HoG-Net: Hierarchical Multi-Organ Graph Network for Head and Neck Cancer Recurrence Prediction from CT Images

Joseph Bae[1 [0000−0002−9440−6655]], Saarthak Kapse[1 [0000−0002−5426−4111]], Lei Zhou[2], Kartik Mani[3], and Prateek Prasanna[1 [0000−0002−3068−3573]]

[1] Department of Biomedical Informatics, Stony Brook University, NY 11794, USA
[2] Department of Computer Science, Stony Brook University, NY 11794, USA
[3] Department of Radiation Oncology, Stony Brook University, NY 11794, USA
{joseph.bae, prateek.prasanna}@stonybrookmedicine.edu

Abstract. In many cancers including head and neck squamous cell carcinoma (HNSCC), pathologic processes are not limited to a single region of interest, but instead encompass surrounding anatomical structures and organs outside of the tumor. To model information from organs-at-risk (OARs) as well as from the primary tumor, we present a **H**ierarchical Multi-**O**rgan **G**raph **Net**work (*HoG-Net*) for medical image modeling which we leverage to predict locoregional tumor recurrence (LR) for HNSCC patients. *HoG-Net* is able to model local features from individual OARs and then constructs a holistic global representation of interactions between features from multiple OARs in a single image. *HoG-Net*'s prediction of LR for HNSCC patients is evaluated in a largest yet studied dataset of N=2,741 patients from six institutions, and outperforms several previously published baselines. Further, *HoG-Net* allows insights into which OARs are significant in predicting LR, providing specific OAR-level interpretability rather than the coarse patch-level interpretability provided by other methods. Code can be found at https://github.com/bmi-imaginelab/HoGNet.

Keywords: Head and Neck Cancer · Graph Neural Networks · Tumor Recurrence

1 Introduction

The human body is a complex system of organs and tissues dependent on one another for normal physiological function. Disease processes including cancer are not limited to well-defined spatial confines, but instead pose wide-ranging disruptions to this interconnected system. Medical images provide a snapshot view into the body, but oftentimes medical image analysis fails to model disease processes within this larger context and instead focus on a single region of interest (ROI). Modeling the full range of surrounding anatomy is particularly relevant for certain cancers including head and neck squamous cell carcinoma (HNSCC)

in which disease migration prominently occurs in structures and organs neighboring the primary tumor. While some methods have attempted to combine imaging features from multiple ROIs, they often suffer from several limitations [12,3,10]. First, many methods require uniformity in image input dimensions, and are inflexible in modeling the highly variable dimensionality of inter-patient organ and anatomical structure sizes. Similarly, there is often variability in patient anatomy or data acquisition that may result in incomplete ROI delineation across patient samples, resulting in samples with different numbers of ROIs for models to analyze. Finally, features from multiple regions are often fused via simple concatenation or averaging which can quickly lead to overwhelming feature dimensionality [17,4], especially in the setting of smaller medical imaging datasets. To overcome these challenges and to holistically model local and global imaging features from tumor and non-tumor regions, we propose a hierarchical graph-based image analysis pipeline which can flexibly model heterogenous patient data while providing interpretable insights into which regions of the head and neck anatomy are significant in predicting tumor recurrence.

HNSCC is among the most common cancers and constitutes 4.5% of cancer diagnoses and 4.6% of cancer deaths each year [5]. It is highly invasive, often presenting with locoregional involvement of nearby structures including cervical lymph node chains and other organs in the head and neck anatomy [9]. Furthermore, though improvements in therapy including chemoradiation have improved outcomes, up to 30% of patients develop locoregional recurrence (LR) of their disease, defined as tumor recurrence within the same primary site or surrounding head and neck anatomy [22]. 3D CT images are regularly acquired for HNSCC patients as part of the normal radiation treatment planning process, and used to contour targeted tumor regions as well as organs-at-risk (OARs) which might be damaged by radiation due to their proximity to tumors. Multiple approaches have been proposed for the prediction of LR after radiation from pre-treatment CT imaging including those leveraging radiomic features [23] and CNNs [6,18]. These studies have exclusively modeled image features from the tumor and have proven suboptimal in demonstrating satisfactory generalizable performance for LR classification [23,18]. The possibility for local changes in specific OARs as a result of HNSCC pathological development and the global interactions between primary tumors and their anatomical environments motivate our proposed hierarchical graph-based approach.

