HectoMixNet: Advancing Automated Head and Neck Tumor Segmentation with Multicenter PET/CT Data

Moona Mazher[1], Steven A Niederer[2] and Abdul Qayyum[2]

[1] Hawkes Institute, Department of Computer Science, University College London, London, United Kingdom
[2] National Heart and Lung Institute, Faculty of Medicine, Imperial College London, London, United Kingdom

Abstract. Head and Neck (H&N) cancers represent one of the most prevalent malignancies worldwide, with accurate delineation of primary tumors and metastatic lymph nodes being essential for radiotherapy planning and treatment outcome prediction. Manual segmentation, though standard, is time-consuming and prone to inter-observer variability, motivating the development of automated deep learning–based methods. In this study, we propose HectoMixNet, a novel 3D encoder–decoder framework augmented with a bidirectional quasiseparable mixing module for robust segmentation of primary tumors and lymph nodes in multimodal PET/CT imaging. The proposed module efficiently integrates local anatomical detail with global volumetric context, addressing the challenges of heterogeneous and spatially complex lesions. We evaluate HectoMixNet on the HECKTOR 2025 Task 1 dataset, the largest multicentric benchmark for H&N cancer analysis to date, comprising over 1,200 patients from 11 international centers. Our model achieved strong performance, with a GTVp Dice score of 0.8812, GTVn Dice of 0.8246, and GTVn F1 of 0.7273 on the validation and leaderboard sets. Compared to state-of-the-art baselines such as 3D Mamba and xLSTM, HectoMixNet demonstrated superior lymph node segmentation accuracy, highlighting the importance of bidirectional quasiseparable mixing for modeling long-range spatial dependencies. These results establish HectoMixNet as an effective and clinically relevant tool for automated H&N cancer segmentation in large-scale, heterogeneous imaging cohorts.

Keywords: Bidirectional Quasiseparable Mixing, State-Space Models, Head and Neck Cancer Tumor Segmentation, Lymph Node Detection, PET/CT Imaging, Radiotherapy Planning, HECKTOR Challenge.

1 Introduction

Head and Neck (H&N) cancers represent the fifth most common malignancy worldwide, with a substantial global disease burden and significant clinical management challenges [1]. Accurate delineation of primary tumors and metastatic lymph nodes plays a crucial role in radiotherapy planning, treatment monitoring, and outcome prediction. Conventionally, segmentation is performed manually by expert radiation oncologists, a process that is both time-consuming and subject to high inter-observer variability. As a result, there is a growing demand for robust, automated segmentation

algorithms that can generalize across heterogeneous, multicenter datasets. The HECKTOR (Head and Neck Tumor Segmentation and Outcome Prediction) challenge, organized annually in conjunction with the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), has become the leading benchmark for automated H&N cancer analysis on multimodal Positron Emission Tomography (PET) and Computed Tomography (CT) data. Following the success of the 2020–2022 editions [1-5], the 2025 HECKTOR challenge expands the dataset to more than 1,200 patients from 11 international centers, making it the largest multicentric cohort available for this task. The 2025 edition introduces refined evaluation metrics to better assess detection and segmentation of multiple distinct lesions, thereby improving clinical relevance. In recent years, numerous segmentation models have been developed for medical image analysis [5-18], demonstrating the potential of deep learning to accurately delineate complex anatomical structures and lesions.

In this work, we present HectoMixNet, a novel deep learning framework for automatic segmentation of primary tumors and lymph nodes in H&N cancer patients. The model builds upon a classical 3D encoder–decoder design but incorporates a specialized bidirectional quasiseparable mixing module at the network bottleneck, enabling effective modeling of both local detail and global spatial context. By integrating fine anatomical information with long-range dependencies across volumetric scans, HectoMixNet is tailored to address the challenges of detecting irregular, heterogeneous tumor structures. We evaluate our method on the HECKTOR 2025 Task 1 dataset and demonstrate its ability to achieve precise and robust segmentation in a challenging multicenter setting.

2 Proposed Method

In this section, we describe the dataset used for training and evaluation, followed by the architecture and implementation details of the proposed HectoMixNet framework.

2.1 Dataset

The dataset used in this study is provided by the HECKTOR 2025 challenge, which focuses on automatic detection and segmentation of head and neck primary tumors and lymph nodes from multimodal PET/CT imaging. The training cohort consists of approximately 850 patients drawn from the HECKTOR 2022 database [5], updated with additional clinical variables including HPV status and extended follow-up information. For the 2025 challenge, two new French cohorts have been added, forming an unseen test set of ~400 patients. Altogether, the dataset comprises more than 1,200 patients from at least 11 medical centers, ensuring significant heterogeneity in imaging protocols and patient demographics [19].

