

DL Techniques and Imaging in Head and Neck

Subjects: Otorhinolaryngology

Contributor: Christos Tsilivigkos, Michail Athanasopoulos, Riccardo di Micco, Aris Giotakis, Nicholas S. Mastronikolis, Francesk Mulita, Georgios-Ioannis Verras, Ioannis Maroulis, Evangelos Giotakis

Deep learning (DL) systems utilize complex algorithms and neural networks featuring numerous intricate layers in order to make decisions and solve advanced problems. Their application in medicine, and specifically in otorhinolaryngology has increased rapidly. The head and neck region is among the most common locations for cancer, with a substantial occurrence of lymph node involvement and metastases observed in both nearby and distant regions.

Keywords: deep learning ; artificial intelligence ; convolutional neural network

1. Head and Neck Imaging

Head and neck surgery relies majorly on imaging, which is often a pre-requisite before any further management. Different techniques offer significant advantages in disease diagnosis but also follow-up. Computed tomography (CT) and magnetic resonance imaging (MRI), during the last decades, have usually been used combinatorically in a large variety of medical conditions to acquire both bone and soft tissue information.

Lately, deep learning algorithms have emerged, which enable the conversion of one imaging modality to another. For example, MRI scans, which involve bone techniques, give us the possibility of a subsequent MRI to CT reconstruction, avoiding exposure to ionizing energy and aiding non-experts in diagnosis at the same time ^[1]. A combination of two generative adversarial networks has also been implemented to generate accurate synthetic CT images from MRI scans ^[2]. On the other hand, non-contrast CT scans can be converted to PET-like images with generative models, eliminating the need for radioactive tracers. The generated PET images demonstrate comparable accuracy to actual FDG-PET images in predicting clinical outcomes ^[3]. It seems rational to hypothesize that such deep learning pipelines can transform head and neck imaging into a one-step-procedure in the future.

Next, CNNs are believed to exhibit superior performance compared to a traditional radiomic framework regarding their ability to detect image patterns, often undetectable by the latter, while systems such as ultra-high-resolution CT with a DL-based image reconstruction engine offer significant amelioration in subjective and objective image quality, with a higher signal-to-noise ratio, lower noise, and lower radiation exposure ^[4].

The DL technique utilized in the analysis of medical images allows the incorporation of both qualitative and quantitative imaging characteristics to create prediction models characterized by exceptional diagnostic accuracy. These principles have been applied generally to HNSCC imaging, but also specifically to specific types of HNSCC. Notably, in the imaging of oral and oropharyngeal cancer, FDG-PET/CT scans can be processed by DL systems to predict local treatment outcomes ^[5], disease-free survival with high sensitivity and specificity ^[6], overall survival ^[7], and they can even assist in differentiating human papillomavirus positive from human papillomavirus negative oropharyngeal carcinomas ^[8].

At the same time, progress in computer vision and deep learning provide potent techniques for creating supplementary tools capable of automatically screening the oral cavity. These cost-effective and non-invasive tools can offer real-time insights for healthcare practitioners during patient assessments and can also facilitate self-examinations for individuals. The automated diagnosis of oral cancer through images is predominantly focused on the utilization of specialized imaging technologies, namely optical coherence tomography ^{[9][10]}, hyperspectral imaging ^[11], and autofluorescence imaging ^[12], but also white-light photographs ^[13]. Such DL techniques can come in the form of mHealth applications, assisting in oral and oropharyngeal lesion detection in both hospitals and resource-limited areas, and enabling telediagnosis ^[14]. Finally, systems offering a real-time estimation of cancer risk and biopsy assistance maps on the oral mucosa are very promising ^[15].

Furthermore, diseases of the nasopharynx have been an area of focus during the last years for DL system developers. From MRI-based applications focusing on the differential diagnosis between benign and malignant nasopharyngeal

diseases [16][17] to the automatic detection of pathological lymph nodes and assessment of the peritumoral area in nasopharyngeal carcinoma, DL algorithms can significantly assist in disease prognosis and treatment planning [18]. Interestingly, peritumoral information, especially the largest areas of tumor invasion, has been shown to provide valuable insights for distant metastasis prediction in individuals with nasopharyngeal carcinoma [19].

Imaging of the salivary glands constitutes another significant challenge for radiologists and otolaryngologists, who have many different imaging modalities in their quiver. Specialized DL algorithms have been developed to assist in differential diagnosis between benign and malignant parotid gland tumors in contrast-enhanced CT images [20], and ultrasonography [21]. MRI remains the gold standard in the diagnosis of salivary gland diseases, where DL models intend to automatically classify salivary gland tumors with very high accuracy [22][23].

Relative to thyroid disease diagnosis, ultrasound (US) is widely acknowledged as the primary diagnostic technique for examining thyroid nodules and assessing papillary thyroid carcinomas (PTCs) before surgery [24]. DL networks with excellent diagnostic efficiency have been deployed to distinguish between benign nodules and thyroid carcinoma [25], improve the detection of follicular carcinoma, differentiate between atypical and typical medullary carcinoma [26], and assess for gross extrathyroidal extension in thyroid cancer [27]. AI systems can be very useful in eliminating the operator dependence of US and ameliorating diagnosis precision, especially in inexperienced radiologists.

