Tumor Micro-environment Interactions Guided Graph Learning for Survival Analysis of Human Cancers from Whole-slide Pathological Images

Wei Shao YangYang Shi Daoqiang Zhang * JunJie Zhou Peng Wan

College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics
Key Laboratory of Brain-Machine Intelligence Technology, Ministry of Education, Nanjing, China
MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, Nanjing, China.

Abstract

*The recent advance of deep learning technology brings the possibility of assisting the pathologist to predict the patients' survival from whole-slide pathological images (WSIs). However, most of the prevalent methods only worked on the sampled patches in specifically or randomly selected tumor areas of WSIs, which has very limited capability to capture the complex interactions between tumor and its surrounding micro-environment components. As a matter of fact, tumor is supported and nurtured in the heterogeneous tumor micro-environment(TME), and the detailed analysis of TME and their correlation with tumors are important to in-depth analyze the mechanism of cancer development. In this paper, we considered the spatial interactions among tumor and its two major TME components (*i.e., *lymphocytes and stromal fibrosis) and presented a Tumor Micro-environment Interactions Guided Graph Learning (TMEGL) algorithm for the prognosis prediction of human cancers. Specifically, we firstly selected different types of patches as nodes to build graph for each WSI. Then, a novel TME neighborhood organization guided graph embedding algorithm was proposed to learn node representations that can preserve their topological structure information. Finally, a Gated Graph Attention Network is applied to capture the survival-associated intersections among tumor and different TME components for clinical outcome prediction. We tested TMEGL on three cancer cohorts derived from The Cancer Genome Atlas (TCGA), and the experimental results indicated that TMEGL not only outperforms the existing WSI-based survival analysis models, but also has good explainable ability for survival prediction.*

1. Introduction

Cancer is an aggressive disease seriously affecting the health and living quality of human beings[16][35]. It is reported that the number of cancer patients will dramatically increased and is expected to achieve 28.4 million cases in 2040 [12]. Hence, it is very important to accurately predict the patients' clinical outcome including the survival time, which in turn can help doctors design individual therapeutic plan at the early stage [3][15].

In the past decade, image-based technologies have shown their great potential in healthcare researches. Among a wide variety of imaging bio-markers including CT images[21], Ultrasound images [28] and MRI images [34] collected from the patients, the whole-slide pathological images (WSIs) are generally considered as the golden standard for the diagnosis and prognosis of human cancers since they can reveal the morphology and arrangement of cells reflecting the progression of human cancers[1][39]. Because of the heterogeneous patterns shown in WSIs [26] that will lead to the large inter-observation variation among different pathologists, it is necessary to design machine learning models to help pathologists assess the survival risks of human cancers.

Recently, with the rapid development of the deep learning technology [23], training WSI-based deep learning models for cancer survival analysis have drawn much attentions [37][33]. In comparison with the traditional survival analysis models requiring the heavy burden of feature engineering, the deep learning models can automatically learn survival-associated representations from WSIs to improve the prognosis results. However, one major challenge for analyzing WSIs is that a WSI is usually with huge size (*e.g.,* 100,000-by-100,000 pixels), and it is impossible to directly feed them into the deep neural network for model training .

To address the above challenges, the existing studies can be categorized into two folds, *i.e.,* patch-level and WSI-level annotation methods. The patch-level annotation methods [6] [42][13] need the experts to annotate the key patches such as the tumor patches from the WSIs at first, and then used those patches for deep model training. However, such

*Corresponding Author

patch pre-selection step requires the extra annotation efforts from the pathologists, which is impractical especially in the case that a great number of WSIs need to be labeled. To reduce such annotation burden, another category of methods only requiring WSI-level annotations are developed. The WSI-level annotation methods [43][5] [25] [36] treat WSIs as bags while the patches in WSIs are regarded as instance. Then, the multi-instance learning (MIL) algorithms are designed to aggregate the patch-level representations to estimate the target of WSIs via the pooling [32] or attention strategy [40]. Besides the MIL methods, some studies attempted to apply the graph convolutional network to learn the global representation of WSIs [41][18][8], which can also achieve satisfied prognosis results.

Although much progress have been achieved, the existing studies have limited ability to capture the complex interactions among tumor and its micro-environment components since the selected patches are randomly generated or derived from the tumor areas of WSIs [41][18]. As a matter of fact, it is widely recognized that cancer is not only achieved by unlimited growth of the tumor cells, but also supported, stimulated, and nurtured by the tumor microenvironment (TME) mainly composed of stromal fibrosis and lymphocytes [11]. For instance, the existing studies have demonstrated that the prominent interactions between lymphocytes and tumors will result in low survival risk because of the immune system could control the tumor growth through activation of adaptive and innate immune mechanisms [30], while the rich intersections between tumor and stromal fibrosis usually promote the tumor progression since they can facilitate the invasion and metastasis of tumors [29]. It can be expected that better prognosis results can be achieved if the developed prognosis model takes the intersections among tumor and different TME components into consideration. Moreover, to the best of our knowledge, all the existing WSI-based graph learning algorithms directly applied the pre-trained network *i.e.,* ResNet101 [44], VGG16 [18] on the selected patches to extract node features. However, such node representations neglected to take the topological structure of TME into consideration, which is closely related to the development of cancer [11]

Accordingly, in this paper, we proposed a Tumor Micro-environment Interactions Guided Graph Learning (TMEGL) algorithm for the prognosis prediction of human cancer (https://github.com/tmegl/tmegl-v1). The main contribution of this study can be summarized into the following three aspects:

1) We developed a graph learning algorithm (*i.e.,* TMEGL) that considers the interactions among tumor and different TME components (*i.e.,* lymphocytes and stromal fibrosis) for survival analysis of human cancers.
2) We proposed a novel graph embedding algorithm preserving the spatial organization of tumor and different TME components extracted from the graph domain.
3) We developed a gated graph attention networks (GGATs) for predicting patients' survival from WSIs, where the gated graph convolution layers are used to update the node representation from the neighbourhood nodes with the same node type, while the graph attention layers are used to characterize the intersections among different types of nodes.