Each patient study contains co-registered FDG-PET and CT scans, with corresponding expert-annotated segmentation masks for the primary tumor (GTVp) and metastatic lymph nodes (GTVn). PET images provide complementary metabolic information to the anatomical details visible in CT, enabling more reliable lesion identification. The segmentation task requires the model to accurately detect and delineate both lesion types, including cases with multiple, spatially distinct lymph node metastases. For

evaluation, the challenge employs updated metrics that measure both overall volumetric overlap (e.g., Dice similarity) and lesion-wise detection performance, reflecting the clinical importance of identifying all tumor sites for accurate treatment planning.

2.2 Preprocessing

For preprocessing, we followed the standardized pipeline provided by the challenge organizers. CT and PET scans were stacked as two-channel volumes and processed as follows:

- Resampling & Alignment: All scans were resampled to an isotropic resolution of 1 mm^3 and aligned using bounding-box coordinates to ensure PET–CT correspondence.
- Region-of-Interest Cropping: A fixed box of 200 × 200 × 310 voxels centered on PET-intensity landmarks was extracted to capture the head and neck region while reducing irrelevant background.
- Normalization: CT intensities were clipped to [−250, 250] HU and scaled to [−6, 6], while PET images were normalized to zero mean and unit variance.
- Reorientation & Padding: All images were reoriented to the RAS convention, foreground-cropped, and padded to the target size to ensure uniform input dimensions.

This preprocessing pipeline, implemented with SimpleITK and MONAI, standardizes resolution, orientation, and intensity distributions across the cohort, while preserving anatomical fidelity and ensuring compatibility with deep learning workflows.

2.3 Postprocessing

To restore predictions to the native imaging space, we applied a structured inverse transformation pipeline. First, symmetric padding introduced during preprocessing was removed, followed by reinsertion into the cropped foreground and neck region of interest. The reconstructed mask in the standardized 1-mm isotropic space was then resampled back to the original CT resolution using nearest-neighbor interpolation to preserve label integrity. This ensured voxel-level alignment of the predicted segmentation with the scanner-resolution images, facilitating accurate evaluation and downstream clinical usability.

2.4 Model

1. HectoMixNet: Proposed Model for 3D heterogeneous H&N tumors and lymph nodes

To address the challenges of segmenting heterogeneous H&N tumors and lymph nodes, we propose HectoMixNet, a 3D convolutional encoder–decoder network augmented with a bidirectional quasiseparable mixing module. The architecture is designed to combine local feature extraction with global spatial context modeling, a critical requirement for distinguishing small, irregular lesions as well as capturing large-scale disease patterns.

The encoder consists of hierarchical 3D convolutional blocks with strided convolutions, batch normalization, and non-linear activations. This stage progressively reduces spatial resolution while extracting rich multi-scale representations of tumor and anatomical

structures. Skip connections are preserved to ensure high-fidelity recovery of spatial details. At the network bottleneck, feature maps are passed through the HectoMix module, which reshapes the volumetric features into a sequence and applies bidirectional quasiseparable mixing using semiseparable state-space models. This operation efficiently propagates information from both past (causal) and future (anti-causal) spatial positions, while preserving self-connections for stability. In doing so, each voxel representation integrates contextual knowledge from the entire 3D volume.