Nevertheless, plenty of other DL techniques are associated with thyroid gland evaluation. Thus, apart from thyroid gland contouring in non-contrast-enhanced CT images [28], special applications used intraoperatively to assist surgeons in recurrent laryngeal nerve [29] and parathyroid gland identification have been designed. Such algorithms have the potential to improve surgical workflows in the intricate environment of open surgery.

The head and neck region is among the most common locations for cancer, with a substantial occurrence of lymph node involvement and metastases observed in both nearby and distant regions. The identification of distant metastases is linked to an unfavorable prognosis, often resulting in a median survival period of around 10 months [30]. The role of imaging in metastasis diagnosis is uncontroversial and novel convolutional neural networks have been developed in this direction. For example, extended 2D-CNN and 3D-CNN models have been deployed to perform time-to-event analysis for the binary classification of distant metastasis in head and neck cancer patients. These models result in the generation of distant metastasis-free probability curves and stratify patients into high- and low-risk groups [31]. CNN are generally able to detect image patterns that can be untraceable with traditional methods. Thus, it has been shown that CNN can be trained to forecast the treatment results for individuals with HNSCC, relying exclusively on the information from CT scans conducted prior to treatment [32].

CNN assessing pre-treatment MRI scans to predict the possibility of distant metastases in individuals with nasopharyngeal carcinoma can also be useful, since the occurrence of a metastasis is the main reason for radiotherapy failure in this patient group. Predicting the high risk for distant metastasis in a patient can lead to a more aggressive treatment approach [19]. Moreover, pre-therapy MRI scans have been used in patients with advanced (T3N1M0) nasopharyngeal carcinoma to guide the clinicians in deciding between induction chemotherapy plus concurrent chemoradiotherapy or concurrent chemoradiotherapy alone [33].

DL models diagnosing lymph node involvement can boost clinical decision-making in the future. A relative model has been developed that detects pathological lymph nodes in individuals with oral cancer [34], while another one predicts lymph node involvement in patients with thyroid cancer through the interpretation of their multiphase dual-energy spectral CT images [35].

The utilization of deep learning techniques allows for the complete automation of image analysis providing the user with multiple possibilities (Table 1). Nevertheless, it demands a substantial volume of accurately labeled images. Additionally, prediction-making necessitates detailed patient endpoint data, a process that is both expensive and time-intensive. Developing more effective models with constrained datasets stands as a critical challenge in the field of AI today.

Table 1. The contributions of deep learning systems in head and neck imaging and radiotherapy.

Deep Learning Contributions in Head and Neck	
Imaging	Generation of an imaging modality from another
	Prediction making based on imaging
	Automated diagnosis of malignant and benign diseases

Deep Learning Contributions in Head and Neck	
	Automated diagnosis of pathological lymph nodes
	Automated diagnosis of metastases
	Analysis of specific tumor characteristics
	Contouring of significant structures
	Cancer risk assessment of a lesion
	Biopsy assistance mapping
	Intraoperative surgeon assistance
	Auto-segmentation of structures based on imaging
	Automation of the procedure
	Automated clinical target volume contouring
Radiotherapy	
Endoscopy and laryngoscopy	Image quality improvement
	Segmentation of images
	Optical detection
	Pathological pattern detection
	Endoscopic image classification
	Lesion histology (benign/malignant) prediction
	Self-screening tumor recurrence detection
	Intra-operative endoscopic lesion detection
	Anatomical structure and lesion automatic segmentation
	Automatic assessment of aspiration and dysphagia
	Evaluation of laryngeal mobility

2. Head and Neck Radiotherapy

Radiotherapy (RT) stands as a fundamental pillar in head and neck cancer (HNC) treatment, whether administered independently, post-surgery, or concurrently with chemotherapy. Defining organs at risk (OARs) and clinical target volumes represents a crucial phase in the treatment protocol. This process typically involves manual work, is time-consuming, and necessitates substantial training. Ideally, these tasks would be substituted by automated procedures requiring minimal clinician involvement, and AI appears competent to undertake this role.

A major challenge and the primary drawback of radiation therapy is that, apart from the cancerous mass, it unavoidably exposes nearby healthy tissues, known as OARs, to some level of radiation. This can potentially result in various adverse effects and toxicities, since contouring organs like the parotid and the submandibular gland and excluding them from radiation intake can be quite arduous ^[36]. Additionally, DL-based automated segmentation of the masticatory area has successfully reduced the incidence of RT-associated trismus ^[37].

Several applications aiming to realize normal tissue structure auto-segmentation from CT images ^{[38][39]} exist. These can include three-dimensional segmentation models and convolutional neural networks for final OAR identification ^[40]. DL pipelines focusing on tumor segmentation in specific organs, such as the oropharynx ^[41], and the salivary glands promise to gradually automatize the RT procedure, and at the same time reduce post-segmentation editing ^[42].

On the other hand, 3D CNNs aim to consistently and precisely generate clinical target volumes contouring for the different lymph node levels in HNSCC RT ^{[43][44]}. Such applications show quicker contouring adjustments in comparison to automated delineations, closely aligning with corrected delineations for specific levels, and reducing interobserver variability.

The possibilities that DL systems offer are countless, with distant metastasis and overall survival prediction in HNSCC using PET-only models without gross tumor volume segmentation ^[45], or automatically delineating the gross tumor volume

in the FDG-PET/CT images of HNSCC patients ^[46]. Overall, DL systems present the potential to offer personalized RT guidance in HNSCC patients, with limited contribution from medical experts.

