MÆDICA - a Journal of Clinical Medicine

ORIGINAL PAPER

MAEDICA – a Journal of Clinical Medicine
2026; 21(2): 320-330
https://doi.org/10.26574/maedica.2026.21.2.320

Deep Learning-Based Detection of Malignant and Equivocal Cervical Lymph Nodes on CT Imaging Using a 2D U-Net++ Model

Iulian-Alexandru TACIUC[a], Mihai DUMITRU[b], Daniela VRINCEANU[b], Andreea MARINESCU[c], Adina Zamfir Chiru ANTON[d], Adrian COSTACHE[a]

[a]"Carol Davila" University of Medicine and Pharmacy, Pathology Department, Bucharest, Romania

[b]"Carol Davila" University of Medicine and Pharmacy, ENT Department, Bucharest, Romania

[c]"Carol Davila" University of Medicine and Pharmacy, Radiology Department, Bucharest, Romania

[d]ENT Department, "Grigore Alexandrescu" Children Hospital, Bucharest, Romania

ABSTRACT

Introduction: *Accurate identification of malignant and equivocal lymph nodes on computed tomography (CT) remains a significant challenge in clinical practice, particularly in borderline cases. Recent advances in artificial intelligence offer potential solutions for improving diagnostic performance and supporting radiological decision-making.*

Materials and methods: *A dataset of 79 CT examinations acquired between January 2024 and December 2025 was retrospectively analyzed, including 538 annotated lymph nodes classified as Node-RADS 3–5. A 2D U-Net++ convolutional neural network was implemented and trained in a slice-wise manner using NIfTI-formatted data. The dataset was divided into training and validation subsets, with data augmentation being applied through random rotations. Model performance was evaluated using Dice Similarity Coefficient (DSC), Intersection over Union (IoU), sensitivity, specificity, precision and accuracy. An independent test set of 12 CT cases was used for qualitative assessment.*

Results: *The model achieved an accuracy of 99.8% on the training set and 99.7% on validation, with specificity consistently reaching 99.9%. Sensitivity was 78% for training and 72% for validation, while precision reached 83% and 79%, respectively. Segmentation performance yielded a DSC of 0.79 (training) and 0.73 (validation), with IoU values of 0.67 and 0.61. The model demonstrated robust anatomical discrimination, correctly excluding common mimics such as accessory spleens, salivary glands, vascular structures and muscular tissue. Qualitative evaluation showed limited false-negative results (three equivocal lymph nodes) and two false-positive detections corresponding to benign (Node-RADS 1) lymph nodes.*

Address for correspondence:
Mihai Dumitru
Email: orldumitrumihai@yahoo.com

Article received on the 15th of May 2026 and accepted for publication on the 20th of May 2026

Conclusions: *The proposed 2D U-Net++ model demonstrates reliable performance in the detection of malignant and equivocal lymph nodes on CT imaging, with high specificity and good anatomical discrimination. Its applicability to both contrast-enhanced and non-contrast CT scans supports its potential as a clinical decision-support tool, particularly for highlighting borderline findings in routine practice.*

Keywords: artificial intelligence, computed tomography, deep learning, head and neck cancer, lymph nodes, medical image segmentation, node-RADS, u-net++.

INTRODUCTION

Head and neck cancers represent a heterogeneous group of malignancies, with squamous cell carcinoma being the most frequent histological subtype (1, 2). Cervical lymph node metastasis is a major prognostic factor, significantly affecting staging, treatment planning, and survival. Previous studies have shown that five-year survival may decrease to approximately 50% in patients with unilateral cervical nodal involvement and to nearly 25% in those with bilateral metastases (3, 4).

Given the complex anatomy of the head and neck region and the high density of lymphatic structures, clinical examination alone is insufficient for reliable nodal assessment. Imaging plays a central role, with computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound (US) being the most commonly used modalities, while positron emission tomography (PET) may provide additional metabolic information, despite more limited availability (5-8). Cervical lymph nodes are classified according to a level-based system, which supports standardized staging and correlates with expected drainage pathways from the primary tumor site (5).

Standardized reporting is essential to improve reproducibility across clinicians and institutions. In this context, the Node Reporting and Data System (Node-RADS) has been proposed as a structured framework for lymph node assessment on CT and MRI, using a five-point scale from very low to very high likelihood of malignancy (9, 10). Recent meta-analytic evidence supports the diagnostic value of Node-RADS for standardized reporting and risk stratification, although further work is needed to establish inter-reader reproducibility (11). Node-RADS incorporates multiple malignancy-related imaging features, including short-axis diameter, rounded morphology, irregular or ill-defined margins, necrosis, heterogeneous enhancement, loss of fatty hilum, nodal clustering, drainage pathway location and clinical context (10, 12). Among these features, nodal clustering was not included in the examples used for network training in the present study, which may limit the model's ability to identify this specific criterion.

Artificial intelligence (AI) has gained increasing relevance in otorhinolaryngology, particularly for risk stratification, diagnostic support and treatment optimization (13). In imaging, AI-based tools may assist workflow management, image classification, segmentation and quantitative analysis, potentially reducing observer variability and improving diagnostic consistency (14). For CT-based detection and segmentation tasks, convolutional neural networks (CNNs) are among the most widely used architectures because of their ability to process spatial image information and recognize complex imaging patterns. In lymph node assessment, AI models may be trained to identify features suggestive of malignancy, including size, shape, border irregularity, necrosis and extranodal extension, thereby supporting earlier detection and more accurate staging (13, 15).