2. Related Work

2.1. Analyze WSIs based on Patch-level Annotation.

The patch-level annotation methods required the pathologists to select a bunch of patches that can represent the aggressiveness of the cancer for model training. Based on the selected patches derived from WSIs, the study in [7] has designed a computational pipelines that can extract various types of hand-crafted features to conduct the diagnosis and prognosis tasks. Recently, with the remarkable success of deep learning, many patch-level methods are implemented under deep neural network frameworks [42][13] [6]. For instance, Zhu et al [42] firstly developed a deep CNN model for predicting patients' survival from the pathological images. Cheng et al [6] applied the deep autoencoder to learn the cell representation from the extracted patches, followed by applying the Delaunay triangulation graph to explore the difference of cell-cell interactions between low and high survival risk patients. Moreover, by considering the number of annotated patches is inadequate for training an effective deep learning model, some patch-based methods also adopted the data augmentation [14][24] or transfer learning strategy [38] to help improve the prognosis performance of human cancers.

2.2. Analyze WSIs based on WSI-level Annotation.

Different from the patch-level methods requiring the huge annotation efforts from the pathologists, the WSI-level methods only relying on WSI-level annotations for analyzing pathological images[43][5][25][36]. Most of the existing WSI-level methods are designed under the multi-instance learning (MIL) framework, where a WSI is treated as bag and the instances refer to the patches sampled from the WSI. In order to efficient aggregate patch-level information into WSI-level representations, the studies in [40][22][20] designed various ways of global pooling over patch-based instances. In addition, Li *et al* [9] have proposed a dual-stream multiple instance learning network for the classification of WSIs and Chen *et al* [5] have presented the Transformer based method taking the position information into consideration for survival analysis of human cancers. Besides the MIL framework, several studies [18][41][8][10][10] considered the spatial structure among the extracted patches and applied the graph convolutional

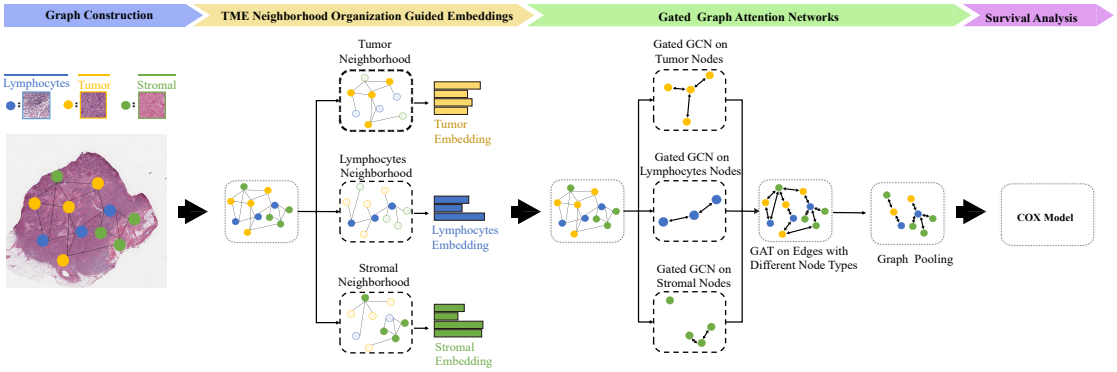

Figure 1. The flowchart of the proposed TMEGL method. Firstly, the patches with different types (*i.e.,* Lymphocytes, Tumor and Stromal) are extracted from the whole-slide pathological image for graph construction. Secondly, a tumor micro-environment organization guided graph embedding algorithm is proposed to learn node representations from the extracted patches. Thirdly, a gated graph attention network is presented to capture the intersections among different types of nodes. Finally, the Cox model is applied for survival prediction.

network (GCN) to predict the clinical outcome of cancer patients. However, these graph learning methods either adopted the pre-trained networks or graph structure information for node representation, which have limited ability to reveal the topological organization of the tumor micro-environment for survival prediction.

3. Method

Fig. 1 shows the schematic diagram of our proposed method, which is consisted of the following four major steps. Specifically, for each whole-slide pathological image (WSI), we firstly extracted different types of image patches (*i.e.,* Tumor, Lymphocytes and Stromal Fibrosis) as nodes to construct a graph. Then, we considered the tumor micro-environment (*i.e.,* TME) structure of each node, and proposed a TME neighborhood organization guided graph embedding algorithm to learn node representations. Based on the learned node embeddings, a gated graph attention networks (GGAT) is applied to characterize the intersections among tumor and different TME components. Finally, the Cox model is applied for the survival prediction task.

3.1. WSI Pre-processing and Graph Construction

Since each of the collected WSIs is of large size (*e.g.,* 100,000-by-100,000 pixels). We first divided it into non-overlap patches with the size of 512×512. Then, by considering the generated patches may not contain enough tissues, we selected the patches whose image density are larger than 0.7 for further analysis [32]. Here, the image density is calculated as the percentage of non-white (at least one of the red, green, and blue values was below 200 in the 24-bit RGB color space) pixels in that patch. Next, based on the annotated dataset released by [1], we followed [1] and trained a U-net++ network for the semantic segmentation of tumor, stromal and lymphocytes regions in each patch. Both the experimental results reported in [1] and our study indicated that the utilized tissue segmentation model could achieve to the dice ratio of 0.858, which are promising for further analysis. After getting the pixel-level segmentation results of each valid patch, we calculated its tumor, lymphocytes and stromal area ratios and selected 300 patches with the largest ratios for each patch type.

Based on the extracted tumor, lymphocytes and stromal patches, we constructed a graph for each WSI. Here, given patches as nodes, we adopted the K-Nearest Neighbors (KNN) algorithm to construct the graph basing on the Euclidean distances between patch pairs, and the number of nearest neighbors for each node is set as 50.