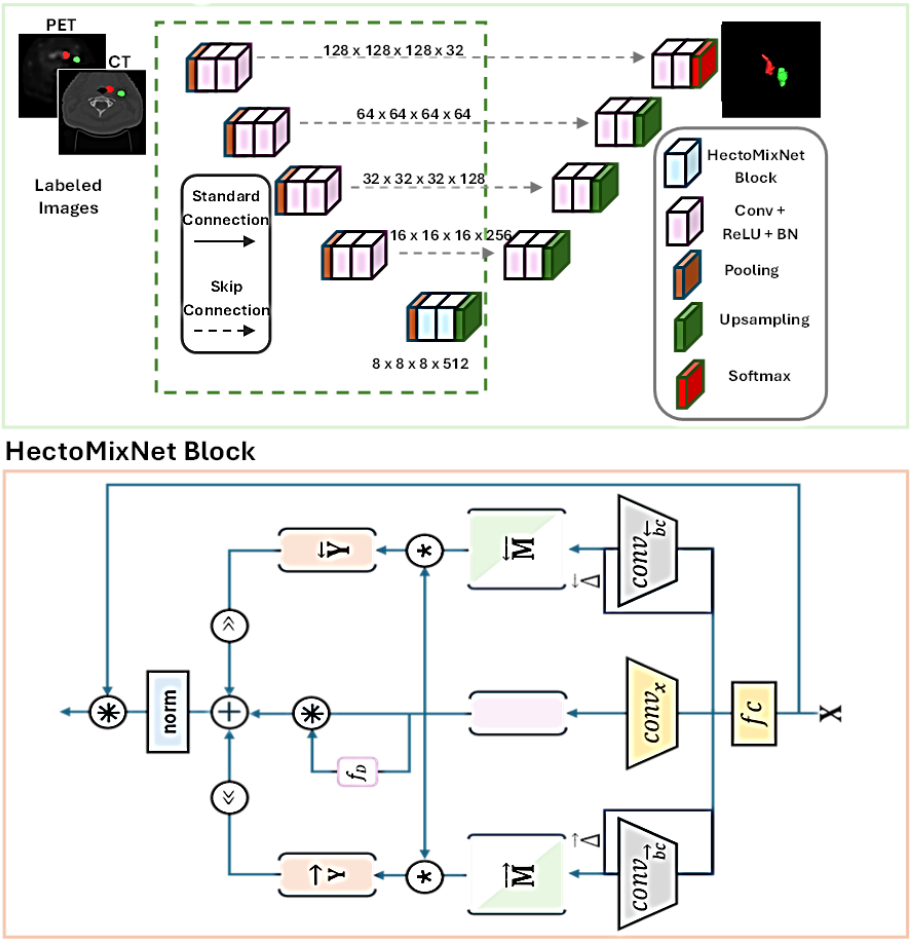

Figure 1. Architectural framework of HectoMixNet, designed for segmentation of complex and heterogeneous head and neck tumors.

The decoder mirrors the encoder, using 3D transposed convolutions to restore spatial resolution. Mixed global features from the bottleneck are fused with skip-connected encoder features at multiple scales, allowing fine-grained lesion boundaries to be

reconstructed while maintaining global consistency. This design enables HectoMixNet to adaptively handle both small, isolated lymph nodes and larger, diffused primary tumors.

By unifying multi-scale local features with efficient long-range dependency modeling, HectoMixNet offers an effective and computationally scalable solution for volumetric medical image segmentation. Its ability to capture global relationships while preserving anatomical detail makes it particularly well-suited for the complexity of head and neck tumor segmentation in the HECKTOR 2025 dataset.

The core mechanism of mixing aggregates information from three sources for each sequence position: past positions (causal context), future positions (anti-causal context), and the current position itself. In the causal pass, a semiseparable SSM operation accumulates information from all preceding positions, with the output shifted to align with the current position. The anti-causal pass reverses the sequence and applies the same causal SSM operation; the result is shifted and reversed back to propagate information from future positions. This bidirectional processing ensures every spatial location benefits from both preceding and succeeding spatial context.

Additionally, a self-connection term applies learnable diagonal parameters to each position's features, preserving local information and stabilizing the mixing process. Mathematically, this operation corresponds to a quasiseparable mixing matrix M structured as follows: the diagonal entries represent self-connection weights, the lower triangular part encodes causal mixing weights for past positions, and the upper triangular part encodes anti-causal mixing weights for future positions. The quasiseparable structure implies a low-rank representation that enables efficient computation without explicitly forming a large dense matrix [20, 21].

Together, this bidirectional quasiseparable mixing yields a powerful, context-aware feature representation that integrates global spatial relationships effectively. This makes mixing highly suitable for 3D medical imaging tasks like TBI, where modeling both local details and global spatial dependencies is critical for precise and reliable analysis.

The equation below describes how the output at each sequence position i is computed by combining three components: contributions from all past positions $j < i$, contributions from all future positions $j > i$, and direct self-connection at the position i. The matrix M organizing these parameters is structured so that their diagonal elements δ_i represent self-connections preserving local features. The lower triangle contains parameters $A_i b_j$ that govern how information flows causally from past positions to the current one, while the upper triangle contains parameters $A'_i b'_j$ for anti-causal mixing, aggregating information from future positions. This quasiseparable structure enables efficient modeling of bidirectional dependencies without explicitly storing a full dense matrix.

For each batch b, position i:

$$Y_{b,i} = \sum_{j<i} A_i b_j X_{b,j} + \sum_{j>i} A'_i b'_j X_{b,j} + \delta_i X_{b,i}$$

Where A_i, b_j: causal parameters, A'_i, b'_j: anti-causal parameters and δ_i: diagonal/self-connection

The matrix M shown for sequence length 16 illustrates how the bidirectional mixing operates. Its diagonal elements δ_i represent self-connection weights that preserve local information at each position i. The lower triangular part (below the diagonal) contains parameters $A_i b_j$that mix information causally from past positions $j < i$ to the current position iii. Conversely, the upper triangular part (above the diagonal) holds parameters $A'_i b'_j$ responsible for mixing information anti-causally from future positions $j > i$ back to position i. This structured matrix efficiently captures dependencies in both directions while maintaining a sparse, low-rank form for fast computation.

$$M = \begin{bmatrix} \delta_1 & A'_1 b'_2 & A'_1 b'_3 & \cdots & A'_1 b'_{16} \\ A_2 b_1 & \delta_2 & A'_2 b'_3 & \cdots & A'_2 b'_{16} \\ A_3 b_1 & A_3 b_2 & \delta_3 & \cdots & A'_3 b'_{16} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{16} b_1 & A_{16} b_2 & A_{16} b_3 & \cdots & \delta_{16} \end{bmatrix}$$

Where,

- Diagonal: δ_i for $i = 1, \ldots, 16$
- Lower triangle: $A_i b_j$ for $i > j$
- Upper triangle: $A'_i b'_j$ for $i < j$
- Lower triangle (below diagonal): Each entry M_{ij} for $i > j$ mixes information from past positions using parameters A_i and b_j (causal direction).
- Upper triangle (above diagonal): Each entry M_{ij} for $i < j$ mixes information from future positions using parameters A'_i and b'_j (anti-causal direction).
- Diagonal: Each M_{ii} is a self-connection parameter δ_i.

