The American Journal of Pathology, Vol. 194, No. 7, July 2024

The American Journal of
PATHOLOGY
ajp.amjpathol.org

MACHINE LEARNING, COMPUTATIONAL PATHOLOGY, AND BIOPHYSICAL IMAGING

Graph Perceiver Network for Lung Tumor and Bronchial Premalignant Lesion Stratification from Histopathology

Rushin H. Gindra,[*†‡] Yi Zheng,[*§] Emily J. Green,[*] Mary E. Reid,[¶] Sarah A. Mazzilli,[*] Daniel T. Merrick,[‖] Eric J. Burks,[**] Vijaya B. Kolachalama,[*§††] and Jennifer E. Beane[*]

From the Departments of Medicine and Pathology and Laboratory Medicine,** Boston University Chobanian and Avedisian School of Medicine, Boston, Massachusetts; the Institute of AI for Health,[†] Helmholtz Zentrum Munich—German Research Center for Environmental Health, Munich, Germany; the Center for Translational Cancer Research,[‡] School of Medicine, Technische Universität München, Munich, Germany; the Department of Computer Science[§] and the Faculty of Computing and Data Sciences,[††] Boston University, Boston, Massachusetts; the Roswell Park Comprehensive Cancer Center,[¶] Buffalo, New York; and the Department of Pathology,[‖] University of Colorado School of Medicine, Aurora, Colorado*

Accepted for publication March 11, 2024.

Address correspondence to Vijaya B. Kolachalama, Ph.D., or Jennifer E. Beane, Ph.D., Boston University Chobanian and Avedisian School of Medicine, 72 E. Concord St., Boston, MA 02118.
E-mail: vkola@bu.edu or jbeane@bu.edu.

Bronchial premalignant lesions (PMLs) precede the development of invasive lung squamous cell carcinoma (LUSC), posing a significant challenge in distinguishing those likely to advance to LUSC from those that might regress without intervention. This study followed a novel computational approach, the Graph Perceiver Network, leveraging hematoxylin and eosin—stained whole slide images to stratify endobronchial biopsies of PMLs across a spectrum from normal to tumor lung tissues. The Graph Perceiver Network outperformed existing frameworks in classification accuracy predicting LUSC, lung adenocarcinoma, and nontumor lung tissue on The Cancer Genome Atlas and Clinical Proteomic Tumor Analysis Consortium datasets containing lung resection tissues while efficiently generating pathologist-aligned, class-specific heatmaps. The network was further tested using endobronchial biopsies from two data cohorts, containing normal to carcinoma *in situ* histology. It demonstrated a unique capability to differentiate carcinoma *in situ* lung squamous PMLs based on their progression status to invasive carcinoma. The network may have utility in stratifying PMLs for chemoprevention trials or more aggressive follow-up. *(Am J Pathol 2024, 194: 1285—1293; https://doi.org/10.1016/j.ajpath.2024.03.009)*

Lung squamous cell carcinoma (LUSC), the second most common type of non—small-cell lung cancer, is preceded by the development of bronchial premalignant lesions (PMLs) that can progress toward invasive carcinoma or regress without intervention. Recently, genomic and proteomic profiling of PMLs has demonstrated that PML progression is associated with impaired immunosurveillance.[1—5] Currently, there are ongoing efforts to build a Lung Pre-Cancer Atlas[6] to understand molecular alterations in their spatial context associated with disease severity and progression. To date, hematoxylin and eosin—stained tissue slides of PMLs have not been fully used as a data source to augment our insights into PML biology. Endobronchial biopsies of PMLs are heterogeneous and contain a range of histologic grades, a variety of structures, including cartilage and submucosal glands, and varying degrees of inflammatory cell infiltration. We hypothesized that computational methods that use digitized hematoxylin and eosin—stained whole slide images (WSIs) may be able to capture PML heterogeneity and stratify PMLs by histologic severity or their ability to progress to invasive carcinoma, providing a

Supported by NIH grants R21-CA253498 (J.E.B., V.B.K., R.H.G., and Y.Z.), U2C-CA233238 (J.E.B., S.A.M., D.T.M., and M.E.R.), R01-HL159620 (V.B.K.), R43-DK134273 (V.B.K.), RF1-AG062109 (V.B.K.), P30-AG073104 (V.B.K.), and 1UL1TR001430 (V.B.K.); Johnson & Johnson Enterprise Innovation, Inc. (R.H.G., Y.Z., E.J.G., D.T.M., E.J.B., V.B.K., J.E.B., S.A.M., and M.E.R.); the American Heart Association 20SFRN35460031 (V.B.K.); and the Karen Toffler Charitable Trust (V.B.K.).

V.B.K. and J.E.B. contributed equally to this work.

https://doi.org/10.1016/j.ajpath.2024.03.009

standardized and informative WSI-level assessment of PMLs.

Towards this end, a Graph Perceiver Network (GRAPE-Net) was developed to characterize PMLs using lung tissue WSIs from public cohorts [The Cancer Genome Atlas (TCGA) and Clinical Proteomic Tumor Analysis Consortium (CPTAC)], consisting of normal (nontumor adjacent), lung adenocarcinoma (LUAD), and LUSC tissue samples. GRAPE-Net was trained using TCGA samples and tested using CPTAC samples. It was then used to learn WSI-level features and classes of endobronchial biopsy-based WSIs from two datasets from the University College London (UCL) and the Roswell Park Comprehensive Cancer Institute (Roswell). The network identified signatures that differentiated the lung tissue types, stratified samples by their predominant histologic pattern within LUSC and LUAD tumors, and identified carcinoma *in situ* (CIS) PMLs that progress to LUSC.

Materials and Methods

Ethics Statement

The Institutional Review Boards at Boston University Chobanian and Avedisian School of Medicine (Boston, MA) and the Roswell Park Comprehensive Cancer Center (Buffalo, NY) approved the study. All subjects provided written informed consent. Data from TCGA, CPTAC, and UCL are publicly available.