3. Endoscopy and Laryngoscopy

Machine learning has been recently experimentally applied to diagnostic ENT endoscopy to leverage meaningful information from digital images, videos, and other visual inputs and take actions or make recommendations based on that information. Mediated from the early experience acquired in the more standardized field of gastrointestinal endoscopy, AI-based video analysis, or videomics ^[47], has been variously applied to improve automatic image quality control, classification, optical detection, and the segmentation of images. After numerous proof-of-concept studies, videomics is rapidly moving to viable clinical approaches for detecting pathological patterns in real-time assistance during the endoscopic evaluation of the upper aerodigestive ways.

A deep learning model consists of complex multilayer artificial neural networks, among which convoluted neural networks are the most popular in the image analysis field. The CNN does not require instructions on which features describe an object and can autonomously learn how to identify it by observing a sufficient number of examples. Various available AI models exist and have been applied ^[48], although a specific comparison between the various algorithm architectures for the task is still lacking. After this preliminary conceptualization phase, the model undergoes a supervised learning session, in which expert otolaryngologists provide the AI human annotated images to transfer their ability in recognizing the lesions. The higher the quality and quantity of items in the validation set, the more accurate the model will be. After the training validation set, the performance of the system is measured on the testing set by comparing the model prediction with the original human annotations. The performance will be evaluated using diagnostic metrics relative to the task analyzed.

AI can be used to classify endoscopic images. In that case, the diagnostic metrics of interest are accuracy (percentage of correctly classified images), precision (positive predictive value), and sensitivity (percentage of correctly identified images compared to all the ones that should have been recognized); F1 score (harmonic mean of precision and sensitivity); and the receiver operating characteristic curve (graphically identifying the true positive rate against the false positive one) ^[49]. In this framework, it is possible to apply AI to classify videos based on their image quality, selecting only the most informative frames for further analysis ^{[50][51]}. Another classification task is the optical biopsy ^[52], predicting the histology of a lesion based on its appearance. At the current state, AI is more accurate in binary classification, e.g., premalignant/malignant ^[53], whereas it loses diagnostic power in multiclass operation ^[54]. By expanding and diversifying the validation dataset, it is possible to achieve high accuracy in simultaneously identifying different conditions such as glottic carcinoma, leucoplakia, nodules, and vocal cord polyps ^[55], outperforming other approaches according to AUC and F1 otolaryngologist trainees ^[56].

Another task the AI is devised for is the automatic detection of lesions during endoscopic evaluation. The main diagnostic metrics for this function are the F1 score, the intersection over union (how well the selected area overlaps with the original annotated area), and the mean average precision (precision and sensitivity according to the chosen IoU). Using narrow band images, AI can be trained to localize mucosal cancerous lesions in the pharynx and larynx during endoscopy ^{[57][58]} ^[59]. This concept has been recently applied to automatically detect laryngeal cancer in real time video-laryngoscopy using the open-source YOLO CNN, achieving 67% precision, 62% sensitivity, and 0.63 mean average precision at 0.5 IoU ^[60], which could be implemented in a self-screening approach for early tumor recurrence detection ^[61]. Based on simple diagnostic endoscopy, the same approach can be applied intraoperatively to detect pathological tissues, such as in endoscopic parathyroid surgery ^{[62][63]}.

Finally, CNN has been used to automatically delineate the boundaries of anatomical structure and lesions in the upper aerodigestive ways. Segmentation performance is evaluated with IoU and the dice similarity coefficient (similarity between the predicted segmentation mask and the ground truth mask). The rationale of segmentation in videomics is to improve lesion follow-up, the definition of tumor resection margins in the operation room, and the area of interest for general laryngology. The automated segmentation of cancer tissue has been successfully attempted in the nasopharynx (DSC 0.78) ^[64], oropharynx (DSC 0.76) ^[65], and laryngeal lesion (DSC 0.814) ^[66]. Aside from cancer pathology, segmentation may be used to select the region of interest for automated functional laryngeal analysis, such as the identification of the glottis angle ^{[67][68]}, glottal midline ^[69], vocal cord paralysis ^[70], postintubation granuloma ^[71], vocal cord dynamics ^{[72][73]}, or in the endoscopic evaluation of aspiration and penetration risk in dysphagia (FESS-CAD, DSC 0.92.5) ^[74].

Building a sufficiently large and heterogeneous training image dataset is a necessary task required to improve the deep learning-based image classifier. The main obstacles remain the lack of standardization of endoscopic techniques and

study structures, hampering a comparison between the different experiences, and the complex anatomy of the upper aerodigestive ways, making image acquisition and standardization difficult. Although deep learning models can be very good at analyzing images belonging to the same group of the training cohort, they may lack accuracy when tested on different populations. To effectively apply videomics in real world situations, future research should focus on validating the trained models with an external dataset, acquired in different institutions and thus being diverse in terms of acquisition technique and population demographics. Although AI-aided endoscopy is still in a preclinical state, the results are promising and may soon efficiently assist the otolaryngologist in many tasks, such as the quality assessment of endoscopic examination, detection of mucosal lesions during endoscopy, optical biopsy of selected lesions, segmentation of cancer margins, and the assessment of laryngeal mobility.