2. Training and optimization

HectoMixNet is a 3D UNet-based architecture implemented in PyTorch, specifically designed for the segmentation of primary tumors and lymph nodes in head and neck PET/CT scans. The network integrates a bidirectional quasiseparable mixing module at the bottleneck, enabling it to capture both local features through convolutional layers and long-range spatial dependencies across the entire volume. All convolutional layers are initialized using He initialization, and the model uses 3×3×3 convolution kernels across five encoding and five decoding levels, with skip connections preserved to maintain high-resolution spatial information.

For training, input patches of size 160 × 160 × 160 were used, and the AdamW optimizer was applied with a cosine annealing learning rate schedule starting at 0.0002. The loss function combines Dice and Cross-Entropy losses, balancing volumetric overlaps and boundary accuracy. Training was conducted for 1000 epochs on all labeled HECKTOR2025 cases, with a batch size of 2 due to GPU memory limitations. A 5-fold cross-validation scheme (~80% training, 20% validation per fold) was employed for robustness, and final predictions were obtained through ensemble averaging across folds. Extensive data augmentation, including rotations, flipping, scaling, Gaussian blur, and additive noise, was applied to improve generalization. Hyperparameters, such

as learning rate, batch size, and network depth, were tuned via grid search, and each fold required approximately 24 hours of training on a single NVIDIA A6000 GPU.

The advantages of HectoMixNet stem from its combination of local convolutional feature extraction with global context modeling via the bidirectional quasiseparable mixing module. This design enables accurate segmentation of both small, irregular lymph nodes and large, diffuse primary tumors. Coupled with standardized preprocessing, extensive data augmentation, and ensemble predictions, HectoMixNet demonstrates robust generalization and high segmentation fidelity. Its architecture ensures precise boundary delineation, effective detection of multiple lesions, and efficient use of multimodal PET/CT information, making it particularly well-suited for the challenges posed by the HECKTOR 2025 Task 1 dataset.

3 Results

Table 1. Segmentation Performance on HECKTOR2025 Task 1

Model	GTVp Dice	GTVn Aggregate Dice	GTVnAggregate F1
HectoMixNet (Test)	**0.7567**	**0.7567**	**0.6407**
HectoMixNet (Validation)	0.7381	0.8021	0.7015
HectoMixNet (Sanity Check)	0.8812	0.8246	0.7273
3D Mamba (Sanity Check)	0.8831	0.8169	0.6667
3D xLSTM (Sanity Check)	0.8626	0.6937	0.4286

HectoMixNet demonstrates strong segmentation performance on HECKTOR2025 Task 1 as shown in Table.1. On the validation set, it achieved a GTVp Dice score of 0.7381, indicating good overlap with the ground truth for primary tumors, while the GTVn aggregate Dice of 0.8021 and F1 score of 0.7015 reflect accurate identification of lymph nodes. In a sanity check on the leaderboard, HectoMixNet further improved, reaching 0.8812 GTVp Dice, 0.8246 GTVn Dice, and 0.7273 GTVn F1, demonstrating its robustness under different test splits. Compared to other models, 3D Mamba slightly outperformed HectoMixNet on GTVp Dice (0.8831 vs 0.8812), but HectoMixNet achieved higher lymph node segmentation scores (GTVn Dice 0.8246 vs 0.8169, F1 0.7273 vs 0.6667). The 3D xLSTM model, while competitive on primary tumor segmentation, lagged significantly in lymph node detection (GTVn Dice 0.6937, F1 0.4286), highlighting the advantage of HectoMixNet's bidirectional quasiseparable mixing module for capturing long-range spatial context. Overall, HectoMixNet provides a balanced and robust segmentation performance for both primary tumors and lymph nodes, making it particularly well-suited for the HECKTOR2025 Task 1 challenge.