Data Acquisition and Preprocessing

Digitized hematoxylin and eosin–stained WSIs of lung resection tissue and lung tumor subtype annotation were obtained from TCGA (534 normal, 740 LUSC, and 808 LUAD) and CPTAC (719 normal, 685 LUSC, and 667 LUAD). Predominant tumor histologic pattern was obtained for the CPTAC samples. The endobronchial biopsy WSIs were from two datasets: UCL (112 CIS samples),[4] where progression was defined as development of LUSC; and Roswell (346 samples, normal to CIS spectrum),[1] where progression was defined as persistence of dysplasia, advancement to mild dysplasia, or a worst histologic grade (Figure 1A). The UCL CIS samples included supplementary details about the samples' lymphocyte counts in CIS and stromal regions. The Roswell samples included molecular phenotypes identified using RNA-sequencing data, such as PML molecular subtype that included four categories, proliferative, inflammatory, secretory, and normal-like, as previously described.[1]

To evaluate the model's explanations, WSIs from CPTAC, Roswell, and UCL were uploaded to a web-based software (PixelView; deepPath, Boston, MA). Several regions on the WSIs were annotated by board-certified thoracic pathologists (D.T.M. and E.J.B.) using the software. In the LUAD specimens, tumor areas were annotated by their histologic patterns (solid, micropapillary, cribriform, papillary, acinar, and lepidic). For both LUAD and LUSC tumors, histologic features, including necrosis, lymphatic invasion, and vascular invasion, were also annotated. Airways present in the tissues were graded on the spectrum of normal to CIS, and nontumor tissue areas were labeled as healthy or inflamed lung tissue, stroma, or necrosis. The LUAD and LUSC annotations were combined into tumor, stroma, normal lung, and necrosis for analysis. For the endobronchial biopsy samples, the histologic grade of the epithelial regions was annotated as normal, hyperplasia, immature squamous metaplasia, squamous metaplasia, mild dysplasia, moderate dysplasia, severe dysplasia, and CIS. These annotations were combined into CIS, dysplasia, and nondysplasia for analysis. The overlap between the model-produced heatmaps and the pathologist annotations was analyzed by exporting them as binary images.

WSIs are inherently large, often exceeding dimensions of tens of thousands of pixels in both width and height. This presents challenges for analyzing them because of computational and memory constraints. To address these limitations, each WSI was passed through a fast-patching pipeline at $\times 20$ magnification, generating nonoverlapping tissue region patches of 256×256 pixels while filtering out background slide information. Key epithelial regions were retained via Otsu thresholding.[7] To prepare the data as input for GRAPE-Net, each WSI was represented as an undirected, unweighted graph with nodes representing the tissue image patches and edges connecting the nodes. The graph followed an eight-connectivity neighborhood structure. The node embeddings in the graph were structured as a matrix of N feature vectors, where N is the total number of patches in the WSI (nodes in the graph). Each vector was of size 768 dimensions. The features for each patch were obtained from CTransPath,[8] a swin-transformer[9] based feature extractor. CTransPath was pretrained on TCGA pan-cancer dataset in a self-supervised manner. The graph neighborhood connections were structured as a binary $N \times N$ matrix (adjacency matrix A) with elements $e_{ij} = 1$ if there existed an edge between nodes i and j. These graphs fed into GRAPE-Net, which was then trained for a WSI classification task (Figure 1B).

Graph Perceiver Network

The architectural design of GRAPE-Net consists of three main components (Figure 2): i) a graph convolution block[10] that retains position information in the input features via edge connections between neighboring nodes, ii) a cross-attention pooling block from perceiver[11] that maps variable number of patches per WSI graph to predefined clusters (or sets), and iii) a self-attention block[12] that learns relevant interactions between the sets.

The goal of GRAPE-Net is to learn a representation of the WSI. The graph convolution block contributes to this by stacking multiple graph convolutional layers to learn

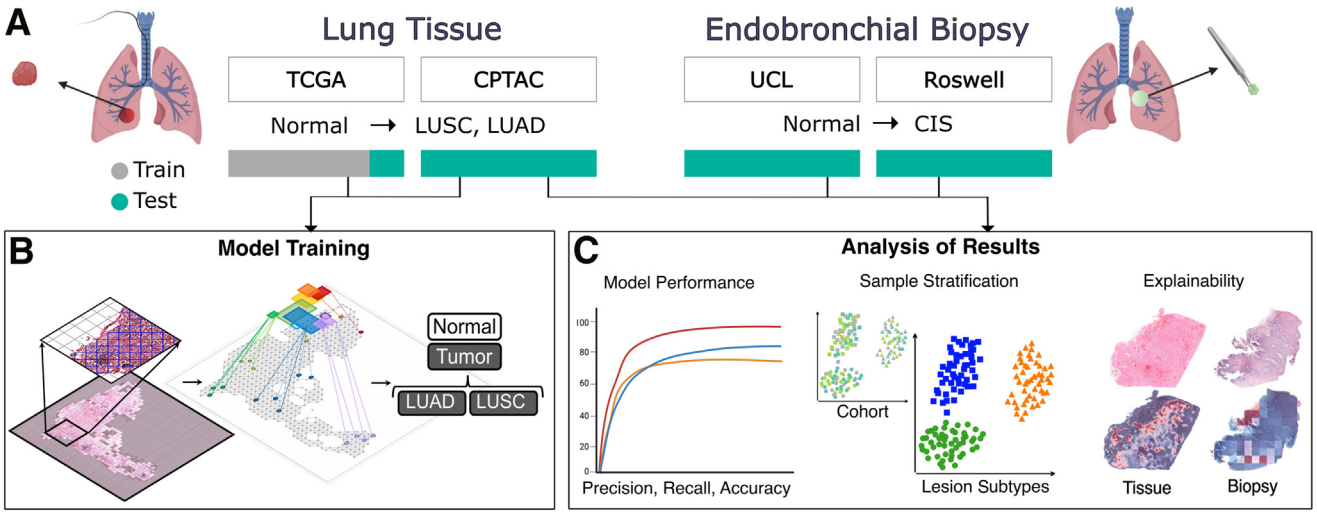

Figure 1 Overview of the study. **A:** Digitized hematoxylin and eosin—stained whole slide images were obtained from four different datasets, including lung resection tissues [The Cancer Genome Atlas (TCGA) and Clinical Proteomic Tumor Analysis Consortium (CPTAC) datasets] with nontumor tissue adjacent to the tumor (Normal), lung adenocarcinoma (LUAD), and lung squamous cell carcinoma (LUSC) and endobronchial biopsies [University College London (UCL) and Roswell Park Comprehensive Cancer Institute (Roswell) datasets] ranging from normal to carcinoma *in situ* (CIS) histology. The model was trained on TCGA data and tested across TCGA, CPTAC, UCL, and Roswell datasets. **B:** Graph Perceiver Network overview trained to differentiate lung cancer subtypes. **C:** Overview of analysis conducted using the results of the model that included evaluation of model performance, calculation of clustering-based metrics using whole slide image features, and model explainability using class-specific heatmaps. Panel **A** was generated with BioRender.com (Toronto, ON, Canada).