The present study aims to adapt and optimize an existing AI model for the detection of malignant or equivocal cervical lymph nodes on CT imaging. This includes fine-tuning training parameters and using a dedicated CT dataset for model training and validation. The main objective is to develop a supportive tool capable of rapidly highlighting suspicious lymph nodes or "red flag" imaging features for radiologist review.

The proposed system is intended to assist, not replace, radiological interpretation, with final diagnosis remaining dependent on clinical correlation and expert assessment. Model robustness and specificity were further evaluated by testing performance against anatomical structures that may mimic lymph nodes and against lymph nodes

from other body regions, in order to assess potential false-positive and false-negative sources. Additionally, lymph nodes were annotated on both contrast-enhanced venous-phase and non-contrast CT images, aiming to maintain model applicability in clinical situations where intravenous contrast administration is contraindicated. ❑

MATERIAL AND METHODS

Dataset

The dataset consists of 79 CT examinations from 79 different patients (67 standard cases and 12 additional cases included for specific testing purposes), which were collected between January 2024 and December 2025. All scans were acquired within the same healthcare institution using four different CT devices, ensuring variability in acquisition parameters while maintaining institutional consistency.

The cases used for training and validation included the 67 standard cases, comprising 40 males and 27 females, with ages ranging from 30 to 83 years (mean age: 62.8 years). The dataset encompassed a spectrum of histologically confirmed head and neck malignancies, including malignant parotid tumors, keratinizing and non-keratinizing squamous cell carcinoma, papillary carcinoma arising in ectopic thyroid tissue, spinocellular carcinoma, non-Hodgkin lymphoma, paraganglioma and anaplastic thyroid carcinoma.

From the 67 patients, a total of 538 lymph nodes were annotated. The dataset included lymph nodes from all seven anatomical cervical levels; however, the model was not explicitly trained to classify nodal levels. Efforts were made to include a wide range of nodal imaging characteristics, while clustered lymph nodes were excluded to preserve labeling clarity and consistency. This generated 1582 axial CT slices with lymph nodes.

Cases were included in the study based on the following criteria: histologically confirmed head and neck malignancy (either via biopsy or surgical resection); availability of a CT examination including at least the head and neck regions; presence of at least equivocal or suspicious lymph nodes; availability of a complete radiological report.

Lymph nodes were included in the study based on the Node-RADS 1.0 system: a. Node-RADS 3 (equivocal); b. Node-RADS 4 (high likelihood of malignancy); and c. Node-RADS 5 (very high likelihood of malignancy).

The following exclusion criteria were applied: CT scans with severe artifacts significantly affecting image quality; lymph nodes classified as Node-RADS 1 or 2; cases without a confirmed primary tumor.

The study intentionally included only lymph nodes classified as Node-RADS 3–5, corresponding to equivocal, suspicious, or highly suspicious findings. Node-RADS 1–2 lymph nodes were excluded because the aim of the model was not to perform comprehensive lymph node classification across the full spectrum of nodal appearances, but rather to detect and segment lymph nodes requiring increased radiological attention. Therefore, the proposed model should be interpreted as a suspicious-node detection and segmentation tool, rather than a universal lymph node classifier.

Image annotation and preprocessing

Computed tomography image annotation was performed using 3D Slicer software (version 5.10.0). Each lymph node was manually segmented in all three anatomical planes (axial, coronal and sagittal) by a nuclear medicine physician, ensuring precise volumetric labeling and consistency. The generated masks were subsequently reviewed with experienced radiologists, who verified lymph node identification and anatomical plausibility of the segmentations. Thus, the final masks were created by one annotator and verified through radiological review.

The annotated datasets were stored in NIfTI (.nii) format, which is commonly used for three-dimensional medical imaging data. In contrast to standard 2D image formats (.jpg/.png), NIfTI preserves volumetric information, spatial orientation and voxel spacing, enabling accurate alignment between CT volumes and their corresponding segmentation masks. The dataset was structured into contrast-enhanced venous-phase CT volumes (venous.nii), their corresponding masks (label_venous.nii), non-contrast CT volumes (native.nii) and their corresponding masks (label_native.nii).

Each label file contained at least one annotated lymph node, enabling the model to learn lymph nodes as individual entities rather than as grouped structures. Importantly, lymph nodes were labeled on both contrast-enhanced and na-

tive CT scans to ensure that the model can generalize to patients in whom intravenous contrast administration is contraindicated, thereby increasing its clinical applicability.

Data augumentation

Considering that one of the objectives of the thesis was to evaluate model performance under conditions of limited preprocessing and minimal augmentation, a conservative augmentation strategy was adopted in this study. This consisted of applying random rotations to axial images within a range of ±10 degrees, without effectively increasing the final size of the dataset.

This approach was chosen to simulate minor variations that may occur in clinical practice, particularly those related to suboptimal head positioning during CT acquisition. The purpose was not to generate additional training samples, but rather to improve the model's generalization ability and reduce its sensitivity to small variations in image orientation.

It is important to note that the model used in this study was trained and applied in a two-dimensional (2D) manner, on axial slices, rather than as a three-dimensional (3D) network. Therefore, augmentation was applied at the axial slice level, in accordance with the operating mode of the network, which is described in the following subsection.

Implementation and model architecture

The model was implemented in Python using a Jupyter Notebook environment. Training was performed on the Kaggle platform, which provides access to NVIDIA Tesla P100 GPUs (with a usage limit of approximately 30 hours per week), along with Intel Xeon CPUs (2.3 GHz, 16 cores) and 13 GB of RAM.