References

1. Bambach, S.; Ho, M.-L. Deep Learning for Synthetic CT from Bone MRI in the Head and Neck. *AJNR Am. J. Neuroradiol.* 2022, 43, 1172–1179.
2. Klages, P.; Benslimane, I.; Riyahi, S.; Jiang, J.; Hunt, M.; Deasy, J.O.; Veeraraghavan, H.; Tyagi, N. Patch-Based Generative Adversarial Neural Network Models for Head and Neck MR-Only Planning. *Med. Phys.* 2020, 47, 626–642.
3. Chandrashekar, A.; Handa, A.; Ward, J.; Grau, V.; Lee, R. A Deep Learning Pipeline to Simulate Fluorodeoxyglucose (FDG) Uptake in Head and Neck Cancers Using Non-Contrast CT Images without the Administration of Radioactive Tracer. *Insights Imaging* 2022, 13, 45.
4. Altmann, S.; Abello Mercado, M.A.; Ucar, F.A.; Kronfeld, A.; Al-Nawas, B.; Mukhopadhyay, A.; Booz, C.; Brockmann, M.A.; Othman, A.E. Ultra-High-Resolution CT of the Head and Neck with Deep Learning Reconstruction—Assessment of Image Quality and Radiation Exposure and Intraindividual Comparison with Normal-Resolution CT. *Diagnostics* 2023, 13, 1534.
5. Fujima, N.; Andreu-Arasa, V.C.; Meibom, S.K.; Mercier, G.A.; Salama, A.R.; Truong, M.T.; Sakai, O. Deep Learning Analysis Using FDG-PET to Predict Treatment Outcome in Patients with Oral Cavity Squamous Cell Carcinoma. *Eur. Radiol.* 2020, 30, 6322–6330.
6. Fujima, N.; Andreu-Arasa, V.C.; Meibom, S.K.; Mercier, G.A.; Truong, M.T.; Hirata, K.; Yasuda, K.; Kano, S.; Homma, A.; Kudo, K.; et al. Prediction of the Local Treatment Outcome in Patients with Oropharyngeal Squamous Cell Carcinoma Using Deep Learning Analysis of Pretreatment FDG-PET Images. *BMC Cancer* 2021, 21, 900.
7. Cheng, N.-M.; Yao, J.; Cai, J.; Ye, X.; Zhao, S.; Zhao, K.; Zhou, W.; Nogues, I.; Huo, Y.; Liao, C.-T.; et al. Deep Learning for Fully Automated Prediction of Overall Survival in Patients with Oropharyngeal Cancer Using FDG-PET Imaging. *Clin. Cancer Res.* 2021, 27, 3948–3959.
8. Fujima, N.; Andreu-Arasa, V.C.; Meibom, S.K.; Mercier, G.A.; Truong, M.T.; Sakai, O. Prediction of the Human Papillomavirus Status in Patients with Oropharyngeal Squamous Cell Carcinoma by FDG-PET Imaging Dataset Using Deep Learning Analysis: A Hypothesis-Generating Study. *Eur. J. Radiol.* 2020, 126, 108936.
9. Yuan, W.; Cheng, L.; Yang, J.; Yin, B.; Fan, X.; Yang, J.; Li, S.; Zhong, J.; Huang, X. Noninvasive Oral Cancer Screening Based on Local Residual Adaptation Network Using Optical Coherence Tomography. *Med. Biol. Eng. Comput.* 2022, 60, 1363–1375.
10. Wilder-Smith, P.; Lee, K.; Guo, S.; Zhang, J.; Osann, K.; Chen, Z.; Messadi, D. In Vivo Diagnosis of Oral Dysplasia and Malignancy Using Optical Coherence Tomography: Preliminary Studies in 50 Patients. *Lasers Surg. Med.* 2009, 41, 353–357.
11. Jeyaraj, P.R.; Samuel Nadar, E.R. Computer-Assisted Medical Image Classification for Early Diagnosis of Oral Cancer Employing Deep Learning Algorithm. *J. Cancer Res. Clin. Oncol.* 2019, 145, 829–837.
12. Song, B.; Sunny, S.; Uthoff, R.D.; Patrick, S.; Suresh, A.; Kolar, T.; Keerthi, G.; Anbarani, A.; Wilder-Smith, P.; Kuriakose, M.A.; et al. Automatic Classification of Dual-Modality, Smartphone-Based Oral Dysplasia and Malignancy Images Using Deep Learning. *Biomed. Opt. Express* 2018, 9, 5318–5329.
13. Fu, Q.; Chen, Y.; Li, Z.; Jing, Q.; Hu, C.; Liu, H.; Bao, J.; Hong, Y.; Shi, T.; Li, K.; et al. A Deep Learning Algorithm for Detection of Oral Cavity Squamous Cell Carcinoma from Photographic Images: A Retrospective Study. *EClinicalMedicine* 2020, 27, 100558.
14. Birur, N.P.; Song, B.; Sunny, S.P.; Mendonca, P.; Mukhia, N.; Li, S.; Patrick, S.; AR, S.; Imchen, T.; Leivon, S.T.; et al. Field Validation of Deep Learning Based Point-of-Care Device for Early Detection of Oral Malignant and Potentially Malignant Disorders. *Sci. Rep.* 2022, 12, 14283.