hierarchical features efficiently. Specifically, the first few layers capture the local, low-level patterns, whereas the deeper layers aggregate them to understand the global structures within the graph. Next, inspired by the recent work on using min-cut pooling on a graph-transformer network for WSI-level classification,[13] and adaptive aggregation functions used in recent graph classification problems,[14] GRAPE-Net uses a cross-attention pooling block from the perceiver. This module simplifies graph-transformer network by replacing the min-cut pooling[15] mechanism with cross-attention layers, which cluster the graph nodes with similar embeddings and spatial proximity to predefined latent clusters, denoted as sets. The number of sets used in the cross-attention pooling module is a hyperparameter C, typically much smaller than the expected number of visual tokens. The formed sets, accompanied by a classification token, are inputs to the self-attention block, mirroring the approach employed in vision transformers.[16] The self-attention block assigns weights to each set embedding for classification. The cross-attention pooling and self-attention block modules are

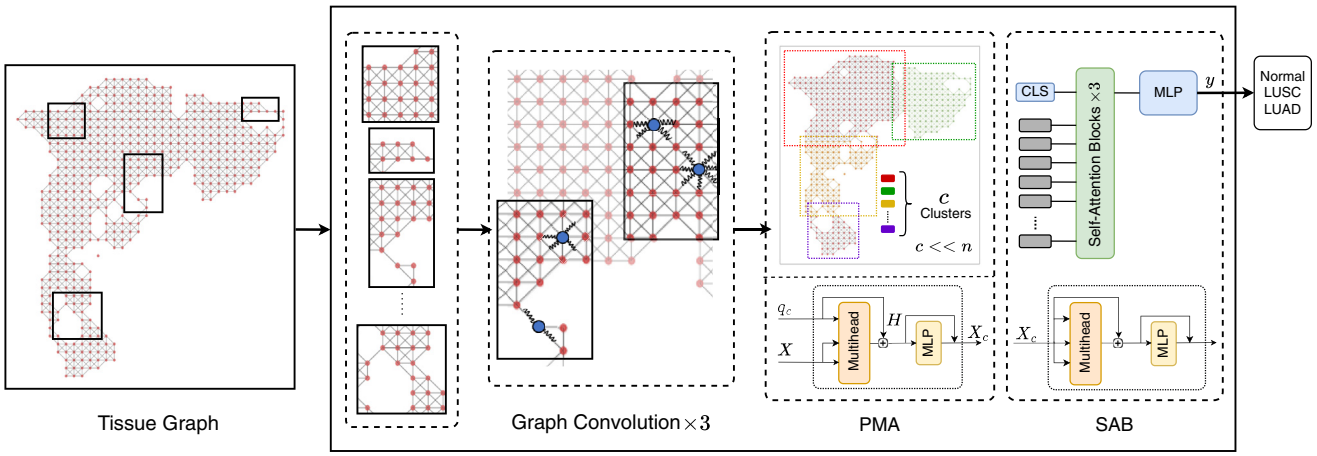

Figure 2 Graph Perceiver Network (GRAPE-Net) architecture. Each whole slide image (WSI) is represented as an undirected, unweighted tissue graph, where each node is an embedding of features in an image patch. Our proposed GRAPE-Net aggregates neighborhood information via graph convolutions while preserving spatial context. To efficiently find the neighborhood interaction between the different tissue regions in the tumor microenvironment, the graph is clustered to C overlapping sets by pooling the nodes using multihead cross-attention pooling (PMA) block. The fixed set embeddings are then given as input to the self-attention block (SAB), which learns morphologic interactions with its neighborhood and relevance of each set toward the specific output label of interest. The aggregated attentions are then given as input to a multilayer perceptron (MLP) for classification of the WSI as normal, lung adenocarcinoma (LUAD), or lung squamous cell carcinoma (LUSC). Here, the layer-specific feature representation of the graph is denoted as H. A classification token (CLS) is added to serve as the entire WSI representation, which is used for classification and computing the relevance of each graph node toward the prediction.

alternatively stacked, allowing the set embeddings to extract information from the input WSI graph based on the previous layer's learnings. This strategy preserves positional information of the graph nodes in the sets. Weight-sharing strategy of the perceiver was followed to optimize computations and reduce memory constraints. The final self-attention block is followed by a multilayer perceptron, which facilitates multiclass prediction. The classification token embedding from the self-attention block represents the slide-level aggregation vector of the weighted sets. The use of the classification token instead of averaging the embeddings of sets facilitates explainability in the network.

The current model introduced novelty by integrating the graph module with the perceiver architecture, enabling sparse graph computations on the visual tokens and computationally efficient modeling of a graph for downstream tasks. For the use case, it served as an end-to-end, amortized clustering algorithm, which clustered distinct regions of the tissue into overlapping sets, preserving vital spatial interactions that enhanced the comprehension of the biological processes withing tumor tissues.

Experimental Design

GRAPE-Net was trained as a three-label classifier (LUAD, LUSC, or normal). The study used TCGA dataset, with stratified patient-level sampling and fivefold cross-validation (Supplemental Figure S1) with internal testing. The receiver operating characteristic and precision-recall curves for each class were computed along with the accuracy, precision, sensitivity, and specificity as performance metrics (Figure 1C). To ensure the classifier is robust to dataset-specific batch effects, it was evaluated on the unseen CPTAC cohort. Model weights of the fold with the highest sensitivity (recall) on the CPTAC cohort was used for further *post hoc* analysis.

