Dual-stream multi-dependency graph neural network enables precise cancer survival analysis

Zhikang Wang [a,b,c], Jiani Ma [d], Qian Gao [a], Chris Bain [e], Seiya Imoto [f], Pietro Liò [g], Hongmin Cai [h], Hao Chen [i], Jiangning Song [b,c,*]

[a] *Xiangya Hospital, Central South University, Changsha, China*
[b] *Biomedicine Discovery Institute and Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia*
[c] *Wenzhou Medical University-Monash Biomedicine Discovery Institute (BDI) Alliance in Clinical and Experimental Biomedicine, Wenzhou, China*
[d] *School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China*
[e] *Faculty of Information Technology, Monash University, Melbourne, Australia*
[f] *Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan*
[g] *Department of Computer Science and Technology, The University of Cambridge, Cambridge, United Kingdom*
[h] *School of Computer Science and Engineering, South China University of Technology, Guangzhou, China*
[i] *Department of Computer Science and Engineering and Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Hong Kong, China*

ARTICLE INFO

ABSTRACT

Histopathology image-based survival prediction aims to provide a precise assessment of cancer prognosis and can inform personalized treatment decision-making in order to improve patient outcomes. However, existing methods cannot automatically model the complex correlations between numerous morphologically diverse patches in each whole slide image (WSI), thereby preventing them from achieving a more profound understanding and inference of the patient status. To address this, here we propose a novel deep learning framework, termed dual-stream multi-dependency graph neural network (DM-GNN), to enable precise cancer patient survival analysis. Specifically, DM-GNN is structured with the feature updating and global analysis branches to better model each WSI as two graphs based on morphological affinity and global co-activating dependencies. As these two dependencies depict each WSI from distinct but complementary perspectives, the two designed branches of DM-GNN can jointly achieve the multi-view modeling of complex correlations between the patches. Moreover, DM-GNN is also capable of boosting the utilization of dependency information during graph construction by introducing the affinity-guided attention recalibration module as the readout function. This novel module offers increased robustness against feature perturbation, thereby ensuring more reliable and stable predictions. Extensive benchmarking experiments on five TCGA datasets demonstrate that DM-GNN outperforms other state-of-the-art methods and offers interpretable prediction insights based on the morphological depiction of high-attention patches. Overall, DM-GNN represents a powerful and auxiliary tool for personalized cancer prognosis from histopathology images and has great potential to assist clinicians in making personalized treatment decisions and improving patient outcomes.

1. Introduction

Cancer is a primary cause of mortality worldwide, with approximately 19.3 million new cancer cases and nearly 10.0 million cancer-related deaths reported in 2020 (Sung et al., 2021). Alarmingly, the global burden of cancer is expected to increase by 47% to 28.4 million cases by 2040 (Sung et al., 2021). Given the diverse range of cancer types and their impact on various organs and systems, precision medicine, particularly survival analysis that accounts for individual patient status and cancer progression, has the potential to significantly reduce cancer prevalence and address the challenges of tumor heterogeneity, thereby holding significant clinical and commercial value.

Generally, histopathology-based survival analysis relies on visually inspecting and quantifying histopathological alterations/features in cell morphology, invasiveness, or inflammation/infiltration in histopathology slides (Gurcan et al., 2009). However, this process becomes highly labor-intensive and time-consuming for pathologists due to the large

---

* Corresponding author.
  E-mail address: Jiangning.Song@monash.edu (J. Song).

gigapixel size of whole slide images (WSIs). Moreover, the final analysis suffers from pathologists' subjective experience and knowledge, rendering the prediction uncertain. In the past few years, the combination of advanced whole slide imaging techniques and deep learning technology has attracted significant research interest in histopathology WSI-based survival analysis, offering promising solutions to overcome these challenges. Compared with conventional methods, the computational pathology methods stand out in terms of efficiency, objectiveness, repeatability, and the possibility of remote diagnostics (Abels et al., 2019; Louis et al., 2016), which might provide new perspectives for alleviating the severe shortage and regional imbalance of qualified pathologists all over the world.