15. Coole, J.B.; Brenes, D.; Mitbender, R.; Vohra, I.; Hou, H.; Kortum, A.; Tang, Y.; Maker, Y.; Schwarz, R.A.; Carns, J.; et al. Multimodal Optical Imaging with Real-Time Projection of Cancer Risk and Biopsy Guidance Maps for Early Oral Cancer Diagnosis and Treatment. *J. Biomed. Opt.* 2023, 28, 016002.
16. Li, S.; Hua, H.-L.; Li, F.; Kong, Y.-G.; Zhu, Z.-L.; Li, S.-L.; Chen, X.-X.; Deng, Y.-Q.; Tao, Z.-Z. Anatomical Partition-Based Deep Learning: An Automatic Nasopharyngeal MRI Recognition Scheme. *J. Magn. Reson. Imaging* 2022, 56, 1220–1229.
17. Ji, L.; Mao, R.; Wu, J.; Ge, C.; Xiao, F.; Xu, X.; Xie, L.; Gu, X. Deep Convolutional Neural Network for Nasopharyngeal Carcinoma Discrimination on MRI by Comparison of Hierarchical and Simple Layered Convolutional Neural Networks. *Diagnostics* 2022, 12, 2478.
18. Li, S.; Wan, X.; Deng, Y.-Q.; Hua, H.-L.; Li, S.-L.; Chen, X.-X.; Zeng, M.-L.; Zha, Y.; Tao, Z.-Z. Predicting Prognosis of Nasopharyngeal Carcinoma Based on Deep Learning: Peritumoral Region Should Be Valued. *Cancer Imaging* 2023, 23, 14.
19. Hua, H.-L.; Deng, Y.-Q.; Li, S.; Li, S.-T.; Li, F.; Xiao, B.-K.; Huang, J.; Tao, Z.-Z. Deep Learning for Predicting Distant Metastasis in Patients with Nasopharyngeal Carcinoma Based on Pre-Radiotherapy Magnetic Resonance Imaging. *Comb. Chem. High. Throughput Screen.* 2023, 26, 1351–1363.
20. Shen, X.-M.; Mao, L.; Yang, Z.-Y.; Chai, Z.-K.; Sun, T.-G.; Xu, Y.; Sun, Z.-J. Deep Learning-Assisted Diagnosis of Parotid Gland Tumors by Using Contrast-Enhanced CT Imaging. *Oral. Dis.* 2022.
21. Tu, C.-H.; Wang, R.-T.; Wang, B.-S.; Kuo, C.-E.; Wang, E.-Y.; Tu, C.-T.; Yu, W.-N. Neural Network Combining with Clinical Ultrasonography: A New Approach for Classification of Salivary Gland Tumors. *Head. Neck* 2023, 45, 1885–1893.
22. Liu, X.; Pan, Y.; Zhang, X.; Sha, Y.; Wang, S.; Li, H.; Liu, J. A Deep Learning Model for Classification of Parotid Neoplasms Based on Multimodal Magnetic Resonance Image Sequences. *Laryngoscope* 2023, 133, 327–335.
23. Gunduz, E.; Alçin, O.F.; Kizilay, A.; Yildirim, I.O. Deep Learning Model Developed by Multiparametric MRI in Differential Diagnosis of Parotid Gland Tumors. *Eur. Arch. Otorhinolaryngol.* 2022, 279, 5389–5399.
24. Guan, Q.; Wang, Y.; Du, J.; Qin, Y.; Lu, H.; Xiang, J.; Wang, F. Deep Learning Based Classification of Ultrasound Images for Thyroid Nodules: A Large Scale of Pilot Study. *Ann. Transl. Med.* 2019, 7, 137.