A convolutional neural network based on the U-Net++ architecture was employed for lymph node segmentation. The model was implemented using the segmentation_models_pytorch library, with the base architecture adapted from publicly available source code. U-Net++ consists of an encoder–decoder structure connected through a series of nested dense convolutional blocks, designed to reduce the semantic gap between encoder and decoder feature maps prior to fusion (16).

The model was further trained (fine-tuned) on the study-specific dataset to achieve improved performance.

Although the input data consisted of volumetric CT scans in NIfTI format, the model operated in a 2D manner. Specifically, the volumes were decomposed into individual axial slices, which were processed independently during training. The resulting slice-wise predictions were subsequently reconstructed into volumetric segmentation maps. Input data were resized to 256 × 256 pixels, with a single-channel configuration (grayscale). The model output consisted of binary segmentation masks corresponding to lymph node regions.

Training protocol and performance analysis

The dataset was divided into training, validation, and test subsets. From the 67 standard cases, 58 were used for training and nine for validation. This split was patient-based, before slice extraction in order to prevent data leakage, and included both native and venous phases of the respective cases in the same partition (train or validation). At slice level, to ensure a balanced learning, the slices were approximately equally distributed between those containing lymph nodes and those without lymph nodes, thus obtaining 2492 axial images for training and 672 slices for validation. Important to clarify that axial images form the training set were not included in the validation set and vice-versa.

The testing set included the 12 special cases, with head and neck adenopathies (n=7), lymph nodes from other anatomical regions (n=3), accesory spleen (n=1) and carotid body paraganglioma (n=1).

Data augmentation was applied to the training dataset in the form of random in-plane rotations within a range of ±10 degrees, in order to improve model generalization.

The model was trained using a batch size of 8 over 17 epochs. Optimization was performed using the Adam optimizer, with a learning rate of 1×10^{-4}. The loss function consisted of a combination of Dice loss and binary cross-entropy loss, allowing for improved performance in segmentation tasks with class imbalance.

Model performance was evaluated using the Dice Similarity Coefficient (DSC), Intersection over Union (IoU), sensitivity, specificity and accuracy. The Dice coefficient was considered the pri-

mary evaluation metric due to its robustness in assessing spatial overlap between predicted and ground truth segmentations.

Because the model was implemented as a 2D slice-wise segmentation network, all quantitative performance metrics were calculated at the axial slice level. CT volumes were decomposed into individual axial slices during training and validation, and each slice was treated as an independent input sample. Therefore, accuracy, sensitivity, specificity, precision, DSC and IoU reflect slice-level performance and were not computed at the patient, scan, or whole-lesion volumetric level.

Additional robustness testing was performed by evaluating the model on anatomical structures that may mimic lymph nodes, as well as on lymph nodes located in other anatomical regions, in order to assess potential sources of false-positive detections. ❑

RESULTS

General results

The model was evaluated on a dataset comprising 79 CT examinations, with a total of 538 annotated lymph nodes. The dataset included lymph nodes classified as Node-RADS 3, 4 and 5, reflecting equivocal to highly suspicious findings.

The model achieved a training accuracy of 99.8% and a validation accuracy of 99.7%. Sensitivity was 78% on the training set and 72% on the validation set, while specificity remained consistently high at 99.9% for both datasets. Precision reached 83% during training and 79% during validation (Figure 1).

In terms of segmentation performance, the model obtained a DSC of 0.79 on the training set and 0.73 on the validation set. The corresponding IoU values were 0.67 and 0.61, respectively. The training loss was 0.2, compared to 0.5 on the validation set (Figure 2).

All quantitative metrics reported for the training and validation subsets were calculated at the axial slice level, consistent with the 2D slice-wise design of the model.

Notably, the model showed consistent performance in identifying both clearly malignant and equivocal (Node-RADS 3) lymph nodes, supporting its potential role in assisting radiologists in detecting borderline findings.

Comparable performance was observed on both contrast-enhanced (venous phase) and non-contrast (native) CT images, suggesting that the model can be effectively applied in patients for whom contrast administration is contraindicated.