Histopathology images have a unique advantage over other medical modalities, e.g., radiology and ultrasonic imaging, in presenting a more detailed and high-resolution depiction of tumor cells and microenvironments (Kumar et al., 2014). Such intricate details enable a comprehensive characterization of cancer in evolution and progression. However, despite the advantages of WSIs in cancer analysis, their high gigapixel resolution poses considerable computational challenges that are not currently feasible with existing hardware for end-to-end processing. Additionally, due to the tumor heterogeneity for which cancerous tissues constitute only a small fraction of WSIs, annotating tumor cells or patches becomes exceedingly challenging. To address the aforementioned challenges, a promising solution is to utilize weakly supervised multiple-instance learning (MIL) (Maron and Lozano-Pérez, 1997). In this approach, each WSI is divided into numerous small patches (instances) and a pre-trained model or encoding strategy converts the patches into feature vectors. Subsequently, further analysis and prediction can be conducted on the computable feature space.

In recent years, a variety of MIL algorithms (Ilse et al., 2018; Li et al., 2021; Shao et al., 2021; Wang et al., 2023a; Yu et al., 2023; Gao et al., 2023; Wang et al., 2024) have been proposed for cancer diagnosis and subtype classification, some of which even achieved performance on par with human experts. The majority of these algorithms are developed to detect class-specific features from WSIs, e.g., tumor cells for cancer diagnosis and specific tumor morphological patterns for subtype classification. Nevertheless, the living status of cancer patients is determined by various factors, e.g., tumor stage (Dunnwald et al., 2007; Elston and Ellis, 1991), tumor size (Zhang et al., 2015; Narod, 2012), tumor-infiltrating lymphocytes (TIL) (Gooden et al., 2011; Deschoolmeester et al., 2010) and necrosis (Richardson et al., 2022). Accordingly, the detection-centric MIL methods are not competent for complicated survival analysis. Recently, there has been a proliferation of algorithms for survival analysis, encompassing convolutional neural network-based methods (Zhu et al., 2017), attention-based methods (Carmichael et al., 2022), graph-based methods (Li et al., 2018; Chen et al., 2021; Wang et al., 2021, 2022), and transformer-based methods (Huang et al., 2021; Wang et al., 2023b). These emerging methods aim to enhance the performance toward survival analysis by leveraging advanced technologies and network architectures. Specifically, attention-based methods can generate the attention weights for each patch and proceed with feature aggregation. Existing graph-based methods often struggle to establish complex correlations between patches, relying primarily on spatial contextual information or morphological correlation. In terms of Transformer-based methods, the patch correlations are determined by the learnable transformation and cross-product operations. This training process can be substantially influenced by the patch morphology similarities and the quantity of the available training data. As a consequence, despite the advancements in the above methods, we still face challenges in modeling the complex correlations between numerous morphologically diverse patches in WSIs, thereby failing to achieve a profound understanding of patient status.

To address the above issues, here, we propose a novel approach, termed dual-stream multi-dependency graph neural network (DM-GNN) for cancer patient survival analysis from histopathology images with only slide-level labels as weak supervision. DM-GNN is capable of automatically correlating numerous patches with diverse morphological patterns within each WSI, thereby enabling a more comprehensive analysis of cancer patients. Specifically, DM-GNN comprises the feature updating branch (FUB) and global analysis branch (GAB) to model each WSI as two graphs based on global morphological affinity and co-activating dependencies, respectively. The feature distance-based affinity matrix (AFM) and the attention generation module (AGM) estimate the co-activating matrix (CAM) function as the edges of the graphs, quantifying the dependencies into a computable form. Graph convolutional neural (Kipf and Welling, 2017) is employed to achieve the message-passing between nodes. As the two dependencies focus on distinct correlations between patches, we can achieve multi-view modeling of the complex correlations between patches. Furthermore, we enhance the utilization of dependency information in graph construction by introducing the affinity-guided attention recalibration (AARM) module as the readout function. This module increases robustness against feature perturbation and thus can enable more reliable and stable predictions.

To validate the performance of DM-GNN, benchmarking experiments are conducted on five different TCGA tumor datasets, including Bladder Urothelial Carcinoma, Breast Invasive Carcinoma, Glioblastoma & Lower Grade Glioma, Lung Adenocarcinoma, and Uterine Corpus Endometrial Carcinoma datasets. Experimental results demonstrate that the proposed DM-GNN can accurately model the risk function of the population and outperform other state-of-the-art methods. Moreover, we provide the morphological depiction of high-attention patches to interpret the prediction insights of our trained model.

