supplementary Table 5). Fig. 4 shows a comparison of ground truth labels with predictions of lymph node levels and elective lymph nodes in one patient.

3.4. Transfer learning of the CNN

The nnU-Net trained with UMCU training set generated lower accuracies on the MSKCC test set than on the UMCU test set regarding DSC (0.57 vs. 0.72, p = 0.002) and recall (0.56 vs. 0.76, p = 0.007). No differences were found for precision (0.79 vs. 0.78, p = 0.837) and F1 score (0.67 vs. 0.78, p = 0.019). Re-training nnU-Net with 10 patients from the MSKCC re-training set improved segmentation results such that no differences were found between the MSKCC and UMCU test set for all evaluation scores (Fig. 5/Supplementary Table 6). Although not statistically significant, there was a trend of a lower DSC, recall, and F1-score and higher precision for the segmentations obtained with the adapted nnU-Net (5 or 10 patients in the training set) compared to the UMCU test results. The size of the re-training set (5 or 10 patients) did not significantly impact the results of DSC (0.64 vs. 0.64, p = 0.796), recall (0.62

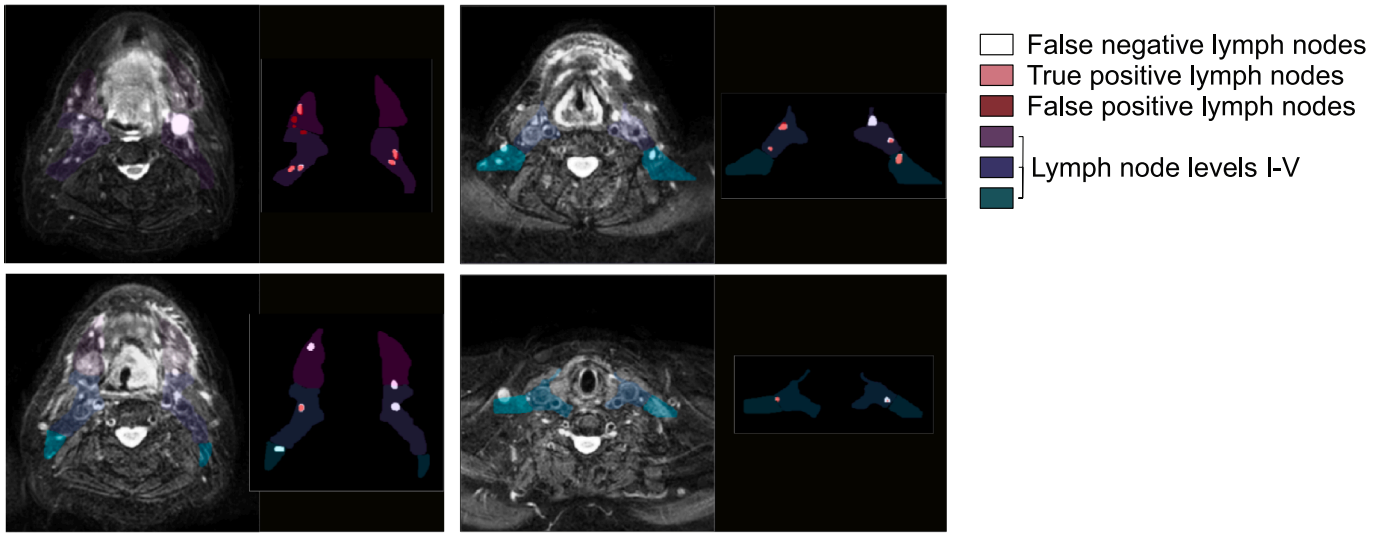

Fig. 4. Ground truth labels (white) and automatic nnU-Net predictions (red) of lymph nodes on consecutive transversal MRI slices from one patient showing true positives voxels (pink), false positives voxels (red), and false negatives voxels (white). Ground truth segmentations of lymph node levels I-V are displayed as an overlay mask (blue/purple/turquoise).

vs. 0.64, p = 0.493), precision (0.85 vs. 0.85, p = 0.853), and F1-scores (0.78 vs. 0.82, p = 0.592).