To evaluate the effectiveness of the proposed network, the study compared the model performance with a state-of-the art lung tumor classification model, graph-transformer network. Because the model leverages graph convolutions, a traditional graph classifier, Graph Isomorphism Network,[10] was selected for comparison. To ensure fair comparison, CTransPath[8] was used as the common feature extractor for all methods and the same cross-validation parameters were used for training. The hyperparameters stated in the respective published research of graph-transformer network and Graph Isomorphism Network were used to achieve the best performance on TCGA and CPTAC cohorts. Optimal hyperparameters for the network configuration involved a hidden dimension of 64, a graph block with three GIN Convolution layers, a cross-attention pooling configuration of 200 sets, and a three self-attention layers per self-attention block. Eight heads were used for multihead attention training. To avoid overfitting, dropouts with a rate of 0.2 were used in feed-forward blocks, and binary cross-entropy loss to enhance robustness. Training was completed over 30 epochs in eight-sample minibatches, using early stopping, a 0.0003 initial learning rate, a step scheduler, and the Adam optimizer[17] for expedited convergence. All the experiments were performed on a single GeForce RTX 2080Ti 11 Gb workstation (Nvidia, Santa Clara, CA).

For *post hoc* analysis, the WSI-level features for each sample were extracted by giving the respective graphs as input to the trained three-label GRAPE-Net classifier. The 64-dimensional features from the final layer of the network were considered as the representations for each WSI and used for further analysis. Uniform manifold approximation and projection manifold representations[18] and principal component analysis (PCA) representations were plotted to illustrate the relationship between image features and sample phenotypes (Figure 1C). Sample phenotypes were grouped as follows: normal adjacent lung tissue, LUSC, and LUAD for the lung tissue data (CPTAC). Within CPTAC, differences were examined in the principal components associated with LUSC predominant histologic patterns (keratinizing or nonkeratinizing) and LUAD predominant histologic patterns grouped as follows: solid/micropapillary/cribriform or lepidic/acinar/papillary. Within the CIS samples, differences in the principal components were examined by progression status. *Post hoc* analysis results suggest that the model is robust to cohort-based batch effects between Roswell and UCL cohorts.

Explainability Analysis

GRAPE-Net learns the classification label-relevant contributions of the tissue regions during training. These explanations are shown as heatmaps using a relevance propagation mechanism[19] (Figure 1C). This mechanism is a generic attention-based model explanation for bimodal transformers. It uses attentions and gradients from the self-attention block and cross-attention pooling blocks to produce relevancy maps for each interaction between the sets C and the assigned nodes N. The relevancy map from cross-attention layers provides the individual contributions of each node (patch) for the final prediction instead of distributing the contribution of each set as an average to all the nodes within the set. These relevancy maps are constructed on the WSI using the adjacency matrix A and coordinates of all the patches on the WSI.

Data Availability

TCGA data used here are in part based on the data generated by TCGA Research Network (*https://www.cancer.gov/tcga*, last accessed April 11, 2021). CPTAC data used here were downloaded via the National Cancer Institute's Cancer Imaging Archive [lung squamous cell carcinoma images: CPTAC-LSCC data set (Version 15). The Cancer Imaging Archive, 2018, *https://doi.org/10.7937/K9/TCIA.2018.6EMUB5L2*; and LUAD: CPTAC-LUAD data set (Version 12). The Cancer Imaging Archive, 2018, *https://doi.org/10.7937/K9/TCIA.2018.PAT12TBS*; both last

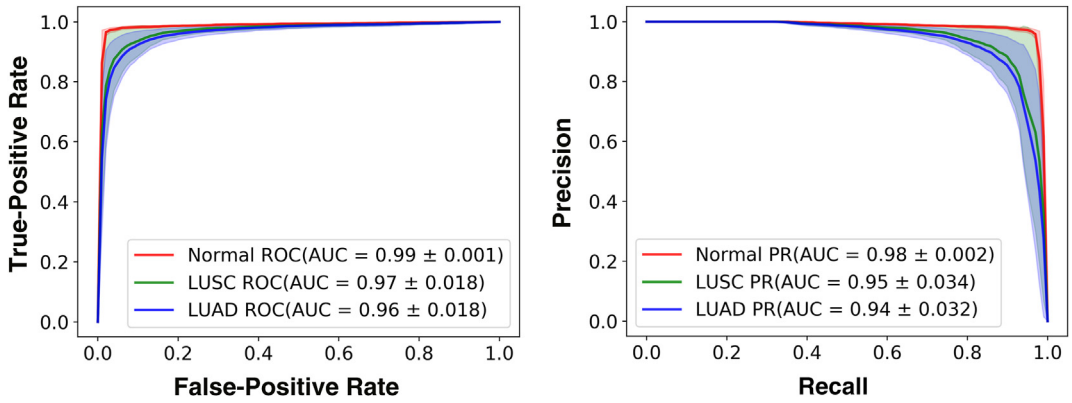

Figure 3 Classification performance. Receiver operating characteristic (ROC) and precision-recall (PR) curves showcasing the performance of Graph Perceiver Network (GRAPE-Net) toward multiclass classification on Clinical Proteomic Tumor Analysis Consortium (CPTAC) external testing dataset. The mean ± SD area under the curve (AUC) score for each label [normal, lung adenocarcinoma (LUAD), and lung squamous cell carcinoma (LUSC)] is provided for the ROC and PR curves.

accessed June 2, 2021]. UCL data can be obtained online (*https://idr.openmicroscopy.org* with IDR0082, last accessed November 23, 2021). Roswell images will be accessible via the Human Tumor Atlas Network Data Portal (*https://humantumoratlas.org*) Data Release V5.1 using the file identifications provided in Supplemental Table S1.

Code Availability

Computer scripts and manuals are made available on GitHub (*https://github.com/vkola-lab/ajpa2024*, last accessed February 13, 2024).

Statistical Analysis

The performance of GRAPE-Net was compared with other approaches for the tumor classification task in lung cancer (train on TCGA test on CPTAC), highlighting its efficiency while performing like the state-of-the-art methods. Uniform manifold approximation and projection and PCA clustering analyses were conducted using Scanpy framework.[20] Clustering performance was evaluated using adjusted Rand score and adjusted mutual index, with higher scores displaying good clustering for user-specified labels. Statistical significance between principal components and sample characteristics was assessed through two-tailed *t*-tests and Spearman correlation coefficient experiments, with analyses performed at a significance $P = 0.05$.

