

INTRODUCTION

- Head and neck cancer is the seventh most common malignancy globally, primarily involving tumor of the upper aerodigestive tract. Esophageal cancer is ranked as the eighth most common cancer worldwide. Collectively, these cancers contribute substantially to the global burden of cancer-related morbidity and mortality.
- Accurate delineation, segmentation, and volumetric measurement of tumors are essential for effective oncologic assessment, treatment planning, and prognostication.
- Precise tumor imaging is critical for radiotherapeutic targeting and surgical planning; however, manual delineation is labor-intensive, time-consuming, and subject to inter- and intra-observer variability.
- Machine learning (ML), particularly deep learning models such as convolutional neural networks (CNNs) and transformer-based models, has demonstrated significant potential for automated segmentation, offering enhancements in efficiency, reproducibility, and accuracy.
- Emerging evidence suggest robust performance of ML models in delineating tumors and metastatic nodes, with increasing generalizability across diverse imaging modalities (e.g., CT, PET, MRI).

STUDY OBJECTIVE

This scoping review examines the current evidence on the application and performance of ML models for tumor segmentation, delineation, and volumetric quantification in head and neck and esophageal cancers.

METHODS

- Lit Review

 - PubMed (January, 2020 – May, 2025).
 - English language articles.
- I/E Criteria

 - *Included:* Tumor segmentation, delineation and volume measurement, oral cavity, pharyngeal, laryngeal, and esophageal cancers.
 - *Excluded:* Skull bone, thyroid, paranasal cavity, parathyroid, myeloma, and lymphoma.
- Data Extracted

 - Cancer type, ML method, imaging modality, segmentation formula and performance.
 - *Metrics (Acceptable Range):* DSC (≥ 0.70), HD95 (≤ 10 mm), AUC (≥ 0.75), ASD (≤ 2 mm), ASSD (≤ 2 mm), IoU (≥ 0.70).

ABBREVIATIONS

CT: Computerized Tomography	DSC: Dice Similarity Coefficient
PET: Positron Emission Tomography	HD95: Hausdorff Distance
MRI: Magnetic Resonance Imaging	AUC: Area Under the Curve
NPC: Nasopharyngeal Cancer	ASD: Average Surface Distance
ESCC: Esophageal Squamous Cell Carcinoma	ASSD: Average Symmetric Surface Distance
OPSCC: Oropharyngeal Squamous Cell Carcinoma	IoU: Intersection over Union

RESULTS

- A total of 26 studies were identified, and 20 studies were eligible based on the inclusion/exclusion criteria.
- *Cancer Types Reported:* Nasopharyngeal carcinoma (n=12), esophageal squamous cell carcinoma (n=5), oropharyngeal squamous cell carcinoma (n=2), and laryngeal carcinoma (n=1).
- *Imaging Modalities:* MRI (n=12), CT (n=7), PET/CT (n=3), and endoscopy (n=1) with several studies using multi-modal imaging (n=6).
- *Machine Learning Models:* Convolutional neural networks (n=10), attention-based models (n=6), transformer-guided architectures (n=2), and multi-branch/domain adaptation frameworks (n=2).
- *Segmentation:* DSC range 0.67–0.876; best-performing models (GloD-LoATUNet, NPCNet, SICNet) achieved DSC >0.83 and HD95 as low as 3.7 mm. *Classification* (OPSCC extranodal extension): AUC up to 0.86, outperforming radiologists in some subgroups. *Endoscopy* (laryngeal cancer): DSC = 0.83, IoU = 0.83, real-time inference (~25 fps).
- Overall, 19 out of 20 studies reported adequate performance (DSC ≥ 0.70 , AUC ≥ 0.75 , IoU ≥ 0.70 , or HD ≤ 10).

Study	Cancer Type	Imaging Modality	Segmentation/Prediction Formula	Performance
Ke et al. (2020)	NPC	MRI	Dice	DSC: 0.77 ± 0.07
Cai et al. (2021)	NPC	MRI	Dice, ASSD	DSC: 0.841 ± 0.011
Wong et al. (2021)	NPC	MRI (non-contrast)	Dice, ASD	DSC: 0.79; ASD: 0.66 mm
Yousefi et al. (2021)	ESCC	CT	Dice, HD95	DSC: 0.79±0.20
Qi et al. (2021)	NPC	CT + MRI	Dice	DSC: 0.719; Accuracy: 0.88
Zhang et al. (2022)	NPC	MRI	Dice	DSC: 0.816
Liao et al. (2022)	NPC	MRI	Dice	DSC: 0.83 (GTVnx), 0.80 (GTVnd)
Ye et al. (2022)	ESCC	CT + PET/CT	Dice, HD95, ASD	DSC: 0.78-0.81
Li et al. (2022)	NPC	MRI	Dice	DSC: 0.73 ± 0.21
Liu et al. (2022)	NPC	MRI	Dice	DSC: 0.81
Yue et al. (2023)	ESCC	PET/CT	Dice, HD	DSC: 0.83-0.86
Meng et al. (2023)	NPC	CT + MRI	Dice, F1-score	DSC: 77.6% ± 6.75
Kann et al. (2023)	OPSCC	CT	AUC, Sensitivity, Specificity	AUC: 0.86; Sens: 90%; Spec: 70-93%
Sampieri et al. (2024)	Laryngeal CA	WL/NBI Endoscopy	Dice, IoU, Accuracy	DSC: 0.83, IoU: 0.83, Accuracy: 0.97
Zhang et al. (2024)	ESCC	CT	Dice, ASD, HD95	DSC: 0.865-0.876
Yuan et al. (2024)	ESCC	PET/CT	AUC	AUC: 0.955 (int), 0.916 (ext)
Luo et al. (2024)	NPC	MRI	Dice, HD95	DSC: 0.70-0.86
Huang et al. (2024)	NPC	DCE-MRI + Ktrans	Dice	DSC: 67.39 ± 15.79
Zhang et al. (2024)	NPC	MRI	Dice, HD95, ASSD	DSC: 74.38 ± 11.99; HD95: 9.31 mm;
Hughes et al. (2025)	OPSCC	CT	Sensitivity, Specificity, AUC	Sens: 41%, Spec: 96%, AUC: 0.75

DISCUSSION

- Machine learning-based imaging models have demonstrated promising and generally acceptable performance in the segmentation, delineation, and volumetric quantification of tumors in head and neck and esophageal cancers.
- Despite encouraging results, the heterogeneity in study design, evaluation metrics, and reporting standards limits direct comparison and clinical translation.
- Validation studies and standardized methodological frameworks are essential to support clinical integration.