Computational graphs are flexible data structures which encode multiple data points (nodes) and their relationship (edges) to one another. Graph neural networks (GNNs) have been extensively studied to model graphs in a variety of domains with common tasks including node classification [13] and graph classification [24,7,16]. Graph representations of images have been previously modeled using GNNs, including in the domains of natural [2] and histopathology images [26]. Graph analysis of medical images has also been explored, commonly studying the lung and brain [7,16]. While in some of these domains graph representations directly correspond to well-defined ROIs such as cell nuclei or brain parcellations, in many cases constructing graphs from images is less straightfor-

ward. Oftentimes, studies face difficulties in identifying meaningful nodes within input images or in building graph representations of different imaging features. For example, image patches or supervoxels are frequently used to create nodes within a graph structure [20,15,2,3]. However, these patches often do not conform to biologically or anatomically defined ROIs and instead are simply arbitrary partitions of imaging data. Node features for these patches are frequently assigned by using simplistic aggregations of pixel- or voxel-wise intensity (i.e. mean or maximum for a patch) or by inputting the patch into an encoder (i.e. ResNet, thereby requiring a uniform patch size for each node). Other recent approaches have also leveraged radiomic features extracted from supervoxels or other ROIs [4,10]. These methods of node identification and node feature extraction incompletely leverage the flexibility of graph structures, and often lose information in the transition from image to graph representations. In turn, the interpretability of such graph-based approaches is lowered as nodes do not correspond to meaningful ROIs and performance can be worsened by the loss of pixel-wise information.

Contributions. We present *HoG-Net*, a hierarchical multi-organ graph network which seamlessly encodes imaging information from physician-defined organs-at-risk (OARs) and tumors into a graph structure for downstream analysis (henceforth we refer to OARs as the set of all organs-at-risk and the primary tumor for simplicity of notation). *HoG-Net* leverages two key innovations. First, we develop an OAR_{enc} module which effectively models *intra-OAR* heterogeneity by encoding OAR imaging information into nodes and node features, irrespective of differences in input dimensionality. Second, we propose a *SuperGraph* module predicated on modeling HNSCC in the global context of OAR interconnectivity (i.e. *inter-OAR* heterogeneity) to predict LR from pre-treatment CTs. *HoG-Net* uniquely addresses the difficulty of modeling non-uniform inter-patient anatomy (OAR_{enc}) while building a holistic representation (*SuperGraph*) of HNSCC pathology on imaging in order to improve classification of patient outcomes. *HoG-Net* is evaluated on the largest dataset studied to date for HNSCC locoregional recurrence prediction (N=2,741) and outperforms several previously published state-of-the-art (SOTA) methods.

2 Methodology

In this section we first describe graph convolution and graph attention mechanisms. We then describe how OAR_{enc} extracts local imaging features from multiple OARs from each patient and encodes each into a graph representation. We finally describe the creation of a patient-level *SuperGraph*, encoding global relationships between different OARs of a single patient. An overview of the proposed framework is shown in Figure 1 and Algorithm 1.

2.1 Preliminary

Graph neural networks have been extensively studied in multiple domains to model graph data structures for various downstream tasks. Given nodes $\mathcal{V} =$