In the case of lymph node levels, similar results were obtained, with no differences in evaluation scores (DSC/HD95) for all levels (except level V) between the UMCU and MSKCC test set after re-training the CNN with 5 or 10 MSKCC patients (Supplementary file 2).

3.5. Observers vs. CNN performance of individual lymph node segmentations

In the UMCU test set, nnU-Net produced segmentations with higher recall (p = 0.002) and F1 score (p = 0.012) compared to two independent radiation oncologists. No differences were found regarding DSC (p = 0.018) and precision (p = 0.138) (Fig. 6/Supplementary Table 7). The

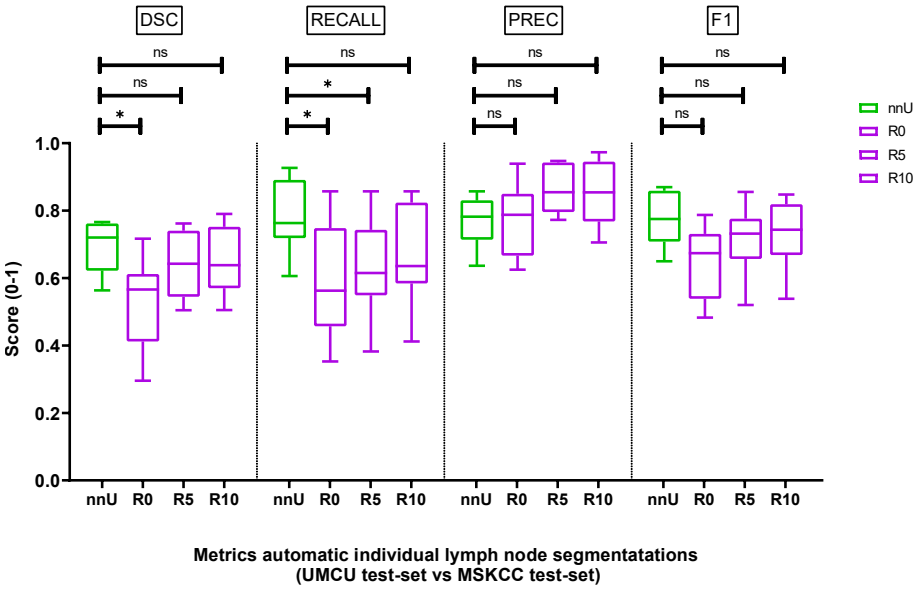

Fig. 5. Performance automatic segmentations of individual lymph nodes produced by a trained neural network (nnU-Net) in 10 HNSCC patients of the UMCU test-set (green) and the MSKCC test-set (purple) regarding dice (DSC), Recall, Precision, and F1 score. The evaluation on the MSKCC test-set was done with the same network without re-training (R0), re-training with an additional five (R5) and ten patients (R10). Boxes represent the first, second, and third quartiles, whereas the whiskers represent the 95 percentile of all values. The dots represent outliers above or below the 2.5–97.5 percentile. Significant p-values of testing evaluation scores are marked with * (p < 0.0125).