2. Related work

Existing computational methods for histopathology image-based survival analysis can be generally divided into two broad categories: regions of interest (ROI)-based methods and WSI-based methods. These two major categories of methods are described in detail in the following subsections.

2.1. ROI-based methods

ROI-based methods rely on the manually selected patches of the ROI for survival prediction. In the early stage, these methods typically utilize hand-crafted features (Barker et al., 2016; Cheng et al., 2018; Wang et al., 2014; Yao et al., 2016; Yu et al., 2016; Yuan et al., 2012) to represent the images, by especially focusing on the cellular morphology and patch-level characteristics. In this regard, Wang et al. (2014) proposed an integrated framework with cell detection, segmentation, and classification for cancer diagnosis and survival analysis. Yao et al. (2016) employed a deep cell subtype classification to annotate the cells and introduced a set of quantitative features to describe cellular information for survival analysis. Considering that the progression of tumors is not only related to the growth of tumor cells but also the interaction between the tumor and its microenvironment, Cheng et al. (2018) further proposed a novel bioimage informatics pipeline for automatic characterization of the topological organization of different cell patterns in the tumor microenvironment. Yu et al. (2016) extracted 9879 quantitative image features and used regularized machine-learning methods to select the top features and to distinguish shorter-term survivors from longer-term survivors with stage I adenocarcinoma. The extensive validation verified the potential of computational pathology in medicine. In recent years, advances in deep learning and big data techniques have enabled the development of deep learning-based survival analysis methods (Zhu et al., 2016; Mobadersany et al., 2018). For instance, Zhu et al. (2016) for the first time developed an early-stage deep convolutional neural network (DeepConvSurv) with pathological images for survival analysis. With the novel optimization strategy for the regression model, DeepConvSurv achieved better performance than
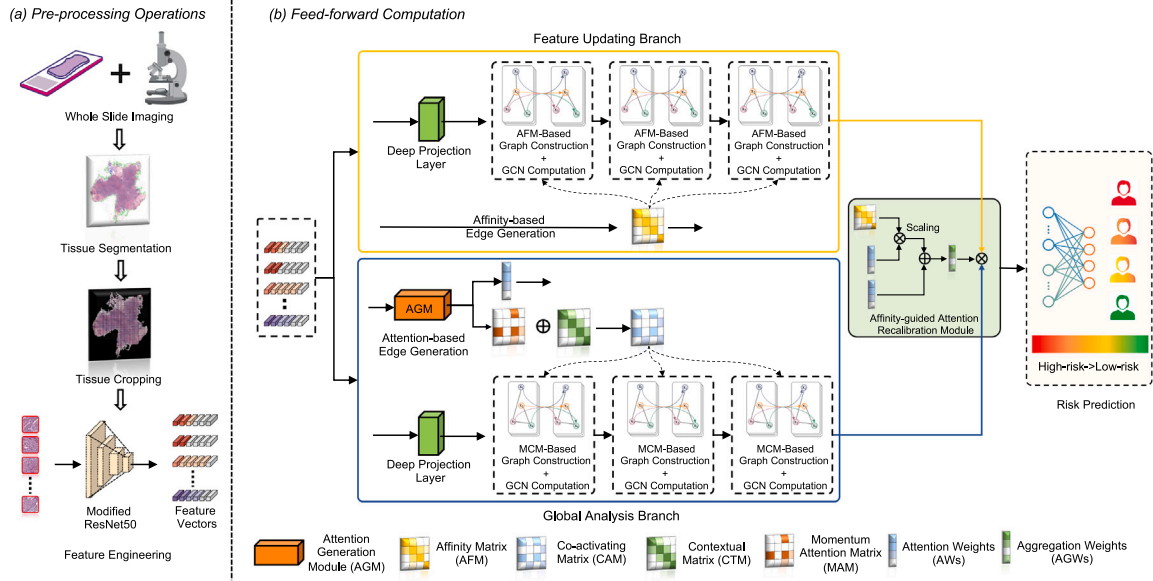
Fig. 1. Overview of the proposed dual-stream multi-dependency graph neural network (DM-GNN). The framework comprises pre-processing operations to convert WSIs into bags of features and feed-forward computation to predict hazard rates. Regarding the network, it is constructed by the feature updating branch and global analysis branch for representation generation and affinity-guided attention recalibration module for graph-level feature aggregation. Eventually, one linear layer will predict the risk of each patient.

machine learning-based methods. Later, Mobadersany et al. (2018) proposed survival convolutional neural networks (SCNNs) for survival analysis. Particularly, SCNNs are able to integrate both histopathology images and genomic biomarkers for final prediction.