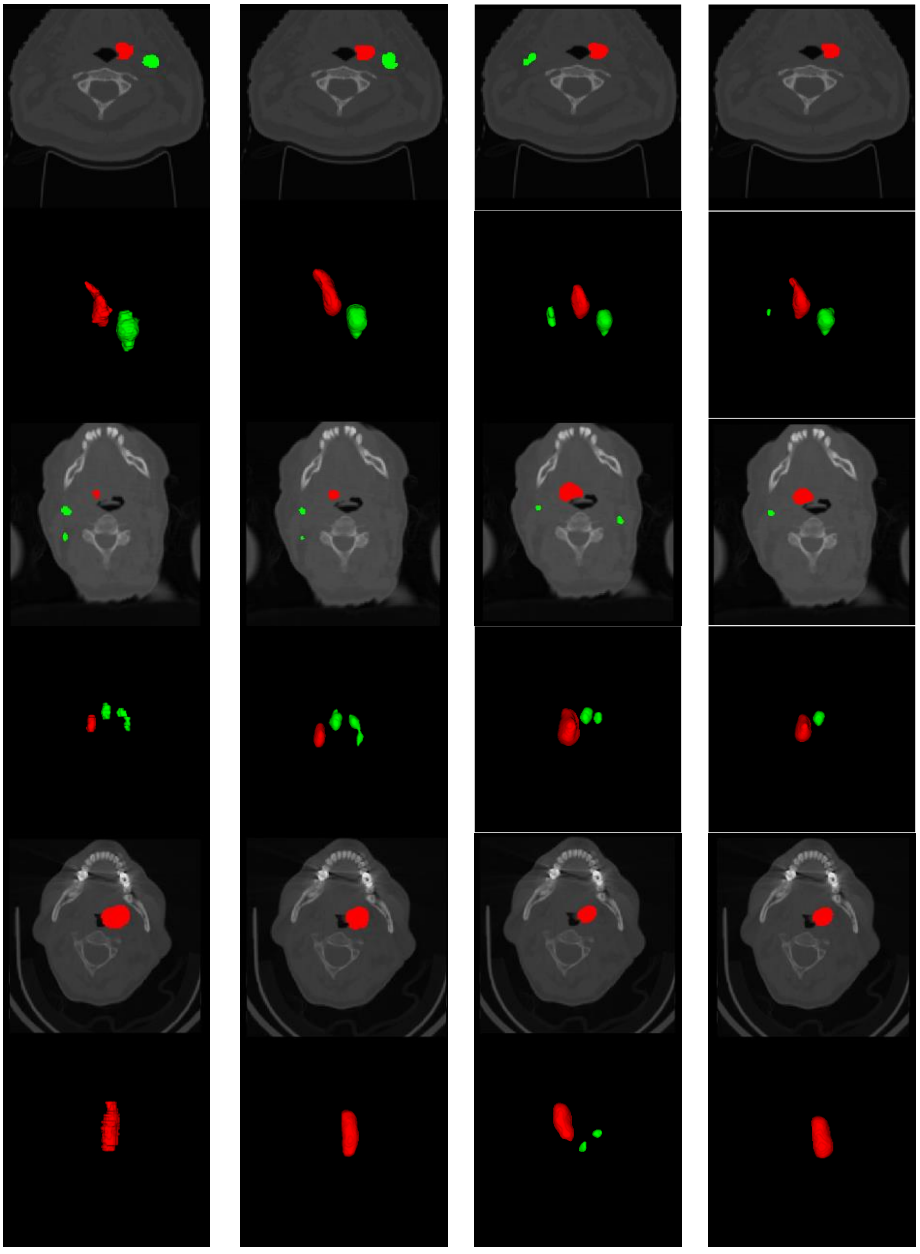

Figure 2. Columns display CT axial overlays with corresponding segmentations. The first column shows the ground truth (GT), followed by predictions from the proposed HectoMixNet, 3D Mamba, and 3D xLSTM. Red highlights the primary tumor volume (GTVp), while green denotes lymph nodes (GTVn).