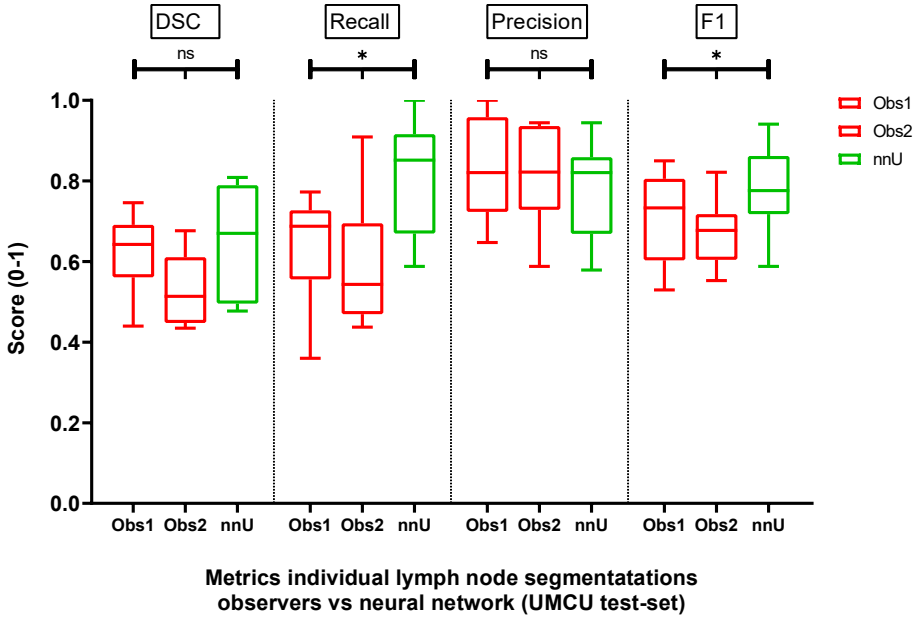

Fig. 6. Performance of individual lymph node segmentations produced by nnU-Net compared to two independent radiation oncologists regarding dice (DSC), Recall, Precision, and F1 score. The nnU-Net (nnU) scores are in green, and the observer's (Obs1/2) scores are in red. Boxes represent the first, second, and third quartiles, whereas the whiskers represent the 95 percentile of all values. The dots represent outliers above or below the 2.5–97.5 percentile. Significant p-values of testing evaluation scores are marked with * (p < 0.0125).

comparison of the segmentations of observer 1 and observer 2 resulted in a median DSC of 0.57 which was lower but not significantly different from nnU-Net segmentations (Supplementary Table 7). The number of segmented lymph nodes per patient varied across segmentations of the ground truth, observer 1/2, and nnU-Net with differences up to 12 lymph nodes (Supplementary Table 8).

4. Discussion

This study found that using a 3D CNN (nnU-Net) for automatic individual lymph node segmentation in HNSCC patients on MRI led to better recall and F1-scores than manual segmentation, supporting individual elective lymph node irradiation with potential OAR sparing. However, similar results could not be achieved in another medical center without retraining the CNN.

DSC scores of lymph node level segmentations aligned with the limited literature on MRI-based automatic segmentations of these structures [38,39]. Prior MRI studies focused on lymph node levels II and III, while this study included levels Ib-V.

While some studies examined CT-based neural networks to distinguish between pathological and benign cervical lymph nodes in HNSCC patients [40–45], only one developed [46] a CT-based CNN for automatic individual lymph node detection and segmentation. That model had similar recall (0.77 vs. 0.76), but higher precision (0.98 vs. 0.78) and F1 score (0.86 vs. 0.78). However, our metrics may have been negatively affected since MRI typically detects more small lymph nodes than CT [14], which are harder to predict by a CNN [47]. Additionally, our MRI had a larger slice thickness (3/5/6 vs. 2 mm), providing less contextual information during training. For the purpose of irradiating elective lymph nodes it is of utmost importance to detect as many lymph nodes as possible which is comparable with this other study.

In this study, a median of 56 lymph nodes (IQR = 46–68) were segmented on MRI, lower than the 34–46 lymph nodes found in several pathology studies [48–51]. However, we excluded lymph nodes only visible in one transverse MRI slice, and small lymph nodes may have been missed due to 3 mm slice thickness.

After re-training the CNN with only 10 patients, no distinct differences were found between evaluation metrics of individual lymph nodes and most lymph node level segmentations of the MSKCC and UMCU test set. A small re-training set of 5–10 patients can significantly improve segmentation results, enabling automatic lymph node detection in medical centers worldwide with limited recourses.