25. Yang, Z.; Yao, S.; Heng, Y.; Shen, P.; Lv, T.; Feng, S.; Tao, L.; Zhang, W.; Qiu, W.; Lu, H.; et al. Automated Diagnosis and Management of Follicular Thyroid Nodules Based on the Devised Small-Datasets Interpretable Foreground Optimization Network Deep Learning: A Multicenter Diagnostic Study. *Int. J. Surg.* 2023, 109, 2732–2741.
26. Zhang, R.; Yi, G.; Pu, S.; Wang, Q.; Sun, C.; Wang, Q.; Feng, L.; Liu, X.; Li, Z.; Niu, L. Deep Learning Based on Ultrasound to Differentiate Pathologically Proven Atypical and Typical Medullary Thyroid Carcinoma from Follicular Thyroid Adenoma. *Eur. J. Radiol.* 2022, 156, 110547.
27. Qi, Q.; Huang, X.; Zhang, Y.; Cai, S.; Liu, Z.; Qiu, T.; Cui, Z.; Zhou, A.; Yuan, X.; Zhu, W.; et al. Ultrasound Image-Based Deep Learning to Assist in Diagnosing Gross Extrathyroidal Extension Thyroid Cancer: A Retrospective Multicenter Study. *EClinicalMedicine* 2023, 58, 101905.
28. He, X.; Guo, B.J.; Lei, Y.; Tian, S.; Wang, T.; Curran, W.J.; Zhang, L.J.; Liu, T.; Yang, X. Thyroid Gland Delineation in Noncontrast-Enhanced CTs Using Deep Convolutional Neural Networks. *Phys. Med. Biol.* 2021, 66, 055007.
29. Gong, J.; Holsinger, F.C.; Noel, J.E.; Mitani, S.; Jopling, J.; Bedi, N.; Koh, Y.W.; Orloff, L.A.; Cernea, C.R.; Yeung, S. Using Deep Learning to Identify the Recurrent Laryngeal Nerve during Thyroidectomy. *Sci. Rep.* 2021, 11, 14306.
30. Pisani, P.; Airoldi, M.; Allais, A.; AluffiValletti, P.; Battista, M.; Benazzo, M.; Briatore, R.; Cacciola, S.; Cocuzza, S.; Colombo, A.; et al. Metastatic Disease in Head & Neck Oncology. *Acta Otorhinolaryngol. Ital.* 2020, 40 (Suppl. 1), S1–S86.
31. Lombardo, E.; Kurz, C.; Marschner, S.; Avanzo, M.; Gagliardi, V.; Fanetti, G.; Franchin, G.; Stancanella, J.; Corradini, S.; Niyazi, M.; et al. Distant Metastasis Time to Event Analysis with CNNs in Independent Head and Neck Cancer Cohorts. *Sci. Rep.* 2021, 11, 6418.
32. Diamant, A.; Chatterjee, A.; Vallières, M.; Shenouda, G.; Seuntjens, J. Deep Learning in Head & Neck Cancer Outcome Prediction. *Sci. Rep.* 2019, 9, 2764.
33. Zhong, L.; Dong, D.; Fang, X.; Zhang, F.; Zhang, N.; Zhang, L.; Fang, M.; Jiang, W.; Liang, S.; Li, C.; et al. A Deep Learning-Based Radiomic Nomogram for Prognosis and Treatment Decision in Advanced Nasopharyngeal Carcinoma: A Multicentre Study. *EBioMedicine* 2021, 70, 103522.
34. Arij, Y.; Fukuda, M.; Nozawa, M.; Kuwada, C.; Goto, M.; Ishibashi, K.; Nakayama, A.; Sugita, Y.; Nagao, T.; Arij, E. Automatic Detection of Cervical Lymph Nodes in Patients with Oral Squamous Cell Carcinoma Using a Deep Learning Technique: A Preliminary Study. *Oral. Radiol.* 2021, 37, 290–296.