Although ROI-based methods have achieved remarkable performance, they still necessitate the involvement of pathologists for patch selection. The pathologists' subjective assessment of the histopathology images can inevitably impact the patch labeling and hence extend to the final prediction. Moreover, patch selection is a labor-intensive task and requires a significant amount of time and effort.

2.2. WSI-based methods

WSI-based methods stand out in comparison to ROI-based methods in processing the original WSIs without any pre-processing. Considering the large giga-pixel sizes of WSIs and the computational power of current hardware, the majority of existing methods process WSIs through weakly-supervised multiple instance learning (MIL). Specifically, weakly supervised methods first crop the tissue regions of WSIs into patches and utilize a pre-trained model to convert patches into feature vectors. Then, MIL algorithms model the feature vectors for the final prediction. Most methods focus on the second part of development. Methods based on attention mechanisms, convolutional neural networks, transformers, and graph neural networks have been proposed. Zhu et al. (2017) proposed the WSISA method, which conducted survival prediction using whole slide images in an end-to-end manner. WISIA first extracted hundreds of patches from each WSI by adaptive sampling and then grouped them into different clusters. Then, an aggregation model was trained to perform patient-level predictions. Chen et al. (2021) proposed the Patch-GCN, which extensively explored the contextual features (tumor microenvironment) with GCN to generate discriminative representations. Wang et al. (2021) proposed a deep learning framework that leverages hierarchical graph-based representations (both patch-based and cell-based pathomic features) to enable more precise prediction of progression-free survival in prostate cancer patients. Huang et al. (2021) integrated the transformer network for long-range dependency establishing and critical feature learning. Moreover, they verified the effectiveness of self-supervised learning for cancer survival analysis. Di et al. (2022) introduced a multi-hypergraph-based learning framework (HGSurvNet) to explore

an informative survival-specific global representation from those WSIs with highly complicated data correlation. Chen et al. (2022) introduced a new vision transformer architecture termed Hierarchical Image Pyramid Transformer (HIPT) for WSI representation learning. The self-supervised learning significantly updated the extracted features and the hierarchical structure ensures a multi-scale modeling of the WSIs.

Although numerous methods have been proposed for the WSI-based cancer prognosis, most of these methods cannot explicitly model the complex correlations/dependencies between numerous morphology-diverse patches, thereby preventing them from achieving a more profound understanding and inference of the patients' survival status.

3. Methodology

In this section, we introduce a dual-stream multi-dependency graph neural network (DM-GNN), which is a new weakly supervised learning approach for cancer survival analysis from histopathology images. An overview of the framework of DM-GNN and a detailed description of the key branches and modules are provided below.

3.1. Overview of the framework

As shown in Fig. 1, the framework of DM-GNN comprises of the pre-processing operations and feed-forward computation. Following these steps, each WSI can be transformed from the original high-dimensional image space to a lower-dimensional feature space and subsequently, follow-up survival analysis can be computationally efficient.

Pre-processing Operations: Initially, we employ an automated segmentation algorithm to differentiate the foreground (tissue region) and background of each WSI, and then crop the tissue region into numerous fixed-size patches. Subsequently, we conduct deep transfer feature engineering by utilizing an ImageNet pre-trained model to embed each image patch into a feature vector. In terms of the pre-trained model, aligned with Patch-GCN (Chen et al., 2021), it is constructed by the initial *Convolution Block* and first three *Residual Blocks* of the ResNet50 model (He et al., 2016). Consequently, each patch is represented as a 1024-dimensional feature vector, and the bag of all instances can be denoted as $F = \{f_1, f_2, \ldots, f_N\} \in \mathbb{R}^{N \times 1024}$. After the feature embedding, the model training and testing can be performed in the low-dimensional feature space, thus substantially reducing the computational burdens.