Results

GRAPE-Net Performance

GRAPE-Net demonstrated high classification performance on TCGA [mean ± SD area under the receiver operating curve (AUROC) = 0.98 ± 0.01; and mean ± SD area under the precision recall curve (AUPRC) = 0.98 ± 0.01] and CPTAC data (mean ± SD AUROC = 0.93 ± 0.01; and mean ± SD AUPRC = 0.95 ± 0.02) (Figure 3), using significantly fewer resources compared with two published frameworks (Table 1) (GRAPE-Net has approximately 202,000 whereas graph-transformer network has approximately 664,000 trainable parameters). Fivefold cross-validation indicated good agreement between the true and predicted tumor classes on both TCGA and CPTAC data, across different folds, displaying the robustness of the model to data bias used for training and testing. For each sample,

Table 1 Performance Metrics for the Three-Label (Normal versus LUAD versus LUSC) Classification Task

		Precision			Recall/sensitivity			Specificity			Accuracy
Method	Data	Normal	LUAD	LUSC	Normal	LUAD	LUSC	Normal	LUAD	LUSC	All
GRAPE-Net	TCGA	97.6 (0.8)	89.7 (2.1)	88.7 (2.6)	99.0 (0.5)	88.8 (2.1)	88.8 (2.2)	99.1 (0.3)	93.5 (1.3	93.8 (1.0)	91.4 (1.0)
	CPTAC	95.9 (2.2)	79.7 (3.2)	88.8 (2.7)	95.3 (2.0)	89.6 (3.6)	78.1 (5.1)	97.8 (1.3)	89.0 (2.4)	95.0 (1.5)	87.8 (1.1
GTP	TCGA	96.0 (1.5)	90.5 (2.3)	90.7 (1.9)	98.5 (1.0)	89.8 (1.5)	89.7 (1.6)	98.6 (0.6)	94.0 (1.4)	94.9 (0.8)	92.0 (1.2)
	CPTAC	92.8 (7.5)	82.9 (4.5)	91.1 (4.3)	96.9 (1.7)	88.9 (5.0)	78.5 (7.4)	95.5 (5.3)	91.0 (3.3)	95.9 (2.6)	88.2 (2.1)
GIN	TCGA	97.3 (1.2)	89.7 (2.5)	85.7 (3.6)	98.4 (1.1)	85.9 (3.3)	89.2 (2.8)	99.0 (0.4)	93.8 (1.6)	92.1 (1.4)	90.3 (1.7)
	CPTAC	95.9 (0.2)	82.5 (0.4)	78.3 (0.3)	89.0 (6.8)	83.5 (3.5)	82.8 (4.4)	97.9 (1.2)	91.5 (2.3)	88.5 (2.4)	85.2 (1.6)

Mean performance metrics (percentages) with SD values in parentheses are reported for the GRAPE-Net method used in this study and two other previously published state-of-the-art methods in TCGA and CPTAC datasets.

CPTAC, Clinical Proteomic Tumor Analysis Consortium; GIN, Graph Isomorphism Network; GRAPE-Net, Graph Perceiver Network; GTP, graph-transformer network; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; TCGA, The Cancer Genome Atlas.