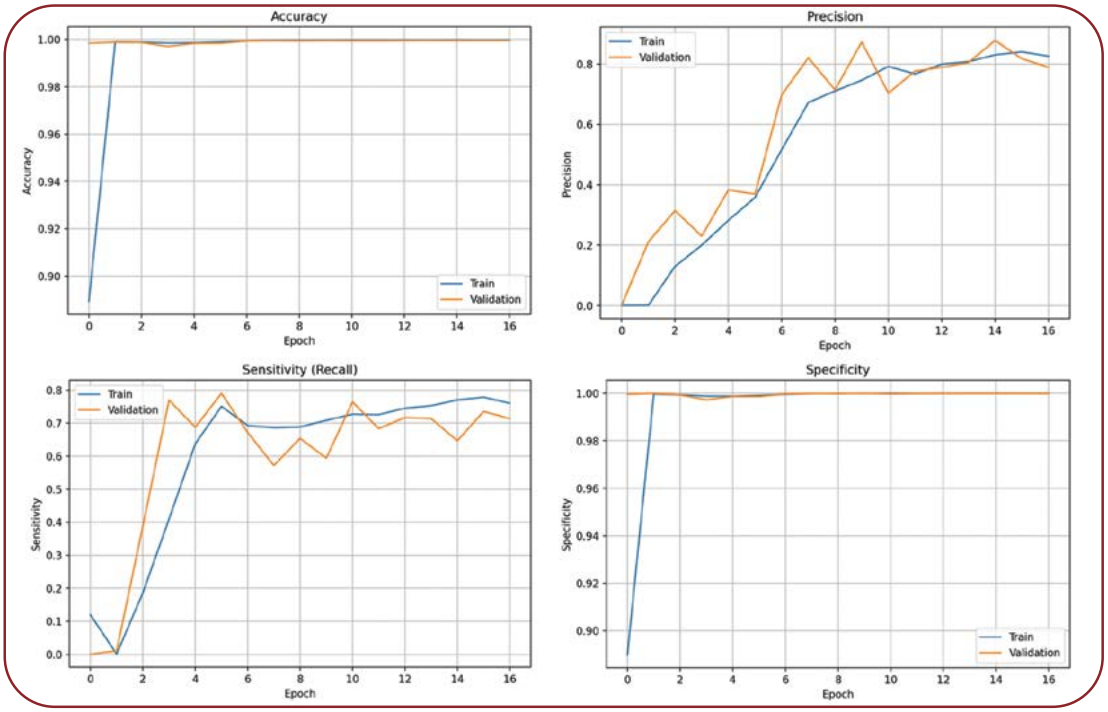

FIGURE 1. Graphical representation of accuracy, precision, sensitivity and specificity for both training and validation datasets

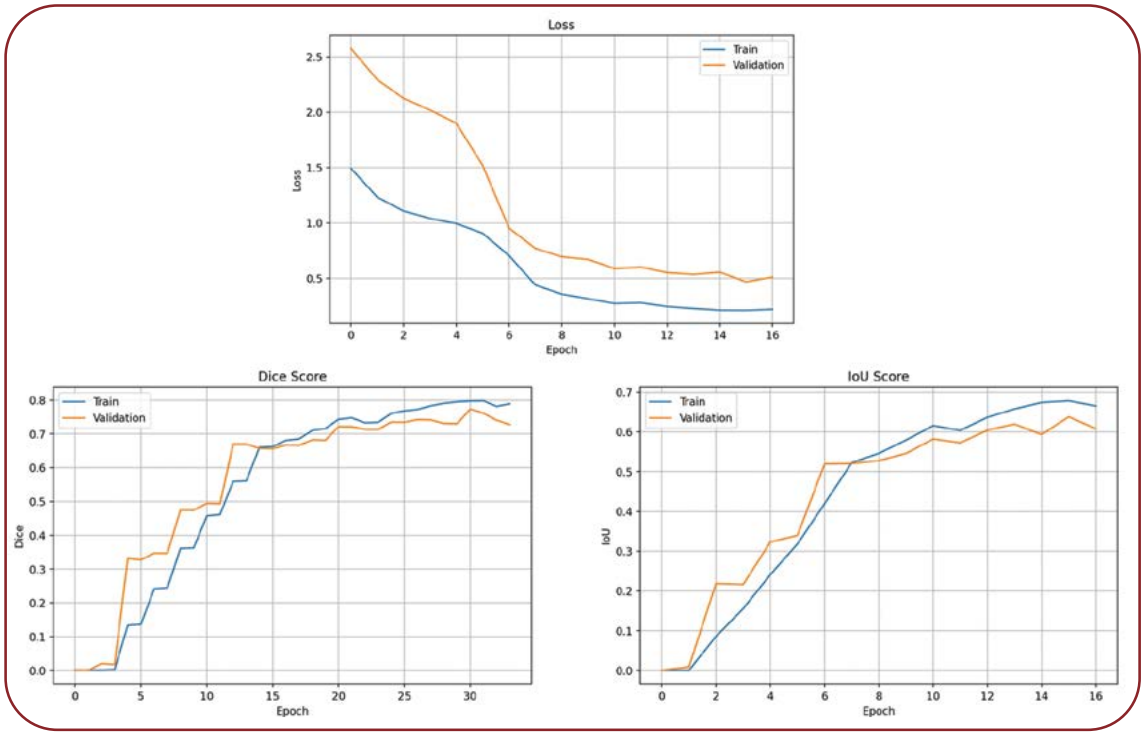

FIGURE 2. Graphical representation of loss, Dice Similarity Coefficient (DSC), Intersection over Union (IoU) for both training and validation datasets