35. Jin, D.; Ni, X.; Zhang, X.; Yin, H.; Zhang, H.; Xu, L.; Wang, R.; Fan, G. Multiphase Dual-Energy Spectral CT-Based Deep Learning Method for the Noninvasive Prediction of Head and Neck Lymph Nodes Metastasis in Patients With Papillary Thyroid Cancer. *Front. Oncol.* 2022, 12, 869895.
36. Zhong, Y.; Yang, Y.; Fang, Y.; Wang, J.; Hu, W. A Preliminary Experience of Implementing Deep-Learning Based Auto-Segmentation in Head and Neck Cancer: A Study on Real-World Clinical Cases. *Front. Oncol.* 2021, 11, 638197.
37. Thor, M.; Iyer, A.; Jiang, J.; Apte, A.; Veeraraghavan, H.; Allgood, N.B.; Kouri, J.A.; Zhou, Y.; LoCastro, E.; Elguindi, S.; et al. Deep Learning Auto-Segmentation and Automated Treatment Planning for Trismus Risk Reduction in Head and Neck Cancer Radiotherapy. *Phys. Imaging Radiat. Oncol.* 2021, 19, 96–101.
38. Kawahara, D.; Tsuneda, M.; Ozawa, S.; Okamoto, H.; Nakamura, M.; Nishio, T.; Saito, A.; Nagata, Y. Stepwise Deep Neural Network (Stepwise-Net) for Head and Neck Auto-Segmentation on CT Images. *Comput. Biol. Med.* 2022, 143, 105295.
39. Oktay, O.; Nanavati, J.; Schwaighofer, A.; Carter, D.; Bristow, M.; Tanno, R.; Jena, R.; Barnett, G.; Noble, D.; Rimmer, Y.; et al. Evaluation of Deep Learning to Augment Image-Guided Radiotherapy for Head and Neck and Prostate Cancers. *JAMA Netw. Open* 2020, 3, e2027426.
40. Cubero, L.; Castelli, J.; Simon, A.; de Crevoisier, R.; Acosta, O.; Pascau, J. Deep Learning-Based Segmentation of Head and Neck Organs-at-Risk with Clinical Partially Labeled Data. *Entropy* 2022, 24, 1661.
41. De Biase, A.; Sijtsema, N.M.; van Dijk, L.V.; Langendijk, J.A.; van Ooijen, P.M.A. Deep Learning Aided Oropharyngeal Cancer Segmentation with Adaptive Thresholding for Predicted Tumor Probability in FDG PET and CT Images. *Phys. Med. Biol.* 2023, 68, 055013.
42. van Rooij, W.; Dahele, M.; Nijhuis, H.; Slotman, B.J.; Verbakel, W.F. Strategies to Improve Deep Learning-Based Salivary Gland Segmentation. *Radiat. Oncol.* 2020, 15, 272.
43. 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, 180–188.
44. Cardenas, C.E.; Beadle, B.M.; Garden, A.S.; Skinner, H.D.; Yang, J.; Rhee, D.J.; McCarroll, R.E.; Netherton, T.J.; Gay, S.S.; Zhang, L.; et al. Generating High-Quality Lymph Node Clinical Target Volumes for Head and Neck Cancer Radiation Therapy Using a Fully Automated Deep Learning-Based Approach. *Int. J. Radiat. Oncol. Biol. Phys.* 2021, 109, 801–812.
45. Wang, Y.; Lombardo, E.; Avanzo, M.; Zschaek, S.; Weingärtner, J.; Holzgreve, A.; Albert, N.L.; Marschner, S.; Fanetti, G.; Franchin, G.; et al. Deep Learning Based Time-to-Event Analysis with PET, CT and Joint PET/CT for Head and Neck Cancer Prognosis. *Comput. Methods Programs Biomed.* 2022, 222, 106948.
46. Moe, Y.M.; Groendahl, A.R.; Tomic, O.; Dale, E.; Malinen, E.; Futsaether, C.M. Deep Learning-Based Auto-Delineation of Gross Tumour Volumes and Involved Nodes in PET/CT Images of Head and Neck Cancer Patients. *Eur. J. Nucl. Med. Mol. Imaging* 2021, 48, 2782–2792.
47. Paderno, A.; Holsinger, F.C.; Piazza, C. Videomics: Bringing Deep Learning to Diagnostic Endoscopy. *Curr. Opin. Otolaryngol. Head Neck Surg.* 2021, 29, 143–148.
48. Cho, W.K.; Choi, S.-H. Comparison of Convolutional Neural Network Models for Determination of Vocal Fold Normality in Laryngoscopic Images. *J. Voice* 2022, 36, 590–598.
49. Sampieri, C.; Baldini, C.; Azam, M.A.; Moccia, S.; Mattos, L.S.; Vilaseca, I.; Peretti, G.; Ioppi, A. Artificial Intelligence for Upper Aerodigestive Tract Endoscopy and Laryngoscopy: A Guide for Physicians and State-of-the-Art Review. *Otolaryngol. Head Neck Surg.* 2023, 169, 811–829.
50. Yao, P.; Witte, D.; Gimonet, H.; German, A.; Andreadis, K.; Cheng, M.; Sulica, L.; Elemento, O.; Barnes, J.; Rameau, A. Automatic Classification of Informative Laryngoscopic Images Using Deep Learning. *Laryngoscope Investig. Otolaryngol.* 2022, 7, 460–466.
51. Patrini, I.; Ruperti, M.; Moccia, S.; Mattos, L.S.; Frontoni, E.; De Momi, E. Transfer Learning for Informative-Frame Selection in Laryngoscopic Videos through Learned Features. *Med. Biol. Eng. Comput.* 2020, 58, 1225–1238.
52. Dunham, M.E.; Kong, K.A.; McWhorter, A.J.; Adkins, L.K. Optical Biopsy: Automated Classification of Airway Endoscopic Findings Using a Convolutional Neural Network. *Laryngoscope* 2022, 132 (Suppl. 4), S1–S8.
53. Xiong, H.; Lin, P.; Yu, J.-G.; Ye, J.; Xiao, L.; Tao, Y.; Jiang, Z.; Lin, W.; Liu, M.; Xu, J.; et al. Computer-Aided Diagnosis of Laryngeal Cancer via Deep Learning Based on Laryngoscopic Images. *EBioMedicine* 2019, 48, 92–99.
54. Zhao, Q.; He, Y.; Wu, Y.; Huang, D.; Wang, Y.; Sun, C.; Ju, J.; Wang, J.; Mahr, J.J.-L. Vocal Cord Lesions Classification Based on Deep Convolutional Neural Network and Transfer Learning. *Med. Phys.* 2022, 49, 432–442.