Feed-forward Computation: In this step, the extracted bag of features F is fed into the proposed DM-GNN to perform survival prediction. As illustrated in Fig. 1(b), the system comprises the feature updating branch (FUB) and global analysis branch (GAB), both of which depict the transformed WSI as two graphs based on the theoretically distinct dependencies. Consequently, the features extracted by the two branches pay attention to the WSIs from different perspectives, thus enabling the multi-view modeling of the complex correlations. Afterward, the parameter-free transitivity-based aggregation module fully utilizes the dependency information during graph construction to generate the final graph representations that are more robust against perturbation. Eventually, a classification layer at the end of the model is employed to predict the hazard rate of each patient.

3.2. Graph convolutional network

Graph convolutional network (GCN) (Kipf and Welling, 2017) has an excellent capability of modeling topological graphs through a sophisticated message-passing mechanism. This mechanism essentially relies on the principle of updating a node's representation by systematically integrating the information extracted from its immediate neighboring nodes, which is akin to creating a dynamic conversation between nodes. The node-updating strategy forms an essential part of the way GCN manipulates and interprets graph data, enabling the GCN to better recognize and leverage the intricate relationships embedded within the graph structure.

In this study, GCN is utilized to model the interdependencies across different patches in each WSI. Its calculation process can be represented as $GCN(X, A)$, where X and A denote the nodes' features and undirected dependency matrix, respectively. Specifically, the function of the graph computation can be formulated as follows:

$$GCN(X, A) = \delta(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} XW), \tag{1}$$

where $\tilde{D}_{ii} = \sum_j A_{ij}$ and W is the trainable weight matrix for feature transformation.

3.3. Feature updating branch

Recent methods (Di et al., 2022; Chen et al., 2022) have specifically involved the spatial contextual/proximity information during the modeling. However, this strategy focuses more on the microenvironment itself rather than the patch-level feature distribution. The feature updating branch (FUB) is used to establish the correlations between the morphologically similar patches from the global perspective. In this way, the updated patch-level features can show better robustness against feature perturbation caused by deep transfer feature engineering in the pre-processing step and intra-tumor heterogeneity. Specifically, we achieve this through the introduced affinity matrix (AFM) which acts as edges for graph construction and GCN for patch-level message passing. The edge generation process is described in detail below.

During the affinity-based edge generation process, we initially calculate the pairwise cosine distances of the patches and then generate a distance matrix, denoted as $dist$. As the ImageNet pre-trained model was utilized to transform the patches into the feature vectors, the similarities between feature vectors tend to be high, which may not reflect the distinct morphological characteristics of the histopathological patches. Therefore, instead of directly adopting the generated $dist$ as the edges for the graph, we implement the min–max normalization to normalize the matrix value to a range between 0 and 1, ensuring the relative differences in morphological characteristics among the patches. This normalization function can be formulated as:

$$dist' = \frac{dist_{i,j} - dist_{min}}{dist_{max} - dist_{min}}. \tag{2}$$

Next, we use the pre-defined threshold to binarize the $dist'$ matrix and generate the final AFM. With this operation, we can regulate the flux

and magnitude of the message passing among nodes, thereby improving the controllability of graph construction.

Prior to graph computation, a deep projection layer (DPL), which comprises two linear layers with the $ReLU$ activation function and the layer normalization (LN) (Ba et al., 2016) in the middle, is used to reduce the feature dimension and recalibrate the feature for histopathology. Following this, the AFM-based graph will be fed to three consecutive GCN layers for feature updating. Each GCN layer is followed by a $ReLU$ and a LN for feature non-linearization and normalization.

3.4. Global analysis branch

The global analysis branch (GAB) focuses on capturing the correlations between informative patches that contribute to the survival analysis from the global perspective. Different from previous methods that directly learn the patch-patch correlation (Shao et al., 2021; Lu et al., 2021), the proposed GAB aims to explore the patch-level attention weights first and then uses the generated weights to establish the potential communication. The two-stage modeling strategy could map more detailed correlations between the numerous tissue patches, thereby improving the prediction performance. For this purpose, the generated co-activating matrix (CAM), which comprises the momentum attention matrix (MAM) and contextual matrix (CTM), aims to explore the short-range (adjacent patches) and long-range (co-activation from the global view) dependencies to model the WSIs. Identical to the FUB, GCN is employed for message passing between patches.