3.2. TME Neighborhood Organization Guided Graph Embedding for Nodes Representations

For each patch in the constructed graph, the existing studies usually applied the pre-trained network (*e.g.,* ResNet-101 and VGG-16) to acquire its node representation[44][18] . However, these feature extraction strategies cannot measure nodes similarity based on their tumor micron-environment (TME) organization, and it is widely recognized that the topological organization of TME plays an important role in the progression of human cancers [11]. For instance, a tumor region surrounded with tumor-infiltration lymphocytes will provide more chance to present immune response [30] to slow down the tumor progression, and thus its representation should be different from regular tumor patches. For another example, since the stromal tissues usually reveal immunosuppressive property during cancer progression [2], the representations of two lymphocyte patches should be similar if they are all surrounded with stromal fibrosis. In order to preserve such topological similarity in TME, we proposed a TME neighborhood organization guided graph embedding algorithm to separately learn the representations

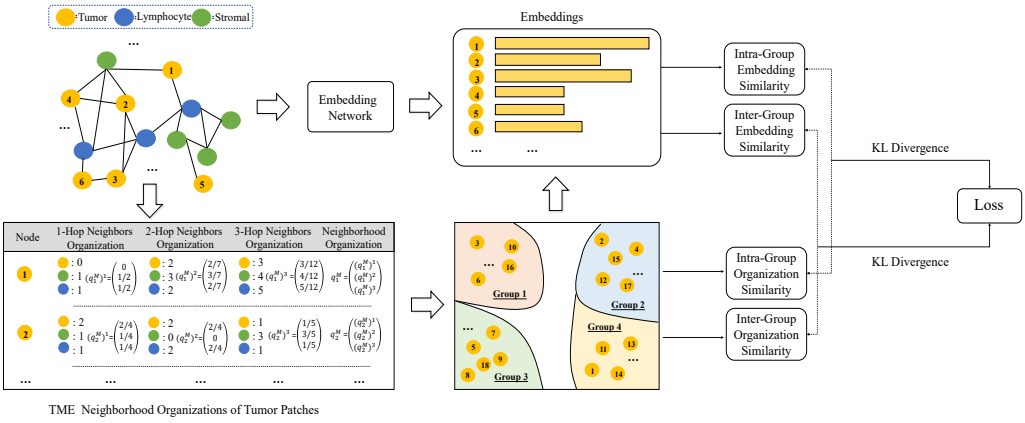

Figure 2. The Scheme of the Proposed TME Neighborhood Organization Guided Graph Embedding Method on Tumor Nodes.

of each type of nodes. In what follows, we will introduce the proposed graph embedding algorithm on tumor nodes (shown in Fig.2), which can be easily extended to other node types (*i.e.,* Lymphocytes and Stromal).

More specifically, let $\{v_1^M, v_2^M, ..., v_{300}^M\}$ denotes the extracted 300 tumor nodes in graph G constructed from a WSI. We explored the similarities among them by considering their TME neighborhood organizations. Here, the TME organization of v_i^M with k-hop can be represented as:

$$(q_i^M)^k = [\frac{(N_i^M)^k}{(N_i)^k}, \frac{(N_i^L)^k}{(N_i)^k}, \frac{(N_i^R)^k}{(N_i)^k}] \tag{1}$$

Where $(N_i)^k$ indicates the number of nodes that k-hop away from x_i^M, $(N_i^M)^k, (N_i^L)^k$ and $(N_i^R)^k$ represent the number of tumor, lymphocytes and stromal nodes in the k-hop neighborhood of x_i^M, respectively. Obviously, $(q_i^M)^k$ indicates the proportions of different TME components that can characterize the topological organizations of v_i^M. In this study, in order to describe the spatial organization of TME at multiple distance levels, we combined the TME neighborhood information from different hop ranges (*i.e.,* 1-hop, 2-hop and 3-hop) to represent each node as follows:

$$q_i^M = [(q_i^M)^1, (q_i^M)^2, (q_i^M)^3] \tag{2}$$

Then, the TME organization similarity between tumor nodes v_i^M and v_j^M can be obtained by:

$$s(v_i^M, v_j^M) = \frac{q_i^M (q_j^M)^T}{\|q_i^M\| \, \|q_j^M\|} \tag{3}$$

Based on the pairwise similarity information among different nodes, we applied the spectral clustering algorithm to aggregate all tumor nodes into B groups, and the elements in the t-th group can be denoted as:

$$P_t^M = \{P_{t1}^M, P_{t2}^M, ...P_{t|P_t^M|}^M\}, t = 1, 2, ..., B \tag{4}$$

where $|P_t^M|$ represents the cardinality of P_t^M. For the t-th group, its centroid node P_{tc}^M refers to the sample, which has the highest average similarity with other data points in P_t^M. Based on the clustering results, we defined a intra-group similarity vector $(S_t^o)^{tra} \in R^{|P_t^M|-1}$ that can measure the neighborhood organization similarity between the centroid node and the remaining elements in P_t^M:

$$(S_t^o)^{tra} = norm([s(P_{t1}^M, P_{tc}^M), ..., s(P_{t|P_t^M|}^M, P_{tc}^M)]) \tag{5}$$

Where the $norm(.)$ operator is used to normalize the vector to ensure that the sum of its elements equals to 1. On the other hand, we also considered the inter-group similarity among the derived B groups by calculating their corresponding centroid nodes similarities. Hence, the inter-group similarity vector for the t-th group $(S_t^o)^{ter} \in R^{B-1}$ can be defined as:

$$(S_t^o)^{ter} = norm([s(P_{1c}^M, P_{tc}^M), ..., s(P_{Bc}^M, P_{tc}^M)]) \tag{6}$$

Next, given the embeddings of the tumor nodes via the embedding network, we applied the same way to calculate their intra-group similarity $(S_t^e)^{tra}$ and inter-group similarity $(S_t^e)^{ter}$ according to the clustering results on graph domain (shown in Eq.(4)). Then, in order to learn the node embeddings that can preserve the TME organization similarity extracted from the graph, the objective loss function takes the form of Kullback–Leibler (KL) divergence on the B groups and is formulated as follows:

$$L = \sum_{t=1}^{B} KL((S_t^e)^{tra}, (S_t^o)^{tra}) + \lambda((S_t^e)^{ter}, (S_t^o)^{ter}) \tag{7}$$

As shown in Eq.(7), we optimize the objective to derive the TME organization associated embeddings that can preserve both intra-group and inter-group similarity information, and

λ is a regularization parameter could be tuned via cross-validation. Such TME neighborhood organization guided graph embedding algorithm can also be applied to calculate the representations for other types of nodes.

3.3. Gated Graph Attention Networks for Survival Analysis of Human Cancers

Although the interactions among different types of TME components is more relevant to cancer prognosis, recent studies also demonstrated that the interactions within the same type of TME components will affect their status [12]. Inspired by the above discovery, we proposed a Gated Graph Attention Network consisted of the Gated Graph Convolution (*i.e.,* GGC) layers and the Graph Attention (*i.e.,* GAT) layers to capture the interactions among different TME components.