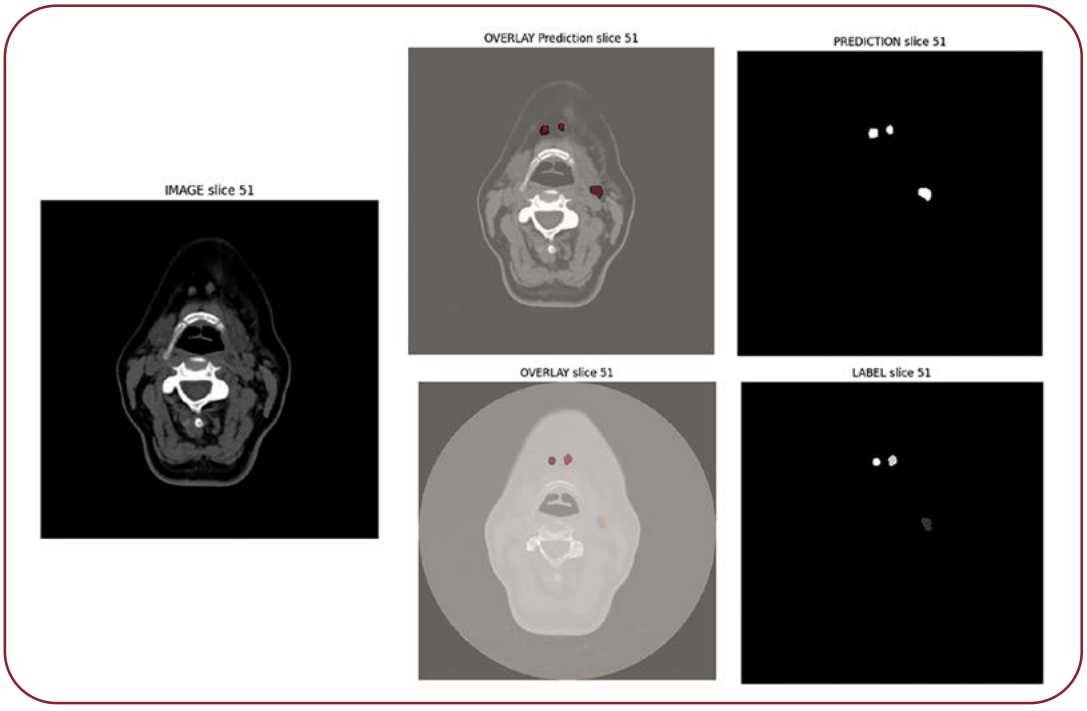

FIGURE 3. Example of model-predicted segmentation compared with the manual ground truth annotations. The manual labels are shown in varyinggrayscale intensities, as each lymph node was annoted as individual entity

Test set performance

In addition to the training and validation datasets, an independent test set comprising 12 CT examinations was used for qualitative evaluation. These cases were not included in the training process and were analyzed in their entirety. Full volume-

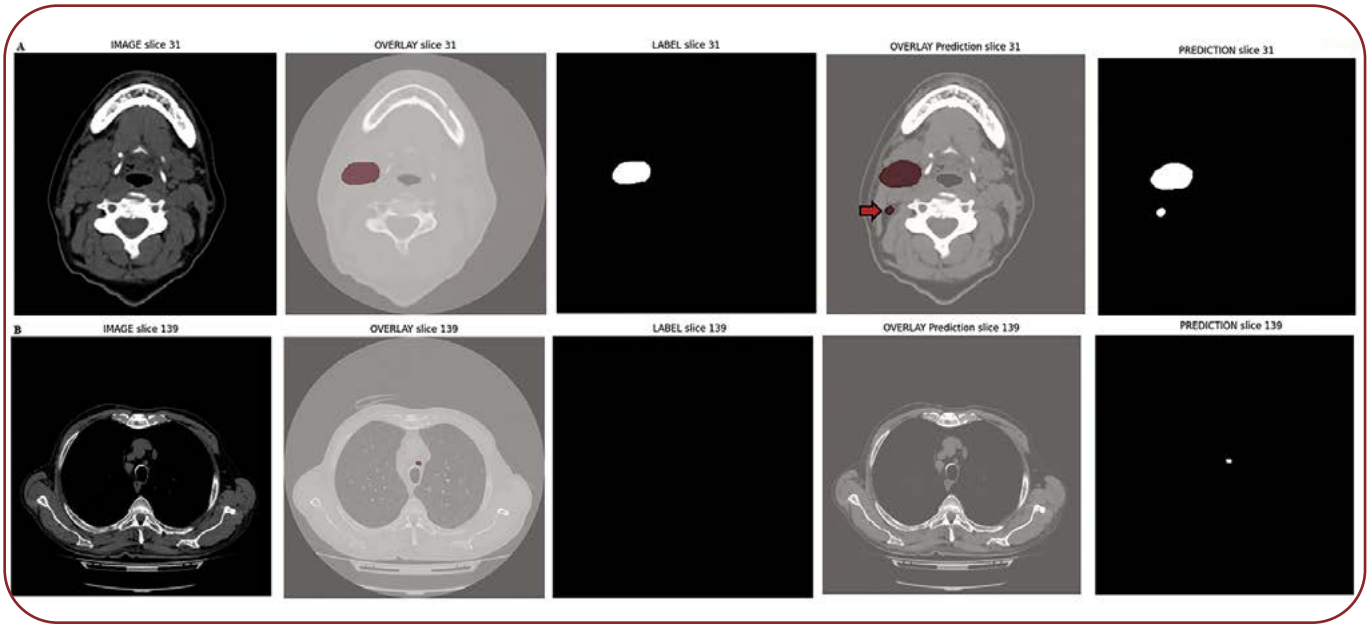

FIGURE 4. (A) False positive lymph node (red arrow); (B) Small mediastinal lymph node. Panels A and B correspond to different patients

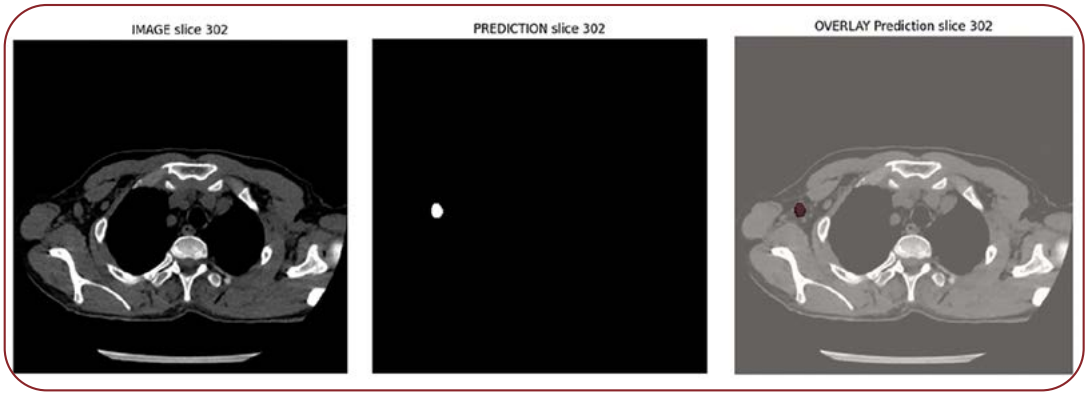

FIGURE 5. Axillary lymph node identified by the model in a patient with history of breast cancer

tric annotations were available for all slices, and no discontinuities in lymph node labeling were observed despite the slice-wise processing approach of the model.