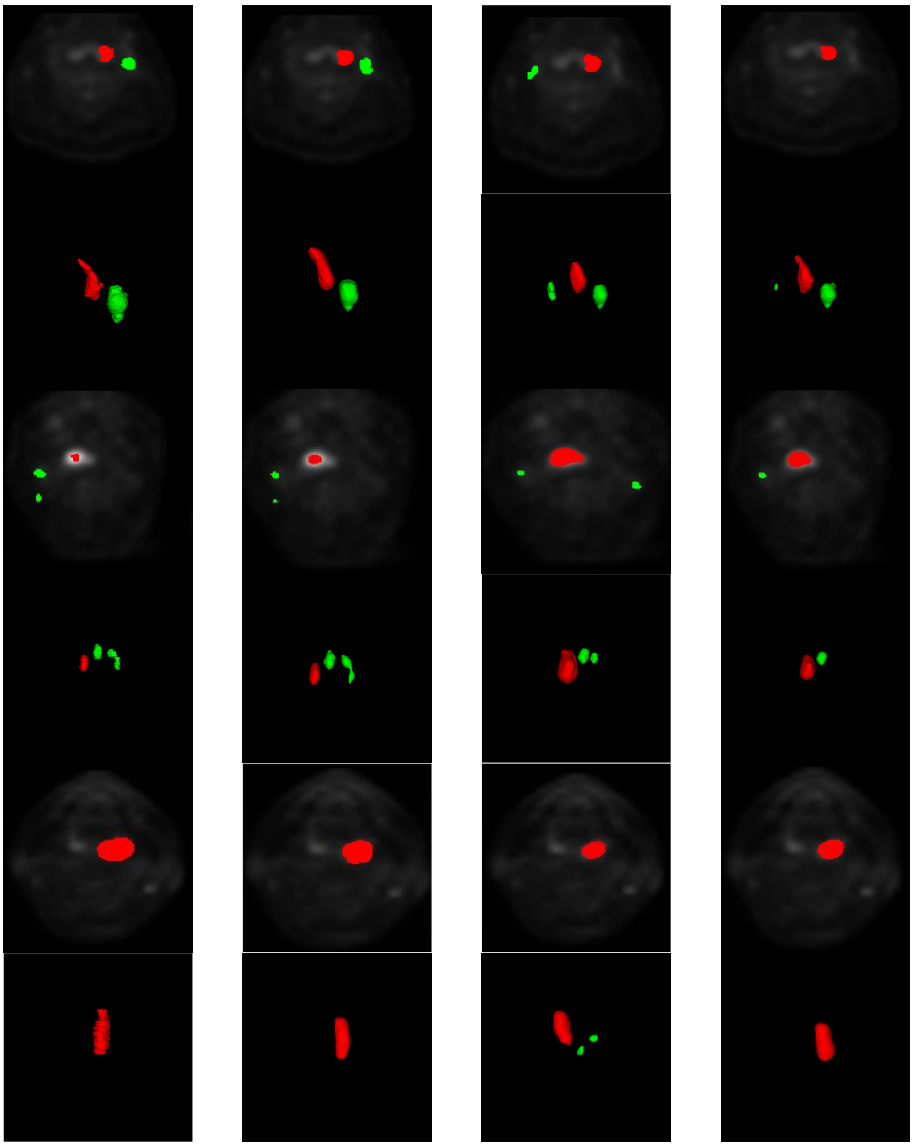

Figure 3. Columns display PET axial overlays with corresponding segmentations. The first column shows the ground truth (GT), followed by predictions from the proposed HectoMixNet, 3D Mamba, and 3D xLSTM. Red indicates the primary tumor volume (GTVp), and green represents lymph nodes (GTVn).

Figures 2 and 3 present segmentation results for three validation subjects from the Hector dataset, comparing ground truth (GT) with predictions from HectoMixNet, 3D Mamba, and 3D xLSTM. Columns show GT (first), HectoMixNet (second), 3D Mamba (third), and 3D xLSTM (fourth), with red representing the primary tumor volume (GTVp) and green indicating lymph nodes (GTVn). Rows show axial slice overlays

(odd) and corresponding 3D volumetric reconstructions (even) for CHUM-014 subject (rows 1–2), HMR-013 subject (rows 3–4), and USZ-003 subject (rows 5–6). Axial slices provide local segmentation detail, highlighting differences in tumor boundary delineation and lymph node detection, particularly for smaller or isolated nodes. 3D views offer volumetric context, demonstrating spatial alignment, tumor coverage, and connectivity of lymph nodes. Overall, HectoMixNet achieves the highest segmentation accuracy, effectively capturing both tumors and small lymph nodes; 3D Mamba provides the second-best performance, showing generally accurate tumor delineation but occasional under-segmentation of smaller nodes; 3D xLSTM ranks third, exhibiting smooth volumetric continuity but less consistent node detection and occasional shape deviations. This figure illustrates both local and global differences in model performance for head and neck tumor and lymph node segmentation.

4 Discussion

4.1 Segmentation Accuracy

HectoMixNet achieved consistently strong segmentation performance across both primary tumors (GTVp) and lymph nodes (GTVn), demonstrating its ability to handle the complex heterogeneity of head and neck cancers. On the validation set, it achieved a GTVp Dice of 0.7381 and GTVn Dice of 0.8021, with further improvements on the leaderboard test set (0.8812 and 0.8246, respectively). Importantly, HectoMixNet maintained balanced performance between tumor and lymph node segmentation, whereas other models tended to favor one task at the expense of the other. The superior voxel-level overlap reflects the effectiveness of bidirectional quasiseparable mixing in capturing both fine lesion boundaries and global anatomical consistency. While 3D Mamba slightly outperformed HectoMixNet on GTVp Dice, our model showed clear advantages in lymph node detection and delineation, a more challenging task given their small size, variable shapes, and spatial dispersion.

4.2 Lesion Detection

Accurate lesion detection is clinically critical for comprehensive treatment planning, as missed lymph nodes can lead to undertreatment. HectoMixNet's ability to capture long-range dependencies translated into robust lesion-wise detection, with a GTVn F1-score of 0.7273 on the leaderboard, surpassing both 3D Mamba (0.6667) and xLSTM (0.4286). This demonstrates that the bidirectional mixing module not only improves volumetric overlap but also enhances the model's ability to identify multiple distinct lesions. By integrating local detail with broader context, HectoMixNet reduces false negatives while maintaining precision, ensuring that clinically relevant lymph node metastases are rarely overlooked.