Specifically, given patches as vertices for each graph G constructed from a WSI. We defined G^M, G^L, G^R as tumor, lymphocytes and stromal sub-graph of G, and all the nodes in $G^w(w \in \{M, L, R\})$ are with the same type. Based on the learned node embeddings in each sub-graph $d_i^w \in R^d (i = 1, 2, ...300)$, we firstly applied the Gated Graph Convolution layer [19] to update the representation of each node from its neighbourhood nodes with the same type as follows:

$$\begin{aligned} (h_i^w)^0 &= d_i^w || 0 \\ (a_i^w)^t &= \sum_{j \in N(i)} u_{ij} (h_j^w)^t \\ (h_i^w)^{t+1} &= GRU((a_i^w)^t, (h_i^w)^t) \end{aligned} \tag{8}$$

As shown in Eq.(8), $(h_i^w)^0 \in R^{2d}$ is initialized by copying the node embeddings into its first components and pads the rest with zeros. The remaining are Gated recurrent unit(GRU)-like updates that can incorporate information from the neighboring nodes $j \in N(i)$ with weight u_{ij} and the previous step to update the status of nodes i. Then, we applied the GAT layer to capture the interactions among different types of nodes. The inputs for GAT layers are $H = [H^M, H^L, H^R] \in R^{d \times 900}$, where $H^w = [h_1^w, h_2^w, ..., h_{300}^w] \in R^{d \times 300}, w \in \{M, L, R\}$ refers to the output features for each types of node after the GGC layer. The GAT layer produces a new set of node features $H^{'} = [(H^M)^{'}, (H^L)^{'}, (H^R)^{'}]$ via a linear transformation matrix $T \in R^{d \times d'}$ and $(H^w)^{'} \in R^{d' \times 300}$, $w \in \{M, L, R\}$. Then, a shared attention mechanism $\mathbb{R}^{d'} \times \mathbb{R}^{d'} \to \mathbb{R}$ is applied to calculate the normalized attention coefficients between node i with type w and node j with type z as follows:

$$\mathrm{e}_{ij}^{w,z} = a(Th_i^w, Th_j^z), w \neq z \tag{9}$$

Moreover, GAT normalizes e_{ij}'s values of node j by applying the Softmax function:

$$\alpha_{ij}^{w,z} = \frac{\exp(e_{ij}^{w,z})}{\sum\limits_{\substack{j \in N_i \\ v \neq w}} exp(e_{ij}^{w,v})} \tag{10}$$

Where N_i refers to the neighbors of node i. Then, the output feature for every node after the GAT layer can be represented as:

$$h_i^w = \sigma(\sum_{\substack{j \in N_i \\ w \neq z}} \alpha_{ij}^{w,z} T h_j^z) \tag{11}$$

As shown in Eq.(11), different from the GGC layer passing information between the same type of nodes, the GAT layer can capture the interactions among different types of nodes. In addition, in order to avoid over-fitting and identify survival-associated interactions among different types of TME components, we adopted SAGPooL[17] to select top $s(s \in [0, 100])$ percentage of nodes with high attention scores. Finally, at the last layer of GAT, a global attention pooling strategy introduced in [19] is applied to aggregate the node features into WSI-level feature for the following survival analysis task. Here, we followed the methods in [31] [8] applying the widely used Cox proportional hazard model to conduct the survival analysis of human cancers. We evaluate the performance of different survival analysis models based on the metrics of Concordance Index(CI) and AUC [31]. Both CI and AUC values range from 0 to 1, and the larger values indicate better prognosis performance.

3.4. Time Complexity Analysis

Suppose we have n nodes for each node type in a graph, the complexity of TMEGL is dominated by three components. Firstly, the complexity for calculating the node similarity within one node type is $O(n^2 d)$, and d refers to the dimensions of the embedding. Secondly, the complexity of applying the spectral clustering algorithm to aggregate the nodes into g groups is $O(n^3)$. Finally, we calculate the inter-group and inter-group similarities with the complexity of $O(g^2)$. In summary, the time complexity of our TMEGL is $O(n^2 d + n^3 + g^2)$. Although the complexity for spectral clustering is relatively high, we set n as 300 in this study that is still effective for computation.

4. Experimental Results

4.1. Dataset

In this study, we conducted experiments on three cancer cohorts *i.e.,* Breast Invasive Carcinoma (BRCA), Kidney Renal Clear Cell Carcinoma (KIRC) and Lung Squamous Cell Carcinoma (LUSC) derived from The Cancer Genome Atlas (TCGA). The demographic information of these three cancer cohorts are listed in Table 1:

Table 1. Demographics and clinical information of the dataset.

	BRCA	KIRC	LUSC
Number of Patients:	690	511	472
Censored Non-censored	627/63	340/171	270/202
Age (Mean)	54.6	66.17	59.20
Follow-up Time (Month)	25.3	31.8	36.1

4.2. Experimental Settings

For each cancer cohort, the 5-fold cross validation strategy is applied to evaluate the performance of different methods. For the training dataset, we randomly selected 25% samples as validation set for parameter tuning. For the graph embedding algorithm introduced in Section 3.2, we applied the pre-trained ResNet-101 to extract node features at first, and then feed them into the embedding network to get the node embeddings with the dimension of 128. The embedding network is consisted of two fully connected layers, whose dimensions are 512 and 128, respectively. For each type of node, the spectral clustering algorithm is applied to aggregate them into $B = 4$ groups, and the regularization parameter λ applied to balance the intra-group and inter-group similarities (shown in Eq. (7)) is tuned from 1 to $1e-5$. The Gated Graph Attention Networks introduced in Section 3.3 is consisted of 3 GAT layers, 3 Gated Graph Convolution (GGC) layers and 2 graph pooling layers. The dimensions for all the three GAT layers are set as 128, and the number of recurrent step t in each GGC layer is tuned from 1 to 6. Moreover, the percentage of the preserved nodes in each SAGPool layer is tuned from 0.3 to 1, Finally, we empirically set the number of epoch as 200, and the learning rate is tuned from the range of 1e-8 to 1e-2.