Individual lymph node segmentations from the re-trained CNN on the MSKCC images showed a trend of lower DSC, recall, and F1-score. Moreover, level-V segmentations showed worse DSC and HD95. One reason for lower performance scores may be that four patients in the MSKCC database had post-radiotherapy images, resulting in smaller, less visible lymph nodes, which were included due to data scarcity. Additionally, the larger slice thickness in the MSKCC dataset may have hindered the detection of small lymph nodes. The poorer performance in predicting level V in the MSKCC test set might be caused by the different scanning setups using a head coil instead of flex coils without immobilization mask, which might have resulted in lower signal quality in the lower neck. Since increasing the re-training set from 5 to 10 patients did not significantly improve results, aligning scanning protocols between medical centers should be the first focus in enhancing the outcomes for MSKCC use of automatic individual lymph node segmentation. Nonetheless, a larger re-training set would be beneficial, and repeating the experiment with more patients would be worthwhile.

Comparing independent observers' and CNN segmentations revealed no significant difference in DSC and precision, but the CNN had higher recall and F1-scores. Radiation oncologists might have missed lymph nodes due to time constraints, while the CNN's higher detection rate ensures more thorough assessments. Future studies should explore its potential to improve efficiency and reduce oncologists' workload.

The interobserver analysis of lymph node segmentations showed a trend of lower DSC scores between observer 1 and 2 compared to nnU-Net, indicating difficulties in identifying correct structures. This likely caused variations in the number of segmented lymph nodes among the observers, nnU-Net, and the ground truth, particularly with smaller lymph nodes. Since small elective lymph nodes (<4 mm) pose a low risk of occult metastasis but have high selection variability, they are excluded in the feasibility study for individual elective lymph node irradiation.

Our study has several limitations. The most important is the relatively small dataset used to train the CNNs, still the UMCU dataset includes over 2800 lymph nodes. Another limitation is that the GT was defined by two experts, allowing for interobserver variation. Additionally, the scan quality of the multi-slice T2 weighted mDixon TSE scans was suboptimal in the cranial and caudal regions, making it difficult to distinguish small lymph nodes from other structures, particularly in levels IV and V. Improving MRI quality might decrease interobserver variability of individual lymph node segmentations and enhance CNN performance.

Future work will explore the influence of automatic lymph node segmentations on the dosimetric evaluation of our new treatment concept. Also, observing the intra-observer variation of the primary observer (FR) could provide insights into the challenges of segmenting individual lymph nodes. For transfer learning, future studies could incrementally increase the training group by one patient to determine the number needed for acceptable performance. A bootstrap process using multiple random re-training sets could also yield more robust results.

The CNN developed in this study is currently being used in the feasibility study, where the first patients are currently being treated with individual elective lymph node irradiation on the MR-linac [52].

This is the first study describing the automatic segmentation of individual lymph nodes in HNSCC patients on MRI using state-of-the-art CNNs. Notably, nnU-Net had a higher detection rate of lymph nodes than manual segmentations, supporting the implementation of individual elective lymph node irradiation with improved sparing of OARs. Retraining was required when applying the model to images from another scanner with a different acquisition protocol, highlighting a common challenge in automatic segmentation on MRI due to varying acquisition parameters.

Declaration of interest

This study was possible due to the sponsored international fellowship of the first author by the following unrestricted grants: 1. Stichting Hanarth Fonds. 2. Prins Bernard Cultuur fonds, 3. Stichting de drie lichten, 4. Hendrik Muller Fonds, 5. Girard de Mielet van Coehoorn Stichting.

Funding statement

The research performed in MSKCC was partly sponsored by the following unrestricted grants: - Prins Bernard Cultuurfonds, grant ID 40042552 - Stichting De Drie Lichten, grant ID 08 2022 - Hendrik Muller fonds, grant ID N.A. - Hanarth Fonds, grant ID N.A. - Girard de Mielet van Coehoorn Stichting, grant ID, N.A.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The trained nnU-Net models for individual elective lymph nodes and lymph node levels are available upon request for research groups

investigating similar topics.

Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2024.100655.

References

[1] Christianen MEMC, Verdonck-De Leeuw IM, Doornaert P, Chouvalova O, Steenbakkers RJHM, Koken PW, et al. Patterns of long-term swallowing dysfunction after definitive radiotherapy or chemoradiation. Radiother Oncol 2015;117:139–44. https://doi.org/10.1016/j.radonc.2015.07.042.

[2] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol 2011;12:127–36. https://doi.org/10.1016/S1470-2045(10)70290-4.

[3] Boomsma MJ, Bijl HP, Langendijk JA. Radiation-induced hypothyroidism in head and neck cancer patients: a systematic review. Radiother Oncol 2011;99:1–5. https://doi.org/10.1016/j.radonc.2011.03.002.

[4] Wilbers J, Dorresteijn LD, Haast R, Hoebers FJ, Kaanders JH, Boogerd W, et al. Progression of carotid intima media thickness after radiotherapy: A long-term prospective cohort study. Radiother Oncol 2014;113:359–63. https://doi.org/10.1016/j.radonc.2014.10.012.

[5] Van Den Bosch S, Dijkema T, Verhoef LCG, Zwijnenburg EM, Janssens GO, Kaanders JHAM. Patterns of recurrence in electively irradiated lymph node regions after definitive accelerated intensity modulated radiation therapy for head and neck squamous cell Carcinoma. Int J Radiat Oncol Biol Phys 2016;94:766–74. https://doi.org/10.1016/j.ijrobp.2015.12.002.

[6] Leeman JE, Li JG, Pei X, Venigalla P, Zumsteg ZS, Katsoulakis E, et al. Patterns of treatment failure and postrecurrence outcomes among patients with locally advanced head and neck squamous cell carcinoma after chemoradiotherapy using modern radiation techniques. JAMA Oncol 2017;3:1487–94. https://doi.org/10.1001/jamaoncol.2017.0973.

[7] Kaanders JHAM, van den Bosch S, Dijkema T, Al-Mamgani A, Raaijmakers CPJ, Vogel WV. Advances in cancer imaging require renewed radiotherapy dose and target volume concepts. Radiother Oncol 2020;148:140–2. https://doi.org/10.1016/j.radonc.2020.04.016.

[8] Sher DJ, Pham NL, Shah JL, Sen N, Williams KA, Subramaniam RM, et al. Prospective phase 2 study of radiation therapy dose and volume de-escalation for elective neck treatment of oropharyngeal and laryngeal cancer. Int J Radiat Oncol Biol Phys 2021;109:932–40. https://doi.org/10.1016/j.ijrobp.2020.09.063.

[9] Deschuymer S, Nevens D, Duprez F, Daisne J-F, Dok R, Laenen A, et al. Randomized clinical trial on reduction of radiotherapy dose to the elective neck in head and neck squamous cell carcinoma; update of the long-term tumor outcome. Radiother Oncol 2020;143:24–9. https://doi.org/10.1016/j.radonc.2020.01.005.

[10] Maguire PD, Neal CR, Hardy SM, Schreiber AM. Single-arm phase 2 trial of elective nodal dose reduction for patients with locoregionally advanced squamous cell carcinoma of the head and neck. Int J Radiat Oncol Biol Phys 2018;100:1210–6. https://doi.org/10.1016/j.ijrobp.2017.12.277.

[11] Tsai CJ, McBride SM, Riaz N, Kang JJ, Spielsinger DJ, Waldenberg T, et al. Evaluation of substantial reduction in elective radiotherapy dose and field in patients with human papillomavirus-associated oropharyngeal carcinoma treated with definitive chemoradiotherapy. JAMA Oncol 2022;8:364–72. https://doi.org/10.1001/jamaoncol.2021.6416.