4.3 Role of State Space Models

Both HectoMixNet and competing SSM-based architectures highlight the value of state space modeling for volumetric medical imaging. Traditional CNNs and recurrent architectures are limited by local receptive fields or vanishing memory, constraining their ability to model the irregular and dispersed nature of H&N lesions. By contrast,

HectoMixNet's bidirectional quasiseparable mixing extends SSM capabilities, enabling efficient propagation of contextual information across the entire 3D volume while preserving local fidelity. The performance gap between HectoMixNet and 3D Mamba underscores that architectural design choices, particularly the integration of self-connections and dual causal/anti-causal processing, are critical for fully exploiting the potential of state space models in medical image segmentation.

4.4 Clinical Implications

From a clinical standpoint, HectoMixNet offers significant advantages. Its high Dice scores and strong lesion-wise detection make it particularly well-suited for radiotherapy planning, where complete and precise delineation of both primary tumors and lymph nodes is essential. The model's robustness across multicenter data also suggests potential for real-world deployment, where imaging protocols and patient characteristics vary widely. Compared to manual segmentation, which is labor-intensive and prone to inter-observer variability, HectoMixNet provides a scalable and reproducible alternative. Importantly, its ability to detect small, spatially separated lymph nodes enhance clinical safety by reducing the likelihood of missed lesions that could compromise treatment outcomes.

4.5 Limitations and Future Work

Despite its promising performance, HectoMixNet has several limitations. First, evaluation was restricted to the HECKTOR 2025 dataset, which, while large and multicentric, may not capture the full variability of clinical practice, such as motion artifacts, rare tumor subtypes, or unusual imaging conditions. Second, the current framework focuses solely on PET/CT; integrating additional modalities such as MRI could provide a richer anatomical and functional context. Third, although quasiseparable mixing is computationally efficient relative to dense global attention, training requirements remain substantial, potentially limiting accessibility for institutions with limited computational resources.

Future directions include extending HectoMixNet for different data fusion, exploring self-supervised pretraining strategies to improve cross-institutional generalization, and developing interpretability mechanisms tailored to state space models to better align automated outputs with clinical decision-making. In addition, longitudinal extensions for outcome prediction could further enhance its value in personalized oncology workflows.

5 Conclusion

In this work, we introduced HectoMixNet, a deep learning framework specifically designed for automatic segmentation of primary tumors and lymph nodes in head and neck cancer patients using multimodal PET/CT imaging. By incorporating a bidirectional quasiseparable mixing module within a 3D encoder–decoder architecture, the model successfully combines fine anatomical detail with efficient global context modeling, enabling robust performance across heterogeneous patient cohorts. Evaluation on the HECKTOR 2025 Task 1 dataset demonstrated that HectoMixNet achieves

competitive segmentation accuracy, with particularly strong performance in lymph node detection and delineation—an area where existing methods often underperform. Compared with state-of-the-art models, HectoMixNet provides a more balanced performance between tumor and lymph node segmentation, underlining the effectiveness of bidirectional quasiseparable mixing in capturing long-range spatial dependencies. Given its scalability and generalization capability, HectoMixNet represents a promising step toward reliable, automated segmentation tools that can reduce inter-observer variability and support precision radiotherapy planning in clinical practice.

References

1. Parkin DM, et al. "Global cancer statistics, 2002." CA: a cancer journal for clinicians 55.2 (2005): 74-108.
2. Bonner, James A., Paul M. Harari, Jordi Giralt, Roger B. Cohen, Christopher U. Jones, Ranjan K. Sur, David Raben, et al. 2010. “Radiotherapy plus Cetuximab for Locoregionally Advanced Head and Neck Cancer: 5-Year Survival Data from a Phase 3 Randomised Trial, and Relation between Cetuximab-Induced Rash and Survival.” The Lancet Oncology 11 (1): 21–28.
3. Chajon E, et al. "Salivary gland-sparing other than parotid-sparing in definitive head-and-neck intensity-modulated radiotherapy does not seem to jeopardize local control." Radiation Oncology 8.1 (2013): 1-9.
4. Bogowicz M, et al. "Comparison of PET and CT radiomics for prediction of local tumor control in head and neck squamous cell carcinoma." Acta Oncologica 56.11 (2017): 1531-1536.
5. Andrearczyk V, et al. "Overview of the HECKTOR Challenge at MICCAI 2021: Automatic Head and Neck Tumor Segmentation and Outcome Prediction in PET/CT Images", in: Head and Neck Tumor Segmentation and Outcome Prediction, 1-37 (2022).
6. Mazher, M., Razzak, I., Qayyum, A., Tanveer, M., Beier, S., Khan, T., & Niederer, S. A. (2024). Self-supervised spatial–temporal transformer fusion based federated framework for 4D cardiovascular image segmentation. Information Fusion, 106, 102256.
7. Wang, Kang, Chen Qin, Zhang Shi, Haoran Wang, Xiwen Zhang, Chen Chen, Cheng Ouyang et al. "Extreme Cardiac MRI Analysis under Respiratory Motion: Results of the CMRxMotion Challenge." arXiv preprint arXiv:2507.19165 (2025).
8. Yang, Kaiyuan, Fabio Musio, Yihui Ma, Norman Juchler, Johannes C. Paetzold, Rami Al-Maskari, Luciano Höher et al. "Benchmarking the cow with the topcow challenge: Topology-aware anatomical segmentation of the circle of willis for cta and mra." ArXiv (2024): arXiv-2312.
9. de la Rosa, Ezequiel, Ruisheng Su, Mauricio Reyes, Roland Wiest, Evamaria O. Riedel, Florian Kofler, Kaiyuan Yang et al. "ISLES'24: Improving final infarct prediction in ischemic stroke using multimodal imaging and clinical data." arXiv preprint arXiv:2408.10966 (2024).
10. Payette, Kelly, Céline Steger, Roxane Licandro, Priscille De Dumast, Hongwei Bran Li, Matthew Barkovich, Liu Li et al. "Multi-center fetal brain tissue annotation (feta) challenge 2022 results." IEEE transactions on medical imaging (2024).