4.3. Comparison of TMEGL with Other Survival Analysis Models

We first compared the performance of TMEGL with the following WSI-based survival analysis models by the measurements of CI and AUC introduced in Section 3.4. 1) WSISA [43]: An ensemble model to make WSI-level predictions based on the combination of individual patch-level deep survival analysis models. 2) DeepAttnMISL [40]: A deep multiple instance learning model to directly learn survival patterns from gigapixel WSIs. 3) DeepGraphSurv [18]: A graph convolutional neural network with attention learning to extract the topological features from WSIs for survival prediction. 4) MCAT[5]: MCAT is the Transformer based method taking the position information into consideration for the survival prediction task based on WSIs. 5) Patch-GCN [4]: A context-aware graph convolutional network that hierarchically aggregates instance-level histology features to learn the global representations of WSIs. 6) CluSiam [36]:A self-supervised learning (SSL) based representation learning method for survival prediction from WSIs. 7) Co-Pilot[25]: A dynamic point-cloud based WSI representation methods for survival prediction of human cancers. 8) HGSurvNet [8]:A hyper-graph based learning framework for survival prediction from WSIs.

As can be observed from Table 2, Firstly, the prognostic power of the graph-based deep learning methods (*i.e.,* DeepGraphSurv, Patch-GCN, TMEGL,HGSurvNet) are generally superior to the multi-instance learning algorithms (*i.e.,* WSISA, DeepAttnMISL). This is because these GCN based methods can take the topological structure of WSI into consideration. Secondly, our proposed TMEGL can achieve the CI values of 0.719, 0.697 and 0.695, and the AUC values of 0.750, 0.716 and 0.736 on BRCA, KIRC and LUSC cohorts, respectively, which is not only significantly better than the comparing graph learning methods but also superior to the SOTA WSI-based representation learning algorithms(*i.e.,* CluSiam [36] and Co-Pilot[25]). This is because the proposed TMEGL can effectively take the topological organization of TME that are closely associated with cancer progression into consideration.

4.4. Comparison of TMEGL with Other Methods for Patient Stratification

For the survival analysis of human cancers, another important task is to stratify patients into different subgroups for personalized treatment. In this section, we followed the method introduced in [44], which applied the K-means clustering algorithm to aggregate the derived WSI-level features of different patients in the testing set into 2 groups. Then, the log-rank test [31] is applied to test if these two groups have significant difference of clinical outcome. Here, better prognosis prediction performance comes with the smaller p-value by the log-rank test, and we showed the results of different methods in Fig.3. As can be seen from Fig.3, our TMEGL can achieve the p-values of $3.32e-2, 2.01e-2$ and $3.01e-3$ on BRCA, KIRC and LUSC datasets, respectively, which are superior to the comparing HGSurvNet [8] and Co-Pilot[25] methods, and these results again validate the advantages of our TMEGL method.

4.5. The Difference of the Interactions Between High and Low survival Risk Patients.

For the stratified patients on different survival groups indicated in Fig.3, we also compared the proportions of the identified survival-associated edges connecting different types of tissues via graph pooling. As shown in Fig.4(a), most of the survival-associated edges on BRCA cohort are relevant to the interactions between lymphocytes and tumor (L-T) in WSIs, and it is obvious that the proportion of edges that connect lymphocytes and tumor regions in low-risk group are higher than that in high-risk group, showing that the immune system could control the tumor growth

Table 2. Comparisons of TMEGL with Other Models by the Mesearments of CI and AUC (along with their standard deviations).

	BRCA		KIRC		LUSC	
	CI	AUC	CI	AUC	CI	AUC
WSISA [43]	0.621(0.02)	0.590(0.05)	0.578(0.04)	0.607(0.05)	0.577(0.04)	0.587(0.05)
DeepAttnMISL [40]	0.617(0.05)	0.637(0.04)	0.629(0.05)	0.654(0.03)	0.654(0.04)	0.677(0.03)
MCAT[5]	0.638(0.05)	0.651(0.04)	0.631(0.05)	0.659(0.05)	0.643(0.05)	0.679(0.05)
DeepGraphSurv[18]	0.653(0.04)	0.665(0.06)	0.649(0.04)	0.671(0.04)	0.651(0.04)	0.683(0.03)
Patch-GCN [4]	0.664(0.04)	0.671(0.03)	0.662(0.03)	0.686(0.03)	0.648(0.05)	0.687(0.04)
CluSiam [36]	0.644(0.07)	0.639(0.05)	0.667(0.04)	0.691(0.02)	0.651(0.04)	0.671(0.03)
Co-Pilot[25]	0.657(0.03)	0.672(0.03)	0.639(0.05)	0.661(0.03)	0.674(0.03)	0.691(0.03)
HGSurvNet [8]	0.671(0.04)	0.681(0.05)	0.659(0.04)	0.681(0.04)	0.667(0.04)	0.701(0.03)
TMEGL	**0.719(0.03)**	**0.750(0.03)**	**0.697(0.03)**	**0.716(0.03)**	**0.695(0.04)**	**0.736(0.04)**

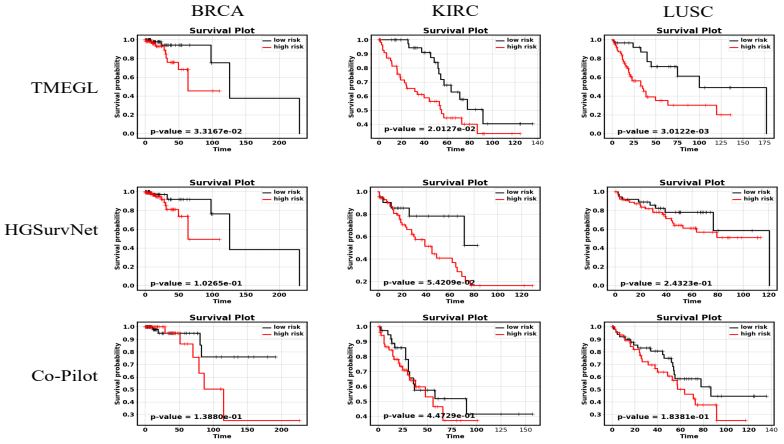

Figure 3. Comparison of the Stratification Performance Among Different Methods.

and reduce the survival risks. Moreover, we found that the proportions of the Stromal-Tumor (S-T) interactions on high-survival risk group is more prominent than that in low-survival risk group. Such observations is consistent with the existing knowledge that the rich interactions between tumor and stromal fibrosis will provide more chance to facilitate the invasion and metasis of tumors, and thus will lead to high survival risk [2].