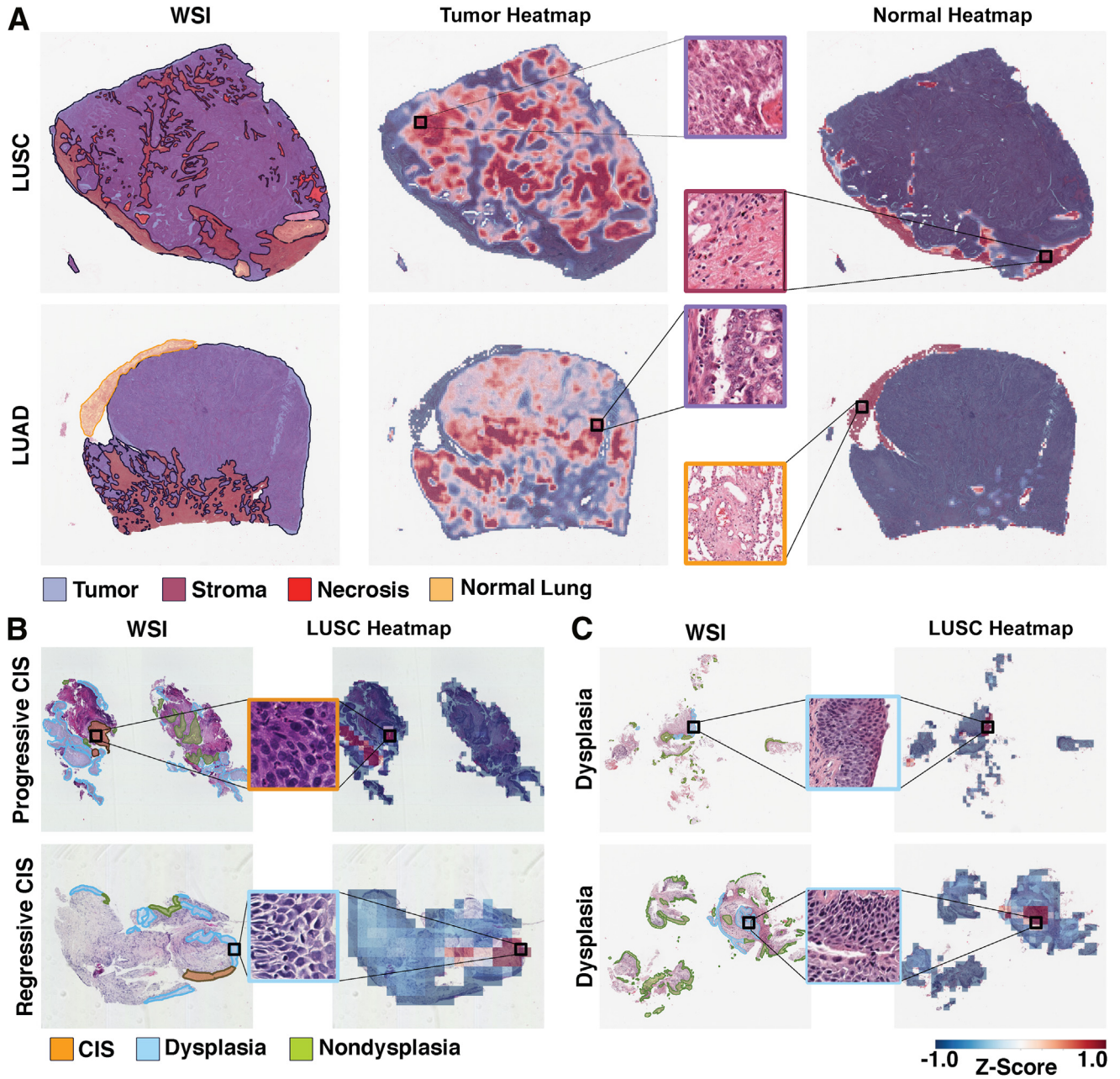

Figure 4 Graph Perceiver Network identifies tissue regions that correspond with pathologic annotations. **A:** The network identified regions annotated as tumor and normal on lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) whole slide images (WSIs) from the Clinical Proteomic Tumor Analysis Consortium (CPTAC). CTPAC images were acquired at ×20 magnification, the LUAD image is 23.6 × 17.9 mm, and the LUSC image is 24.6 × 21.1 mm. Representative patches are 126.5 × 126.5 μm. **B** and **C:** The network was also used to classify premalignant lesions (PMLs) as one of the three labels. The LUSC-specific heat maps for PMLs predicted to be LUSC identified regions of dysplasia and carcinoma *in situ* (CIS) as well as other regions likely contributing to its ability to separate CIS progressors from CIS regressors. **B:** The University College London images were acquired at ×40 magnification, where the progressive image is 8.6 × 5.6 μm, the regressive image is 4.7 × 3.1 μm, and the representative patches are 58.2 × 58.2 μm. **C:** The Roswell Park Comprehensive Cancer Institute images were acquired at ×20 magnification, where the top image is 8.0 × 7.1 μm, the bottom image is 5.0 × 4.8 μm, and the representative patches are 128.8 × 128.8 μm. Z-score color bar represents the contribution of the patches towards the final prediction, with scores close to 1 providing high contributions while scores close to −1 are not relevant to the prediction.

three separate class activation maps were produced to highlight the areas of normal and tumor tissue (LUAD or LUSC) in each WSI. The generated class-specific heat maps were consistent with the expert pathologist annotations from 20 CPTAC samples (selected at random), accurately pinpointing areas corresponding to same-label annotations (Figure 4A and Supplemental Figure S2). These results illustrate that GRAPE-Net could distinguish between the non−small-cell lung cancer tumor subtypes and the normal tissue regions within a WSI. Therefore, its ability to detect subtle

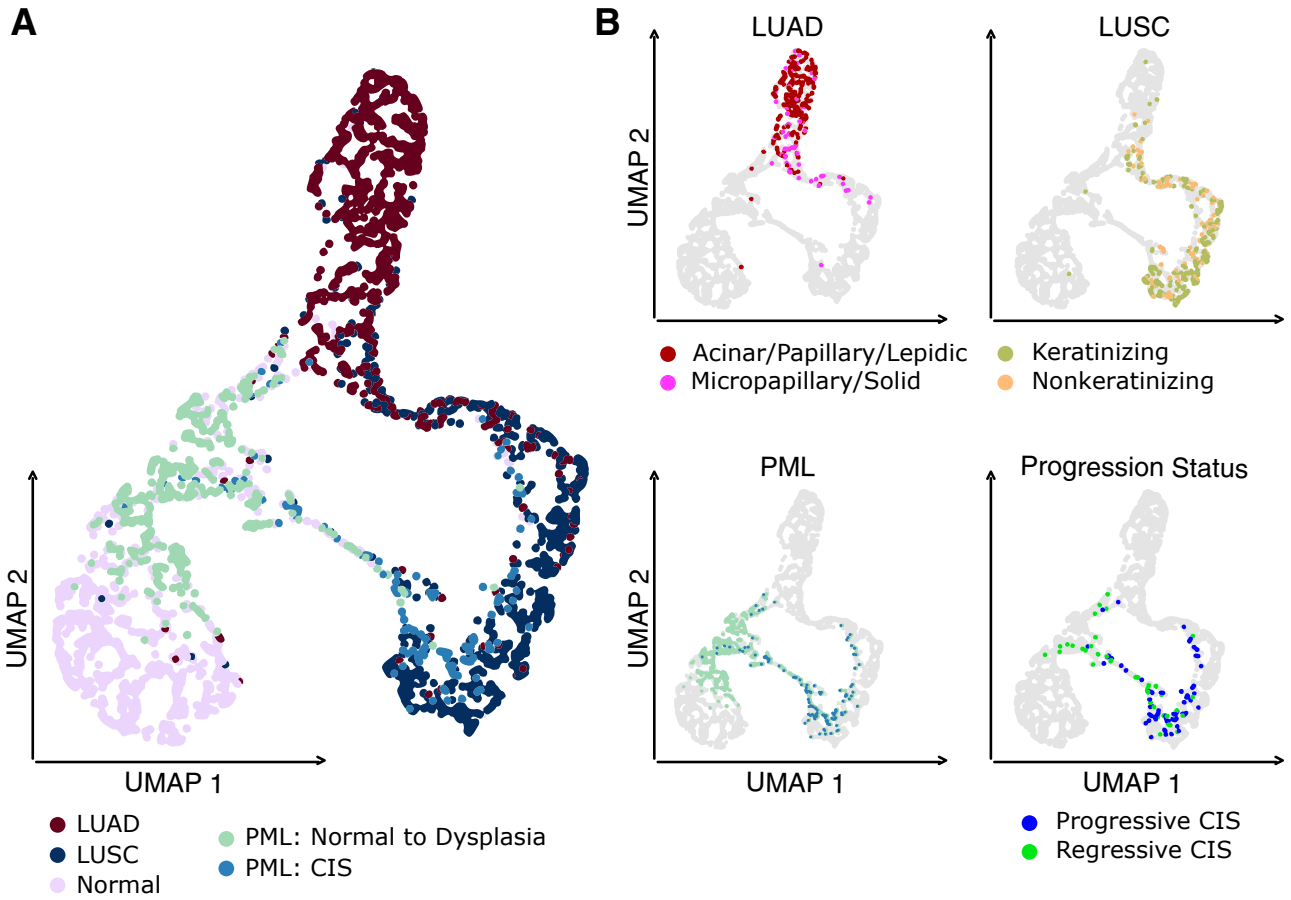

Figure 5 Stratification of lung tumors and premalignant lesions (PMLs) relates to histologic features and outcome. **A:** Uniform manifold approximation and projection (UMAP) plots of Clinical Proteomic Tumor Analysis Consortium, University College London, and Roswell Park Comprehensive Cancer Institute whole slide image (WSI) features, stratified samples from normal to invasive carcinoma. **B:** The model separated lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) WSIs by tumor histologic patterns and carcinoma *in situ* (CIS) PMLs by outcome.

distinctions between tumor histologic patterns and PML histology was explored further.