[12] Suwinski R, Maciejewski B. Dose-response relationship for elective neck irradiation of head and neck cancer–facts and controversies. Neoplasma 1998;45(2):107–12. PMID: 9687892.

[13] Reinders FCJ, Stijnman PRS, De RM, Doornaert PAH, Raaijmakers CPJ, Philippens MEP. MRI visibility and displacement of elective lymph nodes during radiotherapy in head and neck cancer patients. Front Radiol 2022;2:1033521. https://doi.org/10.3389/fradi.2022.1033521.

[14] Reinders FCJ, Heijst TCF, va Mases J, Terhaard CHJ, Doornaert PAH, Philippens MEP, et al. Magnetic resonance guided elective neck irradiation targeting individual lymph nodes: A new concept. Phys Imaging Radiat Oncol 2021;20:76–81. https://doi.org/10.1016/j.phro.2021.10.006.

[15] de Ridder M, Raaijmakers CPJ, Pameijer FA, de Bree R, Reinders FCJ, Doornaert PAH, et al. Target definition in MR-guided adaptive radiotherapy for head and neck cancer. Cancers 2022;14(12):3027. https://doi.org/10.3390/cancers14123027.

[16] Harari PM, Song S, Tomé WA. Emphasizing conformal avoidance versus target definition for IMRT planning in head-and-neck cancer. Int J Radiat Oncol Biol Phys 2010;77:950–8. https://doi.org/10.1016/j.ijrobp.2009.09.062.

[17] Kosmin M, Ledsam J, Romera-Paredes B, Mendes R, Moinuddin S, de Souza D, et al. Rapid advances in auto-segmentation of organs at risk and target volumes in head and neck cancer. Radiother Oncol 2019;135:130–40. https://doi.org/10.1016/j.radonc.2019.03.004.

[18] Mannion-Haworth R, Bowes M, Ashman A, Guillard G, Brett A, Vincent G. Fully automatic segmentation of head and neck organs using active appearance models. Midas J 2016. https://doi.org/10.54294/e86siq.

[19] Anbeek P, Vincken KL, Van Bochove GS, Van Osch MJP, Van Der Grond J. Probabilistic segmentation of brain tissue in MR imaging. Neuroimage 2005;27:795–804. https://doi.org/10.1016/j.neuroimage.2005.05.046.

[20] Han X, Hoogeman M, Levendag P, Hibbard L, Teguh D, Voet P, et al. Atlas-based auto-segmentation of head and neck CT images. Med Image Comput Comput Assist Interv 2008;11:434–41. https://doi.org/10.1007/978-3-540-85990-1_52.

[21] Daisne JF, Blumhofer A. Atlas-based automatic segmentation of head and neck organs at risk and nodal target volumes: a clinical validation. Radiat Oncol 2013;8:1–11. https://doi.org/10.1186/1748-717X-8-154.

[22] Iglesias JE, Sabuncu MR. Multi-atlas segmentation of biomedical images: a survey. Med Image Anal 2015;24:205–19. https://doi.org/10.1016/j.media.2015.06.012.

[23] Chittajallu DR, Shah SK, Kakadiaris IA. A shape-driven MRF model for the segmentation of organs in medical images. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 2010:3233–40. https://doi.org/10.1109/CVPR.2010.5540066.

[24] Yingzi L, Yang L, Yabo F, Tonghe W, Jun Z, Jiang X, et al. Head and neck multi-organ auto-segmentation on CT images aided by synthetic MRI. Med Phys 2020;47(9):4294–302. https://doi.org/10.1002/mp.14378.

[25] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12- June-2015, IEEE Computer Society; 2015, p. 431–40. doi: 10.1109/CVPR.2015.7298965.

[26] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. arXiv.1505.04597 2015. doi: 10.48550/arXiv.1505.04597.

[27] van Dijk LV, Van den Bosch L, Aljabar P, Peressutti D, Both S, Steenbakkers Roel JHM, et al. Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring. Radiother Oncol 2020;142:115–23. https://doi.org/10.1016/j.radonc.2019.09.022.