55. Ren, J.; Jing, X.; Wang, J.; Ren, X.; Xu, Y.; Yang, Q.; Ma, L.; Sun, Y.; Xu, W.; Yang, N.; et al. Automatic Recognition of Laryngoscopic Images Using a Deep-Learning Technique. *Laryngoscope* 2020, 130, E686–E693.
56. Cho, W.K.; Lee, Y.J.; Joo, H.A.; Jeong, I.S.; Choi, Y.; Nam, S.Y.; Kim, S.Y.; Choi, S.-H. Diagnostic Accuracies of Laryngeal Diseases Using a Convolutional Neural Network-Based Image Classification System. *Laryngoscope* 2021, 131, 2558–2566.
57. Inaba, A.; Hori, K.; Yoda, Y.; Ikematsu, H.; Takano, H.; Matsuzaki, H.; Watanabe, Y.; Takeshita, N.; Tomioka, T.; Ishii, G.; et al. Artificial Intelligence System for Detecting Superficial Laryngopharyngeal Cancer with High Efficiency of Deep Learning. *Head. Neck* 2020, 42, 2581–2592.
58. Tamashiro, A.; Yoshio, T.; Ishiyama, A.; Tsuchida, T.; Hijikata, K.; Yoshimizu, S.; Horiuchi, Y.; Hirasawa, T.; Seto, A.; Sasaki, T.; et al. Artificial Intelligence-Based Detection of Pharyngeal Cancer Using Convolutional Neural Networks. *Dig. Endosc.* 2020, 32, 1057–1065.
59. Heo, J.; Lim, J.H.; Lee, H.R.; Jang, J.Y.; Shin, Y.S.; Kim, D.; Lim, J.Y.; Park, Y.M.; Koh, Y.W.; Ahn, S.-H.; et al. Deep Learning Model for Tongue Cancer Diagnosis Using Endoscopic Images. *Sci. Rep.* 2022, 12, 6281.
60. Azam, M.A.; Sampieri, C.; Ioppi, A.; Africano, S.; Vallin, A.; Mocellin, D.; Fragale, M.; Guastini, L.; Moccia, S.; Piazza, C.; et al. Deep Learning Applied to White Light and Narrow Band Imaging Videolaryngoscopy: Toward Real-Time Laryngeal Cancer Detection. *Laryngoscope* 2022, 132, 1798–1806.
61. Kim, G.H.; Sung, E.-S.; Nam, K.W. Automated Laryngeal Mass Detection Algorithm for Home-Based Self-Screening Test Based on Convolutional Neural Network. *Biomed. Eng. Online* 2021, 20, 51.
62. Wang, B.; Zheng, J.; Yu, J.-F.; Lin, S.-Y.; Yan, S.-Y.; Zhang, L.-Y.; Wang, S.-S.; Cai, S.-J.; Abdelhamid Ahmed, A.H.; Lin, L.-Q.; et al. Development of Artificial Intelligence for Parathyroid Recognition During Endoscopic Thyroid Surgery. *Laryngoscope* 2022, 132, 2516–2523.
63. Avci, S.N.; Isiktas, G.; Ergun, O.; Berber, E. A Visual Deep Learning Model to Predict Abnormal versus Normal Parathyroid Glands Using Intraoperative Autofluorescence Signals. *J. Surg. Oncol.* 2022, 126, 263–267.
64. Li, C.; Jing, B.; Ke, L.; Li, B.; Xia, W.; He, C.; Qian, C.; Zhao, C.; Mai, H.; Chen, M.; et al. Development and Validation of an Endoscopic Images-Based Deep Learning Model for Detection with Nasopharyngeal Malignancies. *Cancer Commun.* 2018, 38, 59.
65. Paderno, A.; Piazza, C.; Del Bon, F.; Lancini, D.; Tanagli, S.; Deganello, A.; Peretti, G.; De Momi, E.; Patrini, I.; Ruperti, M.; et al. Deep Learning for Automatic Segmentation of Oral and Oropharyngeal Cancer Using Narrow Band Imaging: Preliminary Experience in a Clinical Perspective. *Front. Oncol.* 2021, 11, 626602.
66. Azam, M.A.; Sampieri, C.; Ioppi, A.; Benzi, P.; Giordano, G.G.; De Vecchi, M.; Campagnari, V.; Li, S.; Guastini, L.; Paderno, A.; et al. Videomics of the Upper Aero-Digestive Tract Cancer: Deep Learning Applied to White Light and Narrow Band Imaging for Automatic Segmentation of Endoscopic Images. *Front. Oncol.* 2022, 12, 900451.
67. Lin, J.; Walsted, E.S.; Backer, V.; Hull, J.H.; Elson, D.S. Quantification and Analysis of Laryngeal Closure From Endoscopic Videos. *IEEE Trans. Biomed. Eng.* 2019, 66, 1127–1136.
68. DeVore, E.K.; Adamian, N.; Jowett, N.; Wang, T.; Song, P.; Franco, R.; Naunheim, M.R. Predictive Outcomes of Deep Learning Measurement of the Anterior Glottic Angle in Bilateral Vocal Fold Immobility. *Laryngoscope* 2023, 133, 2285–2291.
69. Kruse, E.; Dollinger, M.; Schutzenberger, A.; Kist, A.M. GlottisNetV2: Temporal Glottal Midline Detection Using Deep Convolutional Neural Networks. *IEEE J. Transl. Eng. Health Med.* 2023, 11, 137–144.
70. Adamian, N.; Naunheim, M.R.; Jowett, N. An Open-Source Computer Vision Tool for Automated Vocal Fold Tracking From Videoendoscopy. *Laryngoscope* 2021, 131, E219–E225.
71. Parker, F.; Brodsky, M.B.; Akst, L.M.; Ali, H. Machine Learning in Laryngoscopy Analysis: A Proof of Concept Observational Study for the Identification of Post-Extubation Ulcerations and Granulomas. *Ann. Otol. Rhinol. Laryngol.* 2021, 130, 286–291.
72. Yousef, A.M.; Deliyski, D.D.; Zacharias, S.R.C.; de Alarcon, A.; Orlikoff, R.F.; Naghibolhosseini, M. A Deep Learning Approach for Quantifying Vocal Fold Dynamics During Connected Speech Using Laryngeal High-Speed Videoendoscopy. *J. Speech Lang. Hear. Res.* 2022, 65, 2098–2113.
73. Yousef, A.M.; Deliyski, D.D.; Zacharias, S.R.C.; de Alarcon, A.; Orlikoff, R.F.; Naghibolhosseini, M. Spatial Segmentation for Laryngeal High-Speed Videoendoscopy in Connected Speech. *J. Voice* 2023, 37, 26–36.
74. Weng, W.; Imaizumi, M.; Muro, S.; Zhu, X. Expert-Level Aspiration and Penetration Detection during Flexible Endoscopic Evaluation of Swallowing with Artificial Intelligence-Assisted Diagnosis. *Sci. Rep.* 2022, 12, 21689.
-