In addition, we also visualized the sampled tumor, lymphocytes and stromal patches and their connected edges in Fig. 4(b). Here, the thick black lines refer to the edges with higher weights calculated via GAT. As can be seen from Fig.4(b), the lymphocytes density in low-survival risk patient (long survival time) is higher than that in high-survival risk patient (short survival time) while the high-risk patient is enriched with more tumor regions. These observations indicated that the lymphocytes and tumor density are important prognostic factor affecting the patients' survival. In addition, as shown in Fig.4(b), the weights and the number of the edges connecting tumor and TIL regions are higher for patients in low-survival risk group, while the weights and number of the tumor-stromal interactions are more prominent in high-survival risk group. These observations are consistent with the results shown in Fig.4(a), which again validate the good explainable ability of the proposed TMEGL for investigating the association among different TME components on survival prediction task.

4.6. Ablation Study

To further evaluate the effectiveness of TMEGL, we compared TMEGL with the their variants based on average Concordance Index(CI) value within the five fold cross-validation. 1) TMEGL-NE: Directly adopted the Gated Graph Attention Networks for survival prediction without the graph embedding step, and we applied the PCA algorithm to reduce the dimension of the node representation to 128 in order to conduct fair comparison with TMEGL. 2) TMEGL-DW: Applied the Deepwalk [27] to learn node embeddings for survival prediction. 3) TMEGL-GAT: Directly applied the GAT to capture the interactions among different

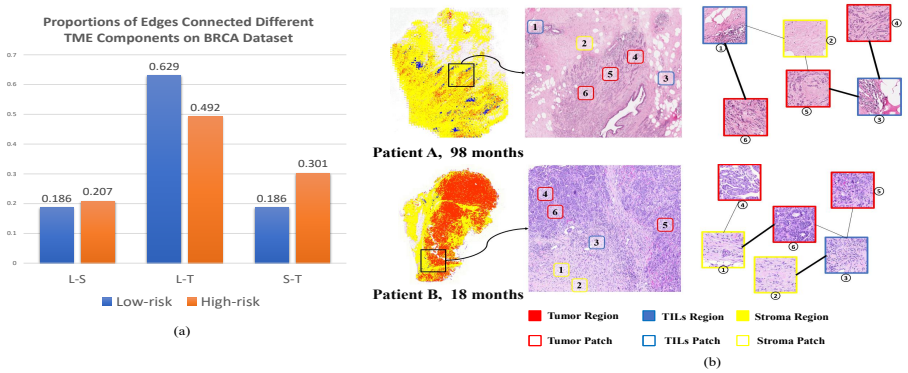

Figure 4. (a):The proportions of edges connected different types of TME components. (b): The interactions among different types of patches. The thick black line indicates the edges with higher weights calculated via graph Attention Network.

nodes without considering the edge is consisted of the same type of nodes or not. 4) TMEGL-N1 and TMEGL-N2: Considering the neighborhood organization within 1-hop and 2-hop ranges for designing the graph-embedding algorithm, respevtively. 5) TMEGL-G2 and TMEGL-G5: Aggregated each type of nodes into $B = 2$ and $B = 5$ groups to calculate the intra-group and inter-group similarity information for graph embedding learning. 6) TMEGL-Intra and TMEGL-Inter: Only applying the intra-group or inter-group similarity information to learn node embeddings. The experimental results are shown in Table 3.

As shown in Table 3, TMEGL is significantly better than TMEGL-NE and TMEGL-DW, suggesting the advantage of the proposed TME neighborhood organization guided graph embedding algorithm. In addition, TMEGL is also superior to TMEGL-GAT, indicating that the prognosis results can be further improved if we take the edge that is consisted of same type or different types of nodes into consideration. Moreover, TMEGL has stronger prognostic power than TMEGL-N1 and TMEGL-N2 since it considers wider hop range to describe the TME organization. Next, we can clear see that the prognosis results of TMEGL will be significantly decreased if the group number is relatively small (TMEGL-G2) for learning the graph embeddings. This is because the small group number may not have adequate ability to reflect the intra-group and inter-group similarities information extracted from the graph. Meanwhile, our results will be stable (TMEGL-G5) if the group number is increased. Finally, TMEGL also yielded better results than TMEGL-Intra and TMEGL-Inter, this implied that the incorporation of both intra-group and inter-group similarity information to learn node embeddings will provide a good solution to improve the prognosis performance.

5. Conclusion

In this paper, we developed a novel WSI-based survival prediction model *i.e.,* TMEGL for the prognosis prediction of human cancers. To the best of our knowledge, TMEGL is the first prospective study taking the interactions between tumor and its TME components into consideration to predict patients' clinical outcome. The experimental results also verifies the effectiveness of the proposed method. In summary, TMEGL is a general WSI-based graph learning framework that can be used to predict cancer survival and explain the mechanism of tumor progression, which opens up new opportunity for personalized treatment. In future, we plan to apply several new machine learning paradigms *e.g,* active learning, semi-supervised learning to train an accurate segmentation model with less annotated data for the pre-selection of important patches in WSIs.

Table 3. Comparison of TMEGL with its variants based on the measurement of CI.