As illustrated in Fig. 1, the CAM, built from the momentum attention matrix (MAM) and the contextual matrix (CTM), functions as the dependency for constructing the graph. Specifically, MAM is generated by the proposed AGM that is based on the gated attention mechanism (Ilse et al., 2018). Firstly, the bag of features F is fed into two parallel linear layers for feature transformation, each of which is followed by the $sigmoid$ or $tanh$ activation functions, respectively. Afterward, the two transformed features are multiplied. The product results are then fed into another linear layer to produce 1-dimensional attention weights (represented as AWs $\in \mathbb{R}^{N \times 1}$) for instances. This computational process can be formulated as follows:

$$AWs = (sigm(W_1 F + b_1) \cdot tanh(W_2 F + b_2))W_3 + b_3, \tag{3}$$

where W and b represent the learnable transformation matrix and bias, respectively, while $sigm$ indicates the $sigmoid$ function. The AWs have two functions: (1) generate the attention matrix for graph construction; (2) serve as the input to AARM (Section 3.5) for node feature aggregation. Regarding the graph construction, each element in the AWs represents the activation level of the corresponding instance and as such, the product value between any two elements indirectly indicates the co-activating status of them. Hence, we initially employ a $sigmoid$ function to range the AWs to the scale of [0,1] and then proceed with these scaled ones using a self-cross-product operation to generate the attention matrix (ATM). Due to the limitations of attention modules in generating reliable attention weights in the initial training phase, the generated attention matrix may overlook critical correlations and affect the model optimization. To address this, we introduce a diminishing momentum to assist the model optimization. Overall, the function can be formulated as follows:

$$MAM = \alpha + (1 - \alpha) \cdot sigm(AWs) \times sigm(AWs)^T, \tag{4}$$

where α is a pre-defined hyper-parameter and will decrease to 0 as the training proceeds. Recent studies have indicated that the spatial contextual information that reflects the tumor microenvironment also plays an important role in cancer development and evolution (Yuan, 2016). Therefore, to better explore such information, we generate the 2-hop contextual matrix (CTM) based on patch-level spatial correlations. The summation of CTM and MAM will then be used as the CAM.

Regarding graph computation, it parallels the FUB, employing a DPL initially for feature transformation, followed by another three GCN layers for feature updating. Each GCN layer is subsequently followed by a $ReLU$ activation function and a layer normalization (LN) (Ba et al., 2016).
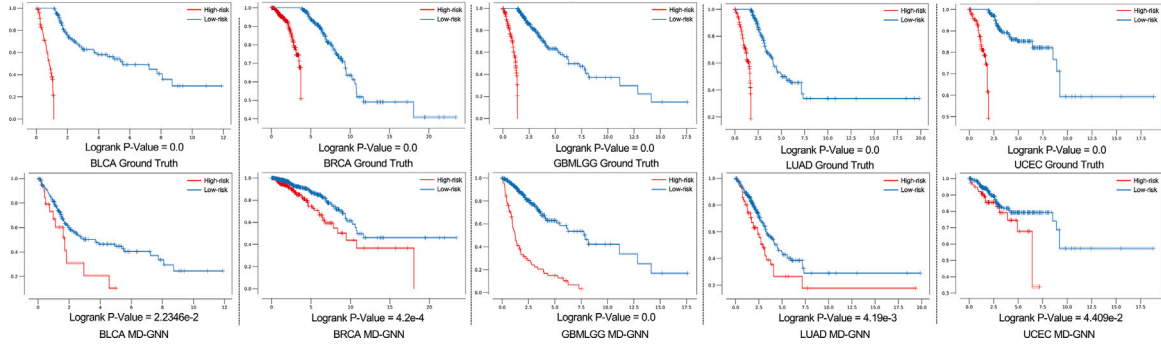
Fig. 2. Kaplan–Meier survival curves of our proposed DM-GNN and ground truth across five cancer types. High-risk and low-risk patients are represented by red and blue lines, respectively. The x-axis shows the time in months and the y-axis presents the probability of survival. The log-rank test is used to evaluate the statistical significance in survival distributions between low-risk and high-risk patients (*P-Value* < 0.05).