11. Imran, Muhammad, Jonathan R. Krebs, Vishal Balaji Sivaraman, Teng Zhang, Amarjeet Kumar, Walker R. Ueland, Michael J. Fassler et al. "Multi-class segmentation of aortic branches and zones in computed tomography angiography: The aortaseg24 challenge." arXiv preprint arXiv:2502.05330 (2025).
12. Qayyum, Abdul, Hao Xu, Brian P. Halliday, Cristobal Rodero, Christopher W. Lanyon, Richard D. Wilkinson, and Steven Alexander Niederer. "Transforming Heart Chamber Imaging: Self-Supervised Learning for Whole Heart Reconstruction and Segmentation." arXiv preprint arXiv:2406.06643 (2024).
13. Qayyum, Abdul, Imran Razzak, Moona Mazher, Xuequan Lu, and Steven A. Niederer. "Unsupervised unpaired multiple fusion adaptation aided with self-attention generative adversarial network for scar tissues segmentation framework." Information Fusion 106 (2024): 102226.
14. Nan, Yang, Xiaodan Xing, Shiyi Wang, Zeyu Tang, Federico N. Felder, Sheng Zhang, Roberta Eufrasia Ledda et al. "Hunting imaging biomarkers in pulmonary fibrosis: benchmarks of the AIIB23 challenge." Medical Image Analysis 97 (2024): 103253.
15. de la Rosa, Ezequiel, Mauricio Reyes, Sook-Lei Liew, Alexandre Hutton, Roland Wiest, Johannes Kaesmacher, Uta Hanning et al. "A robust ensemble algorithm for ischemic stroke lesion segmentation: Generalizability and clinical utility beyond the isles challenge." arXiv preprint arXiv:2403.19425 (2024).
16. de la Rosa, Ezequiel, Mauricio Reyes, Sook-Lei Liew, Alexandre Hutton, Roland Wiest, Johannes Kaesmacher, Uta Hanning et al. "DeepISLES: a clinically validated ischemic stroke segmentation model from the ISLES'22 challenge." Nature Communications 16, no. 1 (2025): 7357.
17. Zhu, Lianghui, Bencheng Liao, Qian Zhang, Xinlong Wang, Wenyu Liu, and Xinggang Wang. "Vision mamba: Efficient visual representation learning with bidirectional state space model." arXiv preprint arXiv:2401.09417 (2024).
18. Qayyum, Abdul, Chun Kit Ang, S. Sridevi, MKA Ahamed Khan, Lim Wei Hong, Moona Mazher, and Tran Duc Chung. "Hybrid 3D-ResNet deep learning model for automatic segmentation of thoracic organs at risk in CT images." In 2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), pp. 1-5. IEEE, 2020.
19. Saeed, Numan, Salma Hassan, Shahad Hardan, Ahmed Aly, Darya Taratynova, Umair Nawaz, Ufaq Khan et al. "A Multimodal and Multi-centric Head and Neck Cancer Dataset for Segmentation, Diagnosis and Outcome Prediction." arXiv preprint arXiv:2509.00367 (2025).
20. Hwang, Sukjun, Aakash Lahoti, Ratish Puduppully, Tri Dao, and Albert Gu. "Hydra: Bidirectional state space models through generalized matrix mixers." Advances in Neural Information Processing Systems 37 (2024): 110876-110908.
21. Luo, Gongning, Mingwang Xu, Hongyu Chen, Xinjie Liang, Xing Tao, Dong Ni, Hyunsu Jeong et al. "Tumor detection, segmentation and classification challenge on automated 3d breast ultrasound: The tdsc-abus challenge." arXiv preprint arXiv:2501.15588 (2025).