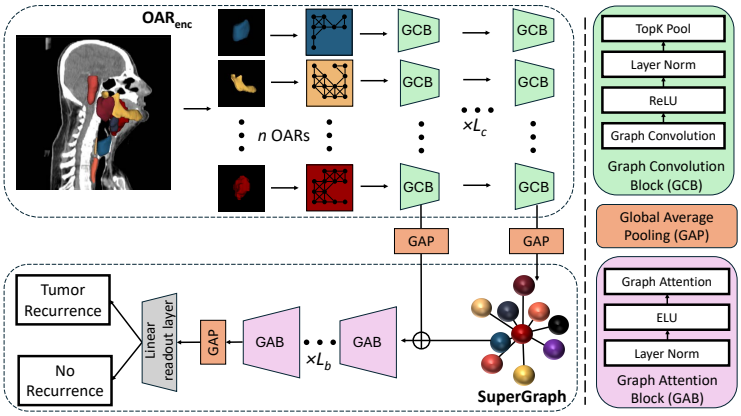

Fig. 1. Framework Overview. An input image with n corresponding organ-at-risk (OAR) segmentations is used to create individual graphs for each OAR. Following, n encoders (each consisting of $\times L_c$ graph convolution blocks) encodes the corresponding n OARs into graph representations. These representations are used as node features for *SuperGraph* yielding an image-level graph for each patient. *SuperGraph* is modeled via $\times L_b$ graph attention blocks for tumor recurrence prediction.

$\{v_0, v_1, ...v_m\}$, associated node feature vectors $\{x_0, x_1, ...x_m\}$, and edges $\mathcal{E} = \{(v_i, v_j)\}$, a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ can be defined. The adjacency matrix is then defined as $\mathcal{A} = 1$ if $(v_i, v_j) \in \mathcal{E}$, else: 0. The diagonal degree matrix is $\mathcal{D}_{ii} = \sum_j \mathcal{A}_{ij}$. $\mathcal{A}$ contains edges between v_i and v_j whereas $\mathcal{D}$ provides the number of edges containing v_i. Both graph convolutional neural networks (GCNs) and graph attention neural networks (GATs) have been extensively studied for graph modeling. Mirroring CNNs commonly used in imaging applications, GCNs employ convolutions to update node features based upon neighboring node features. Kipf and Welling [13] formulated graph convolutions as $X' = \tilde{\mathcal{D}}^{-1/2}\tilde{\mathcal{A}}\tilde{\mathcal{D}}^{-1/2}X\theta$, where $\tilde{\mathcal{A}} = \mathcal{A} + \mathcal{I}$, $\tilde{\mathcal{D}}_{ii} = \sum_j \tilde{\mathcal{A}}_{ij}$, and θ is a learnable weight matrix. Alternatively, GATs leverage widely popularized attention mechanisms for graph data structures. Specifically, Veličković et al. [24] formulated graph attention as a sequence of MLP-based projection of node features followed by computation of attention scores between projected node features using a simple feed-forward layer. Succinctly, $x_i' = \|_{k=1}^{K}\sigma(\sum_j \alpha_{ij}\theta x_j)$ where x_i' denotes an updated node feature, σ denotes an activation function, α is the learned attention score between two nodes, θ is a learnable weight matrix, and K represents multiple attention heads. α is only computed for nodes with edges in $\mathcal{E}$.

Overview of Hierarchical Graph Architecture. To model imaging features from relevant organs-at-risk (OARs) and the primary tumor, we adopt a hierarchical graph approach. First, we leverage an OAR_{enc} module to create local graphs for each of the n pathology-relevant OARs. Graph convolutions are then applied to each OAR graph to produce features which are embedded into a

global *SuperGraph* representation. Finally, GAT layers model *SuperGraph* to produce the binary classification prediction.

Algorithm 1: Overview of $HoG - Net$

Input : CT Image and OAR Contours per Patient
Output: Binary Prediction of Locoregional Recurrence

***OAR*$_{enc}$**: **for** *OAR* **do**
  1. **Create** a graph $\mathcal{G}_o$ from voxel neighbors of OAR;
  2. **Perform** graph convolution;
  3. **Pool** node features to produce OAR feature vector f_o;
end
SuperGraph: **for** *OAR* **do**
  1. **Encode** OAR feature vector f_o into SuperGraph $\mathcal{G}$ node;
  2. **Create** edge between OAR and primary tumor node;
end
Model SuperGraph $\mathcal{G}$ with graph attention layers.