Stratification of Tissues by Morphologic Features

The authors hypothesized that GRAPE-Net could differentiate between predominant histologic patterns of LUSC and LUAD tumors and lung squamous PML histology by discerning subtle morphologic features. To test this, the WSI latent features from the CPTAC, UCL, and Roswell samples were used to perform uniform manifold approximation and projection–based clustering (Figure 5). Clustering segregated normal, LUAD, and LUSC with a high adjusted Rand score of 0.69 and adjusted mutual information score of 0.64. PCA on the latent features (Supplemental Figure S3A) revealed significant principal component (PC) 1 differences between keratinizing and nonkeratinizing LUSC tumors (PC1: $P = 0.01$), with the latter PC2 values like those of LUAD tumors. The LUAD tumors showed significant PC1 differences between aggressive (solid/micropapillary) and nonaggressive (lepidic/acinar/papillary) histologic patterns (PC1: $P = 5.4 \times 10^{-14}$). Notably, LUAD (solid/micropapillary) tumors exhibited PC2 values similar to LUSC tumors, as solid LUAD tumors require immunohistochemical analysis for diagnosis as they can be difficult to distinguish from LUSC tumors (Supplemental Figure S3B).[21]

Uniform manifold approximation and projection and PCA revealed most endobronchial biopsies were proximal to the normal cluster, with some high-grade dysplasia and CIS PMLs clustering with LUAD and LUSC cases ($P < 0.01$ for PML versus LUAD, PML versus LUSC, on both PC1 and PC2). Roswell biopsies did not show significant separation by histology, molecular subtype, or progression status (Supplemental Figure S3D).

All CIS samples from Roswell and some from UCL grouped with LUSC tumor cases (75% of CIS samples were classified as LUSC tumors with $P < 0.01$ for CIS versus LUAD, CIS versus LUSC, on both PC1 and PC2). Among the UCL biopsies with progression status to LUSC, the model robustly stratified regressive versus progressive CIS. GRAPE-Net performance predicting regressive CIS as normals and progressive CIS as tumors was as follows: precision = 0.68, recall = 0.67,

Table 2 Significance Statistics for Comparing Model Feature Space with Sample Characteristics for CIS Premalignant Lesions

Test description	Test statistic	*P* value (<0.05)
t-Tests		
Lymphocyte density in CIS regions (pLUSC vs pNORMAL)	**−5.18**	**1.03×10^{-6}**
Epithelial density in CIS regions (pLUSC vs pNORMAL)	**3.94**	**0.000145**
PC1 (progression vs regression)	**2.82**	**0.005721**
PC2 (progression vs regression)	**−4.59**	**1.18×10^{-5}**
Spearman correlation coefficient		
Lymphocyte density in CIS regions vs PC1	−0.099	0.2932
Lymphocyte density in CIS regions vs PC2	**0.297**	**0.0024**
Epithelial density in CIS regions vs PC1	**0.506**	**0.0002**
Epithelial density in CIS regions vs PC2	**−0.749**	**0.0002**
Classification metrics		
Precision	0.68	—
Recall	0.67	—
Specificity	0.38	—
Accuracy	0.67	—
AUC	0.69	—

To demonstrate Graph Perceiver Network's ability to stratify premalignant lesions, we show the model's classification performance predicting regressive CIS as normals (pNORMAL) and progressive CIS as tumors (pLUSC). Additional test statistics were also done with other metadata characteristics, with $P < 0.05$ considered significant. Significant scores are highlighted in bold.

—, not applicable; AUC, area under the curve; CIS, carcinoma *in situ*; PC, principal component.

specificity = 0.38, accuracy = 0.67, and area under the curve = 0.69 (Table 2). Additionally, in prior work on the UCL dataset,[4] an automated deep learning pipeline was used to quantify the density of lymphocytes, epithelial cells, and stromal cells in the CIS and stromal regions. The study showed a significant decrease in lymphocyte density ($P = 0.02$) and increase in epithelial cell density ($P = 0.005$) in CIS regions in progressive versus regressive samples. Using GRAPE-Net, a similar significant decrease in lymphocyte density ($P = 1.03 \times 10^{-6}$) and an increase in epithelial cell density ($P = 1.4 \times 10^{-4}$) were observed between CIS samples predicted as tumors versus normals. Spearman correlation between the PC values and the respective CIS region lymphocyte and epithelial cell densities suggested that PC1 had a high positive correlation and PC2 had a high negative correlation with epithelial density. PC2 was also positively correlated with lymphocyte density (Table 2). Both PC1 and PC2 were significantly different between CIS progressive versus regressive lesions ($P = 0.005$ and $P = 1.1 \times 10^{-5}$, respectively) (Supplemental Figure S3C). These results suggest that the model learns certain morphologic characteristics pertaining to CIS progression status highlighted using relevance heat maps, as shown in Figure 4B and Supplemental Figure S4.

Discussion

This work introduced the GRAPE-Net that was designed to stratify bronchial PMLs, which are precursors to invasive LUSC. The goal was to distinguish PMLs with a high likelihood of progressing to LUSC from those that may regress without intervention. Using hematoxylin and eosin–stained WSIs, GRAPE-Net stratified PMLs across a continuum from normal to tumor tissues. Using four distinct datasets, the model generated latent features observable across sample types (resected lung tissue versus endobronchial biopsies) that were associated with tumor histologic patterns and CIS PML progression status to invasive carcinoma.

GRAPE-Net's performance was demonstrated through its high classification accuracy on TCGA and CPTAC lung resection tissue datasets, coupled with its efficient generation of pathologist-aligned, class-specific heatmaps. PCA on the WSI latent features demonstrated separation within LUAD and LUSC samples based on predominant histologic pattern. On the endobronchial biopsy WSIs, GRAPE-Net identified CIS lesions that progressed to LUSC. The principal components of the latent features were associated with decreases and increases in lymphocyte and epithelial cell densities within CIS regions, respectively, in progressing versus regressing lesions. GRAPE-Net's predictive capacity on CIS lesions may have clinical utility in prioritizing patients for chemoprevention trials or potentially shortening screening intervals via bronchoscopy.