The model demonstrated consistent performance across the test cases, with representative examples of accurate segmentation shown in Figure 3.

A limited number of false-negative results were identified, involving three equivocal (Node-RADS 3) lymph nodes that were detected by the radiologist but missed by the model.

Two false-positive cases were observed: one located in the head and neck region (Figure 4 A) and one in the mediastinum (Figure 4 B). Notably, these corresponded to true lymph nodes; however, they were classified as low-risk (Node-RADS 1) by the radiologist. This suggests that the model may have a tendency to over-segment or assign suspicion to benign lymph nodes with subtle imaging features.

As an exploratory observation, the model identified an axillary lymph node in a patient with breast cancer, despite not being specifically trained for this anatomical region (Figure 5). This finding should not be interpreted as evidence of cross-anatomical generalizability, but rather as a preliminary observation that warrants further evaluation in dedicated datasets including lymph nodes from other anatomical regions.

Importantly, structures that may mimic lymph nodes, such as accessory spleens, were not incorrectly classified by the model. Similarly, a large carotid body paraganglioma was not misidentified

as a lymph node. In addition, large benign inguinal lymph nodes were not classified as suspicious.

Submandibular salivary glands, which in one case presented with increased size and an oval morphology, were also correctly excluded from lymph node classification. Furthermore, no misclassification was observed in relation to vascular structures or adjacent muscular tissue, indicating that the model was able to distinguish lymph nodes from common anatomical mimics despite operating in a slice-wise (2D) manner. No cases involving thymic tissue were present in the dataset. ❑

DISCUSSIONS

General discussions

A discrepancy was observed between accuracy and overlap-based metrics such as Dice and IoU. This may be partially explained by limitations in the labeling process, as some ground truth masks exhibited pixelation despite the application of smoothing techniques. In contrast, the model generated smoother and more continuous segmentation outputs, which may have contributed to lower overlap scores. Additionally, IoU is known to penalize even minor boundary inconsistencies more strongly, further amplifying these differences (17). These findings highlight the importance of overlap-based metrics in medical image segmentation, as accuracy alone may overestimate model performance in highly imbalanced datasets.

Despite the use of a 2D slice-wise architecture, the model demonstrated a strong ability to differentiate lymph nodes from anatomically similar structures. This was evidenced by the correct exclusion of accessory spleens, submandibular salivary glands with atypical morphology, vascular structures and adjacent muscular tissue. These findings suggest that the model effectively learned discriminative local imaging features, such as texture, attenuation patterns, and boundary characteristics, which are sufficient for distinguishing lymph nodes from common mimics on individual CT slices.

Although 3D architectures may provide additional contextual information by incorporating inter-slice continuity, the present results indicate that, in this setting, local spatial features alone can achieve a high degree of anatomical discrimination. This is particularly relevant in clinical practice, where rapid slice-based assessment remains the standard approach for radiological interpretation.

The detection of an axillary lymph node outside the primary training region is an interesting exploratory finding. However, given that this observation was limited to a single case, it cannot be considered evidence of cross-anatomical generalization. Future studies should evaluate this possibility using dedicated multicenter datasets that include lymph nodes from multiple anatomical regions. At the same time, the occurrence of false-negative results in a small number of equivocal (Node-RADS 3) lymph nodes highlights the inherent difficulty of borderline cases and suggests that such findings should continue to be interpreted in conjunction with clinical and radiological expertise.

The model was not trained on the complete spectrum of benign and malignant lymph node appearances, thus the decision to exclude Node-RADS 1-2 lymph nodes. This choice was intentional, as the purpose of the model was to identify and segment equivocal or suspicious lymph nodes, rather than to classify all visible lymph nodes. Consequently, the model should be regarded as a tool for highlighting potentially relevant "red flag" findings for radiological review, not as a comprehensive lymph node classifier or a replacement for structured radiological assessment.

When compared with previously published studies on lymph node segmentation in head and neck CT imaging, the performance of the proposed model falls within a comparable range, particularly given the complexity of the dataset. However, direct comparison of quantitative metrics should be interpreted with caution, as differences in study design, annotation protocols and target definitions significantly influence reported Dice and IoU values. For example, a machine learning approach based on a U-Net architecture applied to small cervical lymph nodes (5–10 mm) reported a Dice score of approximately 0.80, although it did not evaluate lymph nodes outside the cervical region (18). Diagnosing small lymph node metastases across different cancer types remains a major challenge and recent studies have focused primarily on improving classification performance rather than segmentation (19-22).

Other approaches have incorporated radiomic features to enhance model performance, yielding results that are broadly comparable to those ob-

tained in the present study (22-24). Importantly, the differences in accuracy between studies may reflect variations in study objectives. While many models focus exclusively on clearly malignant lymph nodes, the present study aimed to identify both malignant and equivocal (Node-RADS 3–5) lymph nodes, thereby addressing a more clinically challenging and relevant scenario.

In addition, lymph node characteristics may vary depending on anatomical location (25). For example, mediastinal lymph nodes may enlarge in response to a wide range of benign and malignant conditions, with reported Dice scores in this region ranging from 0.49 to 0.67 (26, 27), further highlighting the variability of segmentation performance across anatomical sites.