2.2 Organ at Risk Graph Encoding

***OAR*$_{enc}$**.Using segmentations for OARs on each patient CT, OAR_{enc} first builds a graph $\mathcal{G}_o = \{\mathcal{V}, \mathcal{E}\}$ for each OAR studied. $\mathcal{V}$ contains the set of all voxels v_i within the mask of OAR o with corresponding node features $x_i = [I_i, c_x, c_y, c_z]$ representing patient-normalized voxel intensity and x, y, and, z coordinates of v_i. The edge set $\mathcal{E}$ is constructed from the 26 voxels v_j nearest v_i, provided they reside within the contour for OAR o. This process results in the creation of n graphs for each patient image with variable node numbers reflecting OAR sizes. Each of these graphs is modeled independently via separate sets of multiple Graph Convolution Blocks (GCBs) composed of graph convolution and TopK pooling layers. This produces a feature vector f_o representing imaging features for a given OAR. The use of GCN and TopK pooling mirrors down-sampling operations of conventional CNNs and concentrates high- and low-level information into feature vectors for *SuperGraph*. Skip connections are created by global average pooling output node features from each GCB and appending them to input node features for *SuperGraph* to retain information from multiple feature levels [21].

2.3 SuperGraph Modeling

SuperGraph Creation. A *SuperGraph* representation is formulated from the combined feature vectors of each OAR produced by OAR_{enc}. Formally, $\mathcal{G}_s = \{\mathcal{V}_s, \mathcal{E}_s\}$ where $\mathcal{V}_s$ contains n OAR nodes (v_o), each representing a single OAR from a given patient CT. Node features f_o are obtained from OAR_{enc} concatenated with residual features from intermediate GCN outputs and the centroid

coordinates for each OAR. Edges for *SuperGraph* are $\mathcal{E}_s = [(v_0, v_o)]_{o=1}^{n}$ where v_0 corresponds to the primary tumor node. This reflects a focus on the pathology of the primary HNSCC tumor and enables *HoG-Net* to model relationships between each OAR and the primary tumor. In this way, one *SuperGraph* is created per patient containing distilled image information from each OAR subgraph.

To model *SuperGraph* representations, we employ multiple Graph Attention Blocks (GABs), each composed of a graph attention layer, layer norm, and ELU activation. A global average pooling and linear readout layer are used for binary classification. The entire *HoG-Net* model including OAR_{enc} GCBs and *SuperGraph* GABs is trained end-to-end using binary cross entropy loss.

3 Experiments

3.1 Datasets and Implementation

The binary classification task studied was presence or absence of locoregional tumor recurrence two years following radiotherapy. This outcome is clinically relevant, and is shared with previous studies for ease of comparison [23,6,18]. The publicly available RADCURE [27] (N=2,346, 385 LR), Head-Neck-PET-CT [23] (HNPET, N=274, 30 LR), and Head-Neck-Radiomics-1 [1] (HN1, N=121, 30 LR) datasets from the Cancer Imaging Archive are studied with a total N=2,741 patient CTs. Clinical variables studied included whether a patient received concurrent chemo-radiotherapy in addition to patient HPV status, sex, age, tumor stage (AJCC 7th edition), Eastern Cooperative Oncology Group (ECOG) performance status, tumor subsite (oropharyngeal, laryngeal, etc.), and tumor volume (assessed using dimensions of primary tumor contours). These variables were concatenated with output features from each studied model prior to the linear layer used for binary classification. To evaluate model generalizability across datasets and institutions, we trained (N=1,502) and validated (N=342) *HoG-Net* on RADCURE and tested held out data from RADCURE, HNPET, and HN1. Training and testing splits for RADCURE mirrored those of the published challenge [11], with the remaining cases used for validation. In total there were 1,502 training, 342 validation, and 897 testing samples. In addition to CT images, each studied patient had a primary tumor segmentation created by the treating physician. Because ground-truth segmentations for OARs were available for only a subset of publicly provided data, nnUNet [8] models were trained and used for OAR segmentation (details in supplementary).

All model training was performed on a single NVIDIA GeForce RTX 2080 Ti with a batch size of 4. *HoG-Net* was implemented with $L_c = 5$ GCBs in OAR_{enc} and $L_b = 2$ GABs for *SuperGraph* modeling.

3.2 Quantitative Prediction of LR

We report AUC and F1-score for LR classification on RADCURE, HNPET, and HN1 test datasets. *HoG-Net* is compared to Mateus et al. [18] and Diamant

Table 1. Metrics for Locoregional Recurrence Prediction.