[28] van der Veen J, Willems S, Bollen H, Maes F, Nuyts S. Deep learning for elective neck delineation: More consistent and time efficient. Radiother Oncol 2020;153:180–8. https://doi.org/10.1016/j.radonc.2020.10.007.

[29] Schouten JPE, Noteboom S, Martens RM, Mes SW, Leemans CR, de Graaf P, et al. Automatic segmentation of head and neck primary tumors on MRI using a multi-view CNN. Cancer Imaging 2022;22:1–9. https://doi.org/10.1186/s40644-022-00445-7.

[30] Rodríguez Outeiral R, Bos P, Al-Mamgani A, Jasperse B, Simões R, van der Heide UA. Oropharyngeal primary tumor segmentation for radiotherapy planning on magnetic resonance imaging using deep learning. Phys Imaging Radiat Oncol 2021;19:39–44. https://doi.org/10.1016/j.phro.2021.06.005.

[31] Ren J, Eriksen JG, Nijkamp J, Korreman SS. Comparing different CT, PET and MRI multi-modality image combinations for deep learning-based head and neck tumor segmentation. Acta Oncol (Madr) 2021;60:1399–406. https://doi.org/10.1080/0284186X.2021.1949034.

[32] Wahid KA, Ahmed S, He R, van Dijk LV, Teuwen J, McDonald BA, et al. Evaluation of deep learning-based multiparametric MRI oropharyngeal primary tumor auto-segmentation and investigation of input channel effects: Results from a prospective imaging registry. Clin Transl Radiat Oncol 2022;32:6–14. https://doi.org/10.1016/j.ctro.2021.10.003.

[33] Glover GH. Multipoint dixon technique for water and fat proton and susceptibility imaging. J Magn Reson Imaging 1991;1:521–30. https://doi.org/10.1002/jmri.1880010504.

[34] Grégoire V, Ang K, Budach W, Grau C, Hamoir M, Langendijk JA, et al. Delineation of the neck node levels for head and neck tumors: A 2013 update. DAHANCA, EORTC, HKNPCSG, NCIC CTG, NCRI, RTOG, TROG consensus guidelines. Radiother Oncol 2014;110:172–81. doi: 10.1016/j.radonc.2013.10.010.

[35] Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. 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.

[36] Lin H, Xiao H, Dong L, Teo KBK, Zou W, Cai J, et al. Deep learning for automatic target volume segmentation in radiation therapy: a review. Quant Imaging Med Surg 2021;11:4847–58. https://doi.org/10.21037/qims-21-168.

[37] Malhotra P, Gupta S, Koundal D, Zaguia A, Enbeyle W. Deep neural networks for medical image segmentation. J Healthc Eng 2022;2022. https://doi.org/10.1155/2022/9580991.

[38] Kawahara D, Tsuneda M, Ozawa S, Okamoto H, Nakamura M, Nishio T, et al. Deep learning-based auto segmentation using generative adversarial network on magnetic resonance images obtained for head and neck cancer patients. J Appl Clin Med Phys 2022;23:1–13. https://doi.org/10.1002/acm2.13579.

[39] Korte JC, Hardcastle N, Ng SP, Clark B, Kron T, Jackson P. Cascaded deep learning-based auto-segmentation for head and neck cancer patients: organs at risk on T2-weighted magnetic resonance imaging. Med Phys 2021;48:7757–72. https://doi.org/10.1002/mp.15290.

[40] Bardosi ZR, Dejaco D, Santer M, Kloppenburg M, Mangesius S, Widmann G, et al. Benchmarking eliminative radiomic feature selection for head and neck lymph node classification. Cancers 2022;14(3):477. https://doi.org/10.3390/cancers14030477.