	BRCA	KIRC	LUSC
TMEGL	**0.719**	**0.697**	**0.695**
TMEGL-NE	0.609	0.629	0.618
TMEGL-DW	0.658	0.631	0.621
TMEGL-GAT	0.662	0.630	0.633
TMEGL-N1	0.644	0.641	0.634
TMEGL-N2	0.671	0.663	0.657
TMEGL-G2	0.643	0.633	0.641
TMEGL-G5	0.694	0.680	0.681
TMEGL-Intra	0.676	0.636	0.641
TMEGL-Inter	0.668	0.623	0.646

6. Acknowledgement

This work was supported by the National Natural Science Foundation of China (Nos.62136004,62272226), by the National Key R&D Program of China (Grant No. 2023YFF1204803), and also by the Key Research and Development Plan of Jiangsu Province (No. BE2022842)

References

[1] Mohamed Amgad, Habiba Elfandy, Hagar Hussein, Lamees A Atteya, Mai AT Elsebaie, Lamia S Abo Elnasr, Rokia A Sakr, Hazem SE Salem, Ahmed F Ismail, Anas M Saad, et al. Structured crowdsourcing enables convolutional segmentation of histology images. *Bioinformatics*, 35(18): 3461–3467, 2019. 1, 3

[2] Manuela Budoni, Alessandra Fierabracci, Rosa Luciano, Stefania Petrini, Vincenzo Di Ciommo, and Maurizio Muraca. The immunosuppressive effect of mesenchymal stromal cells on b lymphocytes is mediated by membrane vesicles. *Cell Transplantation*, 22(2):369–379, 2013. 3, 7

[3] Fengju Chen, Darshan S Chandrashekar, Sooryanarayana Varambally, and Chad J Creighton. Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers. *Nature communications*, 10(1):5679, 2019. 1

[4] Richard J Chen, Ming Y Lu, Muhammad Shaban, Chengkuan Chen, Tiffany Y Chen, Drew FK Williamson, and Faisal Mahmood. Whole slide images are 2d point clouds: Context-aware survival prediction using patch-based graph convolutional networks. In *In Proceedings of Medical Image Computing and Computer Assisted Intervention*, pages 339–349. Springer, 2021. 6, 7

[5] Richard J Chen, Ming Y Lu, Wei-Hung Weng, Tiffany Y Chen, Drew FK Williamson, Trevor Manz, Maha Shady, and Faisal Mahmood. Multimodal co-attention transformer for survival prediction in gigapixel whole slide images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4015–4025, 2021. 2, 6, 7

[6] Jun Cheng, Xiaokui Mo, Xusheng Wang, Anil Parwani, Qianjin Feng, and Kun Huang. Identification of topological features in renal tumor microenvironment associated with patient survival. *Bioinformatics*, 34(6):1024–1030, 2018. 1, 2

[7] Jun Cheng, Zhi Han, Rohit Mehra, Wei Shao, Michael Cheng, Qianjin Feng, Dong Ni, Kun Huang, Liang Cheng, and Jie Zhang. Computational analysis of pathological images enables a better diagnosis of tfe3 xp11. 2 translocation renal cell carcinoma. *Nature Communications*, 11(1):1778, 2020. 2

[8] Donglin Di, Changqing Zou, Yifan Feng, Haiyan Zhou, Rongrong Ji, Qionghai Dai, and Yue Gao. Generating hypergraph-based high-order representations of whole-slide histopathological images for survival prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5):5800–5815, 2022. 2, 5, 6, 7

[9] Kevin W Eliceiri. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, pages 14318–14328, 2021. 2

[10] Henrik Failmezger. Topological tumor graphs: a graph-based spatial model to infer stromal recruitment for immunosuppression in melanoma histology. *Cancer research*, 80(5): 1199–1209, 2020. 2

[11] Thomas F Gajewski, Hans Schreiber, and Yang-Xin Fu. Innate and adaptive immune cells in the tumor microenvironment. *Nature Immunology*, 14(10):1014–1022, 2013. 2, 3

[12] Tobias Hoch, Daniel Schulz, Nils Eling, Julia Martínez Gómez, Mitchell P Levesque, and Bernd Bodenmiller. Multiplexed imaging mass cytometry of the chemokine milieus in melanoma characterizes features of the response to immunotherapy. *Science Immunology*, 7(70):eabk1692, 2022. 1, 5

[13] Le Hou, Dimitris Samaras, Tahsin M Kurc, Yi Gao, James E Davis, and Joel H Saltz. Patch-based convolutional neural network for whole slide tissue image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2424–2433, 2016. 1, 2

[14] Hanna Källén, Jesper Molin, Anders Heyden, Claes Lundström, and Kalle Åström. Towards grading gleason score using generically trained deep convolutional neural networks. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 1163–1167. IEEE, 2016. 2

[15] Sunkyu Kim, Keonwoo Kim, Junseok Choe, Inggeol Lee, and Jaewoo Kang. Improved survival analysis by learning shared genomic information from pan-cancer data. *Bioinformatics*, 36(Supplement_1):i389–i398, 2020. 1

[16] Nicole M Kuderer, Aakash Desai, Maryam B Lustberg, and Gary H Lyman. Mitigating acute chemotherapy-associated adverse events in patients with cancer. *Nature Reviews Clinical Oncology*, 19(11):681–697, 2022. 1

[17] Junhyun Lee, Inyeop Lee, and Jaewoo Kang. Self-attention graph pooling. In *International Conference on Machine Learning*, pages 3734–3743. PMLR, 2019. 5

[18] Ruoyu Li, Jiawen Yao, Xinliang Zhu, Yeqing Li, and Junzhou Huang. Graph cnn for survival analysis on whole slide pathological images. In *Proceedings of Medical Image Computing and Computer Assisted Intervention*, pages 174–182. Springer, 2018. 2, 3, 6, 7

[19] Yujia Li, Richard Zemel, Marc Brockschmidt, and Daniel Tarlow. Gated graph sequence neural networks. In *Proceedings of ICLR'16*, 2016. 5

[20] Tiancheng Lin, Hongteng Xu, Canqian Yang, and Yi Xu. Interventional multi-instance learning with deconfounded instance-level prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 1601–1609, 2022. 2

[21] Le Lu, Jinbo Bi, Matthias Wolf, and Marcos Salganicoff. Effective 3d object detection and regression using probabilistic segmentation features in ct images. In *CVPR 2011*, pages 1049–1056. IEEE, 2011. 1

[22] Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Richard J Chen, Matteo Barbieri, and Faisal Mahmood. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature Biomedical Engineering*, 5(6):555–570, 2021. 2

[23] Mufti Mahmud, Mohammed Shamim Kaiser, Amir Hussain, and Stefano Vassanelli. Applications of deep learning and reinforcement learning to biological data. *IEEE transactions on neural networks and learning systems*, 29(6):2063–2079, 2018. 1