	HN1		**HNPET**		**RADCURE**	
Image Only ↓	AUC	F1	AUC	F1	AUC	F1
Ours	**0.662**	**0.577**	0.649	0.501	**0.721**	**0.604**
Diamant et al.	0.595	0.465	0.488	0.488	0.556	0.526
Mateus et al.	0.604	0.429	0.534	0.471	0.62	0.461
Vallieres et al.	0.645	0.571	0.507	0.479	0.582	0.563
DenseNet (OAR+Tumor)	0.653	0.537	**0.666**	**0.528**	0.673	0.557
DenseNet (Full image)	0.648	0.345	0.664	0.329	0.696	0.495
Image+Clinical ↓						
Ours	**0.809**	**0.651**	**0.759**	**0.582**	**0.807**	**0.658**
Diamant et al.	0.628	0.584	0.534	0.506	0.627	0.572
Mateus et al.	0.793	0.429	0.751	0.471	0.804	0.461
Vallieres et al.	0.66	0.572	0.633	0.385	0.732	0.575
DenseNet (OAR+Tumor)	0.657	0.476	0.689	0.518	0.702	0.594
DenseNet (Full image)	0.665	0.487	0.663	0.407	0.699	0.595

et al. [6] which published CNN models for LR prediction for HNSCC patients using the HNPET dataset. Additionally we add a radiomic baseline mirroring Vallieres et al. [23], one of the original studies for this task. Further, we add two 3D DenseNet baselines (entire image and entire image with non-OAR regions masked as inputs). Finally, we compare the performance of each of these models when trained with only the image as input as well as with clinical variables concatenated.

As seen in the quantitative results in Table 1, *HoG-Net* outperforms baselines in most test sets. This is the largest study to date of imaging-based prediction of locoregional recurrence for HNSCC, and thus our results comprise a baseline for future experiments on these public datasets. As reported by previous studies [23,6,18], the addition of clinical features to imaging features improves performance when compared to imaging features alone. Interestingly, tuned DenseNet models outperformed previous approaches [23,6,18] on several of the test sets studied. This might reflect improved discriminative performance by providing a larger degree of contextual information to models in the form of OAR regions rather than using only the primary tumor as commonly performed in the existing literature [6,18]. *HoG-Net* expands further upon this larger context by explicitly modeling the relationship between OARs and the primary tumor as a *SuperGraph*, and thus achieves both superior performance and improved interpretability when compared to all baselines (Section 3.3).

3.3 OAR Attention

Model attention values were calculated from the final graph attention layer of *HoG-Net* and used to identify which OARs composing *SuperGraph* receive highest attention during model inference (Figure 2). Across all cases, model attention is consistently highest to the esophagus followed by the parotid glands,

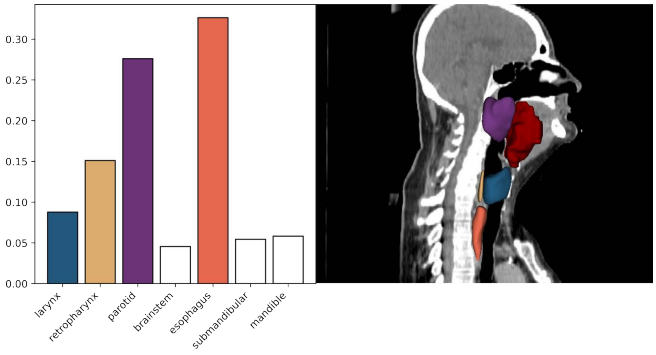

Fig. 2. Model Attention to OARs. Shown (left) are relative graph attention values to each of the OARs studied. Also visualized (right) are the locations of the most highly attended OARs as well as one example of a typical HNSCC primary tumor (red).