3.5. Affinity-guided attention recalibration module

To achieve the final hazard risk prediction of each patient, we still need a readout function to aggregate the node-level features as the graph-level representation. Previous methods typically utilized simple aggregation techniques, e.g., mean or maximum pooling, or relied on generated attention weights for representation aggregation. However, these approaches have shown limitations, either introducing excessive noise into the final representation or exhibiting instability during the training process. In this subsection, we propose a new affinity-guided attention recalibration module (AARM) to stabilize and recalibrate the attention weights of each patch by referring to the morphologically similar ones.

As illustrated in Fig. 1, AARM is featured by four different inputs, including AWs generated by AGM, AFM, as well as features generated by FUB and GAB, respectively. Specifically, the module starts with conducting a cross-product between AWs and AFM. In this way, the weight of each node/patch will be the collection of all the morphologically similar ones, thereby exhibiting better robustness against feature perturbation. Afterward, we scale the collected weights by dividing them by $\sqrt{d_k}$, where d_k is the total number of instances in this bag. Eventually, we operate the softmax operation on the weighted summation of the AWs and scaled collected weights for generating the final aggregation weights (AGWs). We generate the final representation of patients by proceeding with the cross-product operation between the AGWs and concatenated features from the two branches. This process can be formulated as follows:

$$F' = softmax(\beta \frac{AFM \times AWs}{\sqrt{d_k}} + (1-\beta)AWs) \times [F_1, F_2], \quad (5)$$

where β is the pre-defined hyperparameter, $F_1 = FUB(F)$ and $F_2 = GAB(F)$ denote the features from the two branches, and $[,]$ represents the concatenation operation.

3.6. Loss function

In this study, we apply the cross-entropy-based Cox proportional loss (Zadeh and Schmid, 2020) to optimize the model. Generally, the datasets for survival prediction have both censored and uncensored patients. For the model training and testing, we first convert the continuous overall survival time into four non-overlapping bins: $[t_0, t_1)$, $[t_1, t_2)$, $[t_2, t_3)$, and $[t_3, t_4)$, where $t_0 = 0$, $t_4 = \infty$, and t_1, t_2, t_3 correspond to the quartiles of longest survival time of uncensored patients. This binning process serves to discretize the survival time and allows for easier prediction analysis. For a patient j with t_j, we obtain his/her discretized class label y_j by referring to the above bins. The final loss function can be formulated as follows:

$$L_{cox} = -c_j \cdot log(f_{surv}(y_j, F_j))$$
$$- (1-c_j) \cdot log(f_{surv}(y_j - 1, F_j)) \quad (6)$$
$$- (1-c_j) \cdot log(f_{hazard}(y_j, F_j)),$$

Table 1
Data details of the BLCA, BRCA, GBMLGG, LUAD and UCEC datasets with CS, US, and AP representing censored samples, uncensored samples and average patches, respectively.

Number of	BLCA	BRCA	GBMLGG	LUAD	UCEC
Samples	436	1022	1041	515	538
CS	236	889	700	314	460
US	200	133	341	201	78
AP	15,014	9760	7495	10,973	16,142

where F and c_j indicates the bag the features and censorship of the j_{th} patient, respectively while $f_{harzard}$ represents the computation process of hazard rates. In terms of f_{surv}, it can be formulated as (see Fig. 2):

$$f_{surv}(y_j, F_j) = \prod_{i=0}^{y_j}(1 - f_{hazard}(i, F_j)). \quad (7)$$

4. Experimental results and discussion

4.1. Dataset description

In this study, we conducted extensive experiments on 3553 WSIs of five different tumor types obtained from The Cancer Genome Atlas (TCGA) including Bladder Urothelial Carcinoma (BLCA), Breast Invasive Carcinoma (BRCA), Glioblastoma & Lower Grade Glioma (GBMLGG), Lung Adenocarcinoma (LUAD), and Uterine Corpus Endometrial Carcinoma (UCEC) datasets. A detailed summary of the five datasets in terms of censored samples (CS), uncensored samples (UC), and total numbers of samples is provided in Table 1. Moreover, we also showcase the average number of patches of the five datasets.