Traditional criteria such as RECIST are not always reliable in distinguishing individual lymph nodes, particularly in cases of overlapping or clustered structures. For this reason, the present study employed the Node-RADS classification system (28), which provides a more nuanced framework for risk stratification. Recent studies have also explored deep learning approaches for differentiating lymph node pathologies, such as papillary thyroid carcinoma, tuberculosis and HPV-positive oropharyngeal squamous cell carcinoma, achieving accuracies of approximately 0.76 (29).

Alternative imaging approaches have also been investigated. For example, Forghani *et al* proposed a model based on dual-energy CT texture analysis, demonstrating high sensitivity but lower specificity and overall accuracy (30). However, dual-energy CT requires specialized acquisition protocols, which may limit its widespread applicability compared to conventional CT imaging.

More recently, studies have focused on hybrid imaging modalities such as FDG PET-CT, incorporating metabolic parameters such as standardized uptake value (SUV) to improve differentiation between malignant and inflammatory lymph nodes (31-34). The integration of CT-based features with functional imaging biomarkers represents a promising direction for future research and may further enhance diagnostic accuracy.

Limitations of the study

This study has several limitations. First, all cases were collected from a single healthcare institution, which may limit the generalizability of the results. However, efforts were made to reduce potential bias by including data acquired from four different CT scanners.

Second, although image interpretation and label verification were performed by multiple experienced radiologists, but the manual segmentation process was carried out by a single physician, which may introduce operator-dependent variability. This aspect should be addressed in future studies through multi-reader annotation and inter-observer agreement analysis.

Third, because only Node-RADS 3–5 lymph nodes were included, the model was designed to detect and segment equivocal or suspicious lymph nodes rather than to classify the full spectrum of lymph node appearances. Therefore, it should be interpreted as a suspicious-node highlighting tool, not as a comprehensive lymph node classifier.

Additionally, the use of a 2D slice-wise architecture limits the incorporation of inter-slice contextual information, which may affect the detection of small or irregular lymph nodes.

Statements regarding anatomical discrimination and potential generalizability should be interpreted cautiously, as the study was single-center and the independent qualitative test set was limited.

Finally, the relatively limited dataset size may impact model robustness and performance.

Future directions

Future work will focus on further improving model performance through the integration of volumetric deep learning approaches, such as 3D convolutional neural networks, which may better capture spatial relationships between adjacent slices. In particular, the implementation of a 3D U-Net-based architecture could enhance the detection of small or irregular lymph nodes by incorporating inter-slice contextual information.

Additionally, expanding the dataset and incorporating multi-institutional data may improve model generalizability and robustness. Future developments may also include the integration of the model into clinical workflows, with the aim of providing real-time decision support and facilitating earlier detection of suspicious or borderline lymph nodes in routine radiological practice. ❑

CONCLUSION

In this study, a 2D U-Net++-based model demonstrated good performance in the segmenta-

tion and identification of malignant and equivocal lymph nodes on CT imaging, achieving high specificity and overall accuracy.

The model showed particular strength in distinguishing lymph nodes from anatomically similar structures, supporting its potential as a reliable tool for assisting radiologists in routine clinical practice. Additionally, its ability to perform consistently on both contrast-enhanced and non-contrast CT images broadens its applicability, especially in patients for whom contrast administration is contraindicated.

While sensitivity remains moderate, particularly in borderline (Node-RADS 3) cases, the model may serve as a valuable adjunct for highlighting suspicious findings and facilitating earlier detection.

Further improvements, including the use of larger multicentric datasets and the implementation of volumetric (3D) architectures, may enhance performance and generalizability. ❑

Financial support: none declared.
Conflicts of interest: none declared.