retropharynx, and larynx. There are several potential clinical explanations for these findings. First, secondary primary tumors frequently present as locoregional recurrence in HNSCC patients, and these lesions frequently appear in the contiguous structures of the retropharynx, larynx, and esophagus. It is therefore possible that our model may be detecting subtle changes in these OARs suggestive of secondary primary tumor development or locoregional recurrence. Second, each of these structures remains in the vicinity of primary HNSCC tumors; it is encouraging for instance that our model does not direct high attention to the brainstem which is generally not involved in HNSCC pathophysiology. These findings remain a potential avenue for future study, and also demonstrate the unique interpretability of our graph-based approach; other interpretability approaches such as class activation maps provide coarser alternatives to model attention over large regions of an image which cannot be directly assigned to biologically-relevant ROIs. *HoG-Net* is able to definitively identify which OARs are highly attended during model inference on both patient and dataset-population levels since each node of *SuperGraph* exclusively corresponds to a single OAR.

4 Conclusion

In this work we propose a hierarchical graph based approach to holistically model biologically meaningful OARs for LR prediction of HNSCC tumors. We outperform several SOTA methods on the classification task while also providing interpretability into which OARs are significant for model predictions. Individual variations in model attention to these different OARs may be of clinical value in future studies. Recent research has been undertaken to determine what regions of the surrounding head and neck should be targeted during radiotherapy [19], and interpretable computer vision tools such as *HoG-Net* might inform these analyses and perhaps prompt investigation into personalized radiation therapy

planning. There are several limitations of our work. First, in line with previous studies [23,6,18], we perform binary classification of LR. In future studies we might adapt our model for time-to-event analysis to better capture the disease progression of HNSCC. Second, our formulation of *SuperGraph* currently does not involve edge interactions between OARs with one another; we only model edges between each OAR and the primary tumor. In future work we might explore additional OARs and the interactions present between them. Finally, we plan to extend *HoG-Net* to other medical imaging tasks [25,14] where holistically modeling local and global image features may more comprehensively capture biological processes.

Acknowledgments. Funding sources include Stony Brook University OVPR Seed Grant, NIH 5R21CA258493-02, and NIGMS T32GM008444.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Aerts, H.J., Velazquez, E.R., Leijenaar, R.T., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., et al.: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature communications **5**(1), 4006 (2014)
2. Avelar, P.H., Tavares, A.R., da Silveira, T.L., Jung, C.R., Lamb, L.C.: Superpixel image classification with graph attention networks. In: 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). pp. 203–209. IEEE (2020)
3. Bae, J., Mani, K., Noldner, C., Czerwonka, L., Ryu, S., Prasanna, P.: Do spatial-radiomics improve prediction of locoregional recurrence following radiotherapy for hnscc? International Journal of Radiation Oncology, Biology, Physics **118**(5), e69 (2024)
4. Bae, J., Mani, K., Zabrocka, E., Cattell, R., O'Grady, B., Payne, D., Roberson, J., Ryu, S., Prasanna, P.: Pre-treatment spatially-aware mri radiomics can predict distant brain metastases (dbms) following stereotactic radiosurgery/radiation therapy (srs/srt). Advances in Radiation Oncology p. 101457 (2024)
5. Barsouk, A., Aluru, J.S., Rawla, P., Saginala, K., Barsouk, A.: Epidemiology, risk factors, and prevention of head and neck squamous cell carcinoma. Medical Sciences **11**(2), 42 (2023)
6. Diamant, A., Chatterjee, A., Vallières, M., Shenouda, G., Seuntjens, J.: Deep learning in head & neck cancer outcome prediction. Scientific reports **9**(1), 2764 (2019)
7. Guo, J., Qiu, W., Li, X., Zhao, X., Guo, N., Li, Q.: Predicting alzheimer's disease by hierarchical graph convolution from positron emission tomography imaging. In: 2019 IEEE international conference on big data (big data). pp. 5359–5363. IEEE (2019)
8. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
9. Jimenez, L., Jayakar, S.K., Ow, T.J., Segall, J.E.: Mechanisms of invasion in head and neck cancer. Archives of pathology & laboratory medicine **139**(11), 1334–1348 (2015)