4.2. Implementation details

All the WSIs were processed at the magnification level of 20 times (20x). In the pre-processing step, we initially discarded the background patches with a low saturation value (< 15). Subsequently, the tissue region was cropped into a series of 256×256 non-overlapping patches. Then, an ImageNet pre-trained modified ResNet50 model was used to embed original patches into a collection of feature vectors. The hyperparameter α is initially set as 0.3 and will decrease by 0.1 every 5 epochs. In the feed-forward computation, we applied the Adam optimizer to train each model for 20 epochs in an end-to-end fashion. In each iteration, the batch size was set as 1 and the gradients were accumulated over 32 steps for backpropagation. With respect to the super large size WSIs, we randomly sampled their 30,000 patches for modeling in light of the computation burden. All experiments were performed on a single NVIDIA GeForce RTX 3090 Graphics Card.

Table 2
Performance comparison with state-of-the-art methods on TCGA datasets in terms of c-index.

Methods	BLCA	BRCA	GBMLGG	LUAD	UCEC	Overall
MIL (Deep Set) (Zaheer et al., 2017)	0.500 ± 0.002	0.499 ± 0.001	0.573 ± 0.084	0.511 ± 0.019	0.504 ± 0.013	0.517
Attention MIL (Ilse et al., 2018)	0.528 ± 0.047	0.559 ± 0.048	0.772 ± 0.034	0.573 ± 0.073	0.589 ± 0.086	0.604
DeepGraphConv (Li et al., 2018)	0.580 ± 0.043	0.550 ± 0.099	0.785 ± 0.008	0.592 ± 0.045	0.631 ± 0.065	0.628
CLAM-MB (Lu et al., 2021)	0.512 ± 0.029	0.580 ± 0.087	0.736 ± 0.052	0.567 ± 0.038	0.571 ± 0.051	0.593
CLAM-SB (Lu et al., 2021)	0.528 ± 0.027	0.546 ± 0.070	0.745 ± 0.036	0.574 ± 0.050	0.586 ± 0.080	0.596
Patch-GCN (Chen et al., 2021)	0.562 ± 0.042	0.600 ± 0.035	0.806 ± 0.016	0.598 ± 0.053	0.646 ± 0.047	0.642
Patch-GCN+VarPool (Carmichael et al., 2022)	0.571 ± 0.023	0.607 ± 0.069	0.814 ± 0.021	0.605 ± 0.045	0.655 ± 0.040	0.650
TransMIL (Shao et al., 2021)	0.561 ± 0.049	0.641 ± 0.079	0.852 ± 0.020	0.645 ± 0.069	**0.669 ± 0.110**	0.674
CoADS (Zhao et al., 2023)	0.549 ± 0.066	0.631 ± 0.082	0.847 ± 0.032	0.618 ± 0.032	0.641 ± 0.029	0.657
GraphLSurv (Liu et al., 2023)	0.551 ± 0.013	0.586 ± 0.052	0.787 ± 0.021	0.585 ± 0.064	0.649 ± 0.043	0.632
HGT (Hou et al., 2023)	0.575 ± 0.092	0.639 ± 0.070	0.853 ± 0.039	0.642 ± 0.049	0.661 ± 0.052	0.674
Ours	**0.597 ± 0.045**	**0.651 ± 0.076**	**0.862 ± 0.019**	**0.664 ± 0.065**	0.663 ± 0.031	**0.687**

Table 3
Analysis of the number of local prototypes on TCGA datasets in terms of c-index.

Threshold	BLCA	BRCA	GBMLGG
0.75	0.575 ± 0.056	0.619 ± 0.067	0.836 ± 0.030
0.80	0.577 ± 0.047	0.634 ± 0.082	0.839 ± 0.031
0.85	0.582 ± 0.051	0.633 ± 0.078	**0.842 ± 0.024**
0.90	**0.583 ± 0.050**	**0.649 ± 0.068**	0.830 ± 0.035
0.95	0.537 ± 0.046	0.631 ± 0.080	0.835 ± 0.028

Threshold	LUAD	UCEC	Overall
0.75	0.609 ± 0.074	0.598 ± 0.093	0.647
0.80	0.616 ± 0.070	0.600 ± 0.089	0.653
0.85	0.617 ± 0.067	0.583 ± 0.101	0.651
0.90	**0.631 ± 0.070**	0.600 ± 0.109	**0.658**
0.95	0.627 ± 0.071	**0.605 ± 0.110**	0.647

In terms of the data, we randomly took non-overlapping 80% and 20% WSIs of each dataset as training and testing data, respectively. To assess the feasibility and stability of the proposed algorithm, we utilized