[24] Caner Mercan, Selim Aksoy, Ezgi Mercan, Linda G Shapiro, Donald L Weaver, and Joann G Elmore. From patch-level to

roi-level deep feature representations for breast histopathology classification. In *Medical Imaging 2019: Digital Pathology*, pages 86–93. SPIE, 2019. 2

[25] Ramin Nakhli, Allen Zhang, Ali Mirabadi, Katherine Rich, Maryam Asadi, Blake Gilks, Hossein Farahani, and Ali Bashashati. Co-pilot: Dynamic top-down point cloud with conditional neighborhood aggregation for multi-gigapixel histopathology image representation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 21063–21073, 2023. 2, 6, 7

[26] Jackson Nyman, Thomas Denize, Ziad Bakouny, Chris Labaki, Breanna M Titchen, Kevin Bi, Surya Narayanan Hari, Jacob Rosenthal, Nicita Mehta, Bowen Jiang, et al. Spatially aware deep learning reveals tumor heterogeneity patterns that encode distinct kidney cancer states. *Cell Reports Medicine*, 4(9), 2023. 1

[27] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 701–710, 2014. 7

[28] Xuejun Qian, Jing Pei, Hui Zheng, Xinxin Xie, Lin Yan, Hao Zhang, Chunguang Han, Xiang Gao, Hanqi Zhang, Weiwei Zheng, et al. Prospective assessment of breast cancer risk from multimodal multiview ultrasound images via clinically applicable deep learning. *Nature Biomedical Engineering*, 5 (6):522–532, 2021. 1

[29] Erik Sahai, Igor Astsaturov, Edna Cukierman, David G DeNardo, Mikala Egeblad, Ronald M Evans, Douglas Fearon, Florian R Greten, Sunil R Hingorani, Tony Hunter, et al. A framework for advancing our understanding of cancer-associated fibroblasts. *Nature Reviews Cancer*, 20(3): 174–186, 2020. 2

[30] Joel Saltz, Rajarsi Gupta, Le Hou, Tahsin Kurc, Pankaj Singh, Vu Nguyen, Dimitris Samaras, Kenneth R Shroyer, Tianhao Zhao, Rebecca Batiste, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Reports*, 23 (1):181–193, 2018. 2, 3

[31] Wei Shao, Zhi Han, Jun Cheng, Liang Cheng, Tongxin Wang, Liang Sun, Zixiao Lu, Jie Zhang, Daoqiang Zhang, and Kun Huang. Integrative analysis of pathological images and multi-dimensional genomic data for early-stage cancer prognosis. *IEEE Transactions on Medical Imaging*, 39(1): 99–110, 2019. 5, 6

[32] Wei Shao, Tongxin Wang, Zhi Huang, Zhi Han, Jie Zhang, and Kun Huang. Weakly supervised deep ordinal cox model for survival prediction from whole-slide pathological images. *IEEE Transactions on Medical Imaging*, 40(12):3739–3747, 2021. 2, 3

[33] Chetan L Srinidhi, Ozan Ciga, and Anne L Martel. Deep neural network models for computational histopathology: A survey. *Medical Image Analysis*, 67:101813, 2021. 1

[34] Armando Stabile, Francesco Giganti, Andrew B Rosenkrantz, Samir S Taneja, Geert Villeirs, Inderbir S Gill, Clare Allen, Mark Emberton, Caroline M Moore, and Veeru Kasivisvanathan. Multiparametric mri for prostate cancer diagnosis: current status and future directions. *Nature reviews urology*, 17(1):41–61, 2020. 1

[35] Mei-Ling Wang, Wei Shao, Xiao-Ke Hao, and Dao-Qiang Zhang. Machine learning for brain imaging genomics methods: a review. *Machine intelligence research*, 20(1):57–78, 2023. 1

[36] Weiyi Wu, Chongyang Gao, Joseph DiPalma, Soroush Vosoughi, and Saeed Hassanpour. Improving representation learning for histopathologic images with cluster constraints. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 21404–21414, 2023. 2, 6, 7

[37] Yawen Wu, Michael Cheng, Shuo Huang, Zongxiang Pei, Yingli Zuo, Jianxin Liu, Kai Yang, Qi Zhu, Jie Zhang, Honghai Hong, et al. Recent advances of deep learning for computational histopathology: principles and applications. *Cancers*, 14(5):1199, 2022. 1

[38] Yawen Wu, Yingli Zuo, Qi Zhu, Jianpeng Sheng, Daoqiang Zhang, and Wei Shao. Transfer learning-assisted survival analysis of breast cancer relying on the spatial interaction between tumor-infiltrating lymphocytes and tumors. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 612–621. Springer, 2023. 2

[39] Jun Xu, Lei Xiang, Qingshan Liu, Hannah Gilmore, Jianzhong Wu, Jinghai Tang, and Anant Madabhushi. Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. *IEEE transactions on medical imaging*, 35(1):119–130, 2015. 1

[40] Jiawen Yao, Xinliang Zhu, Jitendra Jonnagaddala, Nicholas Hawkins, and Junzhou Huang. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. *Medical Image Analysis*, 65: 101789, 2020. 2, 6, 7

[41] Yu Zhao, Fan Yang, Yuqi Fang, Hailing Liu, Niyun Zhou, Jun Zhang, Jiarui Sun, Sen Yang, Bjoern Menze, Xinjuan Fan, et al. Predicting lymph node metastasis using histopathological images based on multiple instance learning with deep graph convolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4837–4846, 2020. 2

[42] Xinliang Zhu, Jiawen Yao, and Junzhou Huang. Deep convolutional neural network for survival analysis with pathological images. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 544–547. IEEE, 2016. 1, 2

[43] Xinliang Zhu, Jiawen Yao, Feiyun Zhu, and Junzhou Huang. Wsisa: Making survival prediction from whole slide histopathological images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7234–7242, 2017. 2, 6, 7

[44] Yingli Zuo, Yawen Wu, Zixiao Lu, Qi Zhu, Kun Huang, Daoqiang Zhang, and Wei Shao. Identify consistent imaging genomic biomarkers for characterizing the survival-associated interactions between tumor-infiltrating lymphocytes and tumors. In *Proceedings of Medical Image Computing and Computer Assisted Intervention–MICCAI 2022*, pages 222–231. Springer, 2022. 2, 3, 6