This framework currently does not differentiate between biopsies with and without bronchial dysplasia, as observed in Figure 4C and Supplemental Figure S5, likely because of a paucity of training data on the cellular morphology changes characteristic of various dysplasia stages (ie, mild, moderate, and severe). This limitation underlines the necessity for future research to expand model training with a broader array of bronchial PMLs. Additionally, it will be important to test this model on lung adenomatous premalignant lesions (including atypic adenomatous hyperplasia and adenocarcinoma *in situ*) as biopsies of these lesions become available via robotic bronchoscopy. The assessment of GRAPE-Net across more extensive cohorts and the integration of additional bulk, single-cell, and spatial-resolved molecular data into the model is expected to refine PML stratification further.

Acknowledgment

We thank Dr. Erin Kane for coordinating the upload of the Roswell slides to the Human Tumor Atlas Network public data portal.

Author Contributions

J.E.B. and V.B.K. conceptualized the study; R.H.G., Y.Z., V.B.K., and J.E.B. developed methods; R.H.G., V.B.K., J.E.B., and E.J.B. interpreted results; E.J.B., D.T.M., and E.J.G. performed pathologic annotation; M.E.R., S.A.M., and J.E.B. acquired samples and curated clinical data; and R.H.G., V.B.K., and J.E.B. wrote the manuscript.

Disclosure Statement

J.E.B., V.B.K., S.A.M., and M.E.R. received commercial research grants from Janssen Pharmaceuticals.

Supplemental Data

Supplemental material for this article can be found at *http://doi.org/10.1016/j.ajpath.2024.03.009*.

References

1. Beane JE, Mazzilli SA, Campbell JD, Duclos G, Krysan K, Moy C, Perdomo C, Schaffer M, Liu G, Zhang S, Liu H, Vick J, Dhillon SS, Platero SJ, Dubinett SM, Stevenson C, Reid ME, Lenburg ME, Spira AE: Molecular subtyping reveals immune alterations associated with progression of bronchial premalignant lesions. Nat Commun 2019, 10:1856
2. Merrick DT, Edwards MG, Franklin WA, Sugita M, Keith RL, Miller YE, Friedman MB, Dwyer-Nield LD, Tennis MA, O'Keefe MC, Donald EJ, Malloy JM, van Bokhoven A, Wilson S, Koch PJ, O'Shea C, Coldren C, Orlicky DJ, Lu X, Baron AE, Hickey G, Kennedy TC, Powell R, Heasley L, Bunn PA, Geraci M, Nemenoff RA: Altered cell-cycle control, inflammation, and adhesion in high-risk persistent bronchial dysplasia. Cancer Res 2018, 78:4971–4983
3. Mascaux C, Angelova M, Vasaturo A, Beane J, Hijazi K, Anthoine G, Buttard B, Rothe F, Willard-Gallo K, Haller A, Ninane V, Burny A, Sculier J-P, Spira A, Galon J: Immune evasion before tumour invasion in early lung squamous carcinogenesis. Nature 2019, 571:570–575
4. Pennycuick A, Teixeira VH, AbdulJabbar K, Raza SEA, Lund T, Akarca AU, et al: Immune surveillance in clinical regression of preinvasive squamous cell lung cancer. Cancer Discov 2020, 10:1489–1499
5. Ning B, Tilston-Lunel AM, Simonetti J, Hicks-Berthet J, Matschulat A, Pfefferkorn R, Spira A, Edwards M, Mazzilli S, Lenburg ME, Beane JE, Varelas X: Convergence of YAP/TAZ, TEAD and TP63 activity is associated with bronchial premalignant severity and progression. J Exp Clin Cancer Res 2023, 42:116
6. Rozenblatt-Rosen O, Regev A, Oberdoerffer P, Nawy T, Hupalowska A, Rood JE, et al: Human tumor atlas network: the human tumor atlas network: charting tumor transitions across space and time at single-cell resolution. Cell 2020, 181:236–249
7. Otsu N: A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979, 9:62–66
8. Wang X, Yang S, Zhang J, Wang M, Zhang J, Yang W, Huang J, Han X: Transformer-based unsupervised contrastive learning for histopathological image classification. Med Image Anal 2022, 81:102559
9. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B: Swin transformer: hierarchical vision transformer using shifted windows. arXiv 2021, [Preprint] doi: 10.48550/arXiv.2103.14030
10. Xu K, Hu W, Leskovec J, Jegelka S: How powerful are graph neural networks? arXiv 2018, [Preprint] doi: 10.48550/arXiv.1810.00826
11. Jaegle A, Gimeno F, Brock A, Zisserman A, Vinyals O, Carreira J: Perceiver: general perception with iterative attention. arXiv 2021, [Preprint] doi: 10.48550/arXiv.2103.03206
12. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I: Attention is all you need. arXiv 2017, [Preprint] doi: 10.48550/arXiv.1706.03762
13. Zheng Y, Gindra RH, Green EJ, Burks EJ, Betke M, Beane JE, Kolachalama VB: A graph-transformer for whole slide image classification. IEEE Trans Med Imaging 2022, 41:3003–3015
14. Buterez D, Janet JP, Kiddle SJ, Oglic D, Liò P: Graph neural networks with adaptive readouts. arXiv 2022, [Preprint] doi: 10.48550/arXiv.2211.04952
15. Bianchi FM, Grattarola D, Alippi C: Spectral clustering with graph neural networks for graph pooling. arXiv 2019, [Preprint] doi: 10.48550/arXiv.1907.00481
16. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N: An image is worth 16x16 words: transformers for image recognition at scale. arXiv 2020, [Preprint] doi: 10.48550/arXiv.2010.11929
17. Kingma DP, Ba J: Adam: a method for stochastic optimization. arXiv 2014, [Preprint] doi: 10.48550/arXiv.1412.6980
18. McInnes L, Healy J, Melville J: UMAP: uniform manifold approximation and projection for dimension reduction. arXiv 2018, [Preprint] doi: 10.48550/arXiv.1802.03426
19. Chefer H, Gur S, Wolf L: Generic attention-model explainability for interpreting bi-modal and encoder-decoder transformers. arXiv 2021, [Preprint] doi: 10.48550/arXiv.2103.15679
20. Wolf FA, Angerer P, Theis FJ: SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 2018, 19:15
21. Kadota K, Nitadori J-I, Rekhtman N, Jones DR, Adusumilli PS, Travis WD: Reevaluation and reclassification of resected lung carcinomas originally diagnosed as squamous cell carcinoma using immunohistochemical analysis. Am J Surg Pathol 2015, 39:1170–1180