10. Kazmierski, M., Haibe-Kains, B.: Lymph node graph neural networks for cancer metastasis prediction. arXiv preprint arXiv:2106.01711 (2021)
11. Kazmierski, M., Welch, M., Kim, S., McIntosh, C., Rey-McIntyre, K., Huang, S.H., Patel, T., Tadic, T., Milosevic, M., Liu, F.F., et al.: Multi-institutional prognostic modelling in head and neck cancer: evaluating impact and generalizability of deep learning and radiomics. Cancer Research Communications pp. CRC–22 (2023)
12. Keek, S., Sanduleanu, S., Wesseling, F., De Roest, R., Van Den Brekel, M., Van Der Heijden, M., Vens, C., Giuseppina, C., Licitra, L., Scheckenbach, K., et al.: Computed tomography-derived radiomic signature of head and neck squamous cell carcinoma (peri) tumoral tissue for the prediction of locoregional recurrence and distant metastasis after concurrent chemo-radiotherapy. PLoS One **15**(5), e0232639 (2020)
13. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
14. Konwer, A., Bae, J., Singh, G., Gattu, R., Ali, S., Green, J., Phatak, T., Gupta, A., Chen, C., Saltz, J., et al.: Predicting covid-19 lung infiltrate progression on chest radiographs using spatio-temporal lstm based encoder-decoder network. In: Medical Imaging with Deep Learning. pp. 384–398. PMLR (2021)
15. Li, R., Zhou, L., Wang, Y., Shan, F., Chen, X., Liu, L.: A graph neural network model for the diagnosis of lung adenocarcinoma based on multimodal features and an edge-generation network. Quantitative Imaging in Medicine and Surgery **13**(8), 5333 (2023)
16. Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., Scheinost, D., Staib, L.H., Ventola, P., Duncan, J.S.: Braingnn: Interpretable brain graph neural network for fmri analysis. Medical Image Analysis **74**, 102233 (2021)
17. Lv, W., Feng, H., Du, D., Ma, J., Lu, L.: Complementary value of intra-and peri-tumoral pet/ct radiomics for outcome prediction in head and neck cancer. IEEE Access **9**, 81818–81827 (2021)
18. Mateus, P., Volmer, L., Wee, L., Aerts, H.J., Hoebers, F., Dekker, A., Bermejo, I.: Image based prognosis in head and neck cancer using convolutional neural networks: a case study in reproducibility and optimization (2023)
19. Nissi, L., Suilamo, S., Kytö, E., Vaittinen, S., Irjala, H., Minn, H.: Recurrence of head and neck squamous cell carcinoma in relation to high-risk treatment volume. Clinical and Translational Radiation Oncology **27**, 139–146 (2021)
20. Sun, L., Yu, K., Batmanghelich, K.: Context matters: Graph-based self-supervised representation learning for medical images. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 35, pp. 4874–4882 (2021)
21. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence. vol. 31 (2017)
22. Taneja, C., Allen, H., Koness, R.J., Radie-Keane, K., Wanebo, H.J.: Changing patterns of failure of head and neck cancer. Archives of Otolaryngology–Head & Neck Surgery **128**(3), 324–327 (2002)
23. Vallieres, M., Kay-Rivest, E., Perrin, L.J., Liem, X., Furstoss, C., Aerts, H.J., Khaouam, N., Nguyen-Tan, P.F., Wang, C.S., Sultanem, K., et al.: Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. Scientific reports **7**(1), 10117 (2017)
24. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2017)

25. Wang, F., Kapse, S., Liu, S., Prasanna, P., Chen, C.: Topotxr: a topological biomarker for predicting treatment response in breast cancer. In: International Conference on Information Processing in Medical Imaging. pp. 386–397. Springer (2021)
26. Wang, Y., Wang, Y.G., Hu, C., Li, M., Fan, Y., Otter, N., Sam, I., Gou, H., Hu, Y., Kwok, T., et al.: Cell graph neural networks enable the precise prediction of patient survival in gastric cancer. NPJ precision oncology **6**(1), 45 (2022)
27. Welch, M.L., Kim, S., Hope, A.J., Huang, S.H., Lu, Z., Marsilla, J., Kazmierski, M., Rey-McIntyre, K., Patel, T., O'Sullivan, B., et al.: Radcure: An open-source head and neck cancer ct dataset for clinical radiation therapy insights. Medical Physics (2024)