References

1. **Pfister DG, et al.** Head and neck cancers, version 2.2020, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2020;18:873-898.
2. **Amin MB, et al.** The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin* 2017;67:93-99.
3. **Dudea SM, Lenghel M, Botar-Jid C, et al.** Ultrasonography of superficial lymph nodes: Benign *vs.* malignant. *Med Ultrasonogr* 2012;14:294.
4. **Norris CD, Anzai Y.** Anatomy of Neck Muscles, Spaces, and Lymph Nodes. *Neuroimaging Clin N Am* 2022;32:831-849.
5. **Forghani R, Yu E, Levental M, et al.** Imaging evaluation of lymphadenopathy and patterns of lymph node spread in head and neck cancer. *Expert Rev Anticancer Ther* 2015;15:207-224. doi: 10.1586/14737140.2015.978862.
6. **Forghani R, Johnson JM, Ginsberg LE.** Imaging of head and neck cancer. In: *Cancer of the Head and Neck*. J. Myers, E. Hanna, and E. N. Myers Eds., 5th ed., Philadelphia: Wolters Kluwer. 2017, pp 92-148.
7. **Koroulakis A, Jamal Z, Agarwal M.** Anatomy, Head and Neck, Lymph Nodes. (Updated 2022 Dec 11). In: *StatPearls*; StatPearls Publishing: Treasure Island, FL, USA, 2025.
8. **Payabvash S, Meric K, Cayci Z.** Differentiation of benign from malignant cervical lymph nodes in patients with head and neck cancer using PET/CT imaging. *Clin Imaging* 2016;40:101-105. doi: 10.1016/j.clinimag.2015.09.001.
9. **Som PM, Brandwein-Gensler MS.** Lymph nodes of the neck, in head and neck imaging. In: Som PM, Curtin HD, Editors. Mosby; St. Louis, Mo. 2011.
10. **Elsholtz FHJ, Asbach P, Haas M, et al.** Introducing the Node Reporting and Data System 1.0 (Node-RADS): a concept for standardized assessment of lymph nodes in cancer. *Eur Radiol* 2021;31:6116-6124. Erratum in: Eur Radiol 2021;31:7217.
11. **Zhong J, Mao S, Chen H, et al.** Node-RADS: a systematic review and meta-analysis of diagnostic performance, category-wise malignancy rates, and inter-observer reliability. *Eur Radiol* 2025;35:2723-2735.
12. **Hoang JK, Vanka J, Ludwig BJ, Glastonbury CM.** Evaluation of cervical lymph nodes in head and neck cancer with CT and MRI: tips, traps, and a systematic approach. *AJR Am J Roentgenol* 2013;200:W17-W25. doi: 10.2214/AJR.12.8960.
13. **Taciuc IA, Dumitru M, Vrinceanu D, et al.** Applications and challenges of neural networks in otolaryngology (Review). *Biomed Rep* 2024;20:92.
14. **Choy G, Khalilzadeh O, Michalski M, et al.** Current Applications and Future Impact of Machine Learning in Radiology. *Radiology* 2018;288:318-328.
15. **Song B, Leroy A, Yang K, et al.** Deep Learning Model of Primary Tumor and Metastatic Cervical Lymph Nodes From CT for Outcome Predictions in Oropharyngeal Cancer. *JAMA Netw Open* 2025;8:e258094.
16. **Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J.** UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support* (2018) 2018;11045:3-11.
17. **Taciuc I-A, Dumitru M, Marinescu A, et al.** Enhancing Malignant Lymph Node Detection in Ultrasound Imaging: A Comparison Between the Artificial Intelligence Accuracy, Dice Similarity Coefficient and Intersection over Union. *J Mind Med Sci* 2025;12:29.
18. **Al Hasan MM, Ghazimoghadam S, Tunlayadechanont P, et. al.** Automated Segmentation of Lymph Nodes on Neck CT Scans Using Deep Learning. *J Imaging Inform Med* 2024;37:2955-2966.
19. **Ariji Y, Fukuda M, Kise Y, 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;127:458-463.
20. **Lee JH, Ha EJ, Kim JH.** Application of deep learning to the diagnosis of cervical lymph node metastasis from thyroid cancer with CT. *Eur Radiol* 2019;29:5452-5457.
21. **Kann BH, et al.** Multi-institutional validation of deep learning for pretreatment identification of extranodal extension in head and neck squamous cell carcinoma. *J Clin Oncol* 2020;38:1304-1311.
22. **Ho TY, Chao CH, Chin SC, et al.** Classifying neck lymph nodes of head and neck squamous cell carcinoma in MRI images with radiomic features. *J Digit Imaging* 2020;33:613-618.
23. **Mukherjee P, et al.** CT-based radiomic signatures for predicting histopathologic features in head and neck squamous cell carcinoma. *Radiol Imaging Cancer* 2020;2:e190039.
24. **Tomita H, Yamashiro T, Heianna J, 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;31:7440-7449.

25. **Bouget D, Jørgensen A, Kiss G, Leira HO, Langø T**. Semantic segmentation and detection of mediastinal lymph nodes and anatomical structures in ct data for lung cancer staging.
Int J Comput Assist Radiol Surg 2019; 14:977-986.
26. **Mathai TS, Liu B, Summers RM.** Segmentation of mediastinal lymph nodes in CT with anatomical priors.
Int J Comput Assist Radiol Surg 2024;19:1537-1544.
27. **Oda H, Bhatia KK, Roth HR, et al.** Dense volumetric detection and segmentation of mediastinal lymph nodes in chest CT images. SPIE Med Imaging 2018.
28. **Shafiei A, Bagheri M, Farhadi F, et al.** CT Evaluation of Lymph Nodes That Merge or Split during the Course of a Clinical Trial: Limitations of RECIST 1.1.
Radiol Imaging Cancer 2021;3:e200090.
29. **Onoue K, Fujima N, Andreu-Arasa VC, et al.** 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 2021;42:103026.
30. **Forghani R, Chatterjee A, Reinhold C, et al.** Head and neck squamous cell carcinoma: prediction of cervical lymph node metastasis by dual-energy CT texture analysis with machine learning.
Eur Radiol 2019;29:6172-6181.
31. **de Koekkoek-Doll PK, Vogel W, Maas M, et al.** SUVmax values at FDG PET-CT to predict malignancy in lymph nodes aspirated by real time image fused USgFNAC in head and neck squamous cell carcinoma.
Am J Nucl Med Mol Imaging 2021;11:178-187.
32. **Santer M, Zelger P, Schmutzhard J, et al.** The Neck-Persistency-Net: a three-dimensional, convolution, deep neural network aids in distinguishing vital from non-vital persistent cervical lymph nodes in advanced head and neck squamous cell carcinoma after primary concurrent radiochemotherapy.
Eur Arch Otorhinolaryngol 2024;281:5971-5982.
33. **Kuno H, Garg N, Qureshi MM, et al.** CT Texture Analysis of Cervical Lymph Nodes on Contrast-Enhanced (18F) FDG-PET/CT Images to Differentiate Nodal Metastases from Reactive Lymphadenopathy in HIV-Positive Patients with Head and Neck Squamous Cell Carcinoma.
AJNR Am J Neuroradiol 2019;40:543-550.
34. **Droogers E, Pruis IJ, van der Eerden AW, et al.** Clinical evaluation of (18F) fluoro-2-deoxy-d-glucose positron emission tomography-magnetic resonance imaging ((18F)FDG PET-MRI) and (18F)FDG PET-computed tomography ((18F)FDG PET-CT) for lymph node assessment in head and neck squamous cell carcinoma.
Clin Radiol 2025;91:107095.