[41] Onoue K, Fujima N, Andreu-Arasa VC, Setty BN, Sakai O. Cystic cervical lymph nodes of papillary thyroid carcinoma, tuberculosis and human papillomavirus positive oropharyngeal squamous cell carcinoma: utility of deep learning in their differentiation on CT. Am J Otolaryngol – Head Neck Med Surg 2021;42:103026. https://doi.org/10.1016/j.amjoto.2021.103026.

[42] Tomita H, Yamashiro T, Heianna J, Nakasone T, Kimura Y, Mimura H, et al. Nodal-based radiomics analysis for identifying cervical lymph node metastasis at levels I and II in patients with oral squamous cell carcinoma using contrast-enhanced computed tomography. Eur Radiol 2021;31:7440–9. https://doi.org/10.1007/s00330-021-07758-4.

[43] Kann BH, Aneja S, Loganadane GV, Kelly JR, Smith SM, Decker RH, et al. Pretreatment identification of head and neck cancer nodal metastasis and extranodal extension using deep learning neural networks. Sci Rep 2018;8:1–11. https://doi.org/10.1038/s41598-018-32441-y.

[44] Seidler M, Forghani B, Reinhold C, Pérez-Lara A, Romero-Sanchez G, Muthukrishnan N, et al. Dual-energy CT texture analysis with machine learning for the evaluation and characterization of cervical lymphadenopathy. Comput Struct Biotechnol J 2019;17:1009–15. https://doi.org/10.1016/j.csbj.2019.07.004.

[45] Ariji Y, Fukuda M, Kise Y, Nozawa M, Yanashita Y, Fujita H, et al. Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using a deep learning system of artificial intelligence. Oral Surg Oral Med Oral Pathol Oral Radiol 2019;127:458–63. https://doi.org/10.1016/j.oooo.2018.10.002.

[46] Ariji Y, Kise Y, Fukuda M, Kuwada C, Ariji E. Segmentation of metastatic cervical lymph nodes from CT images of oral cancers using deep-learning technology. Dentomaxillofac Radiol 2022;51:20210515. https://doi.org/10.1259/dmfr.20210515.

[47] Zhao X, Xie P, Wang M, Li W, Pickhardt PJ, Xia W, et al. Deep learning–based fully automated detection and segmentation of lymph nodes on multiparametric-mri for rectal cancer: A multicentre study. EBioMedicine 2020;56:102780. https://doi.org/10.1016/j.ebiom.2020.102780.

[48] Nurimba M, Hines W, Sinha U, Mathew A, Kokot N, Swanson M. Evaluation of lymph node ratio and lymph node yield as prognosticators of locoregional recurrence in p16-associated oropharyngeal squamous cell carcinoma. Head Neck 2020;42:2811–20. https://doi.org/10.1002/hed.26324.

[49] Sheppard SC, Frech L, Giger R, Nisa L. Lymph node yield and ratio in selective and modified radical neck dissection in head and neck cancer—impact on oncological outcome. Cancers 2021;13(9):2205. https://doi.org/10.3390/cancers13092205.

[50] Marres CCM, De Ridder M, Hegger I, Van Velthuysen MLF, Hauptmann M, Navran A, et al. The influence of nodal yield in neck dissections on lymph node ratio in head and neck cancer. Oral Oncol 2014;50:59–64. https://doi.org/10.1016/j.oraloncology.2013.09.014.

[51] Iocca O, Di Maio P, De Virgilio A, Pellini R, Golusiński P, Petruzzi G, et al. Lymph node yield and lymph node ratio in oral cavity and oropharyngeal carcinoma: preliminary results from a prospective, multicenter, international cohort. Oral Oncol 2020;107:104740. https://doi.org/10.1016/j.oraloncology.2020.104740.

[52] Reinders FCJ, de Ridder M, Doornaert PAH, Raaijmakers CPJ, Philippens MEP. Individual elective lymph node irradiation for the reduction of complications in head and neck cancer patients (iNode): A phase-I feasibility trial protocol. Clin Transl Radiat Oncol 2022;39:100574. https://doi.org/10.1016/j.ctro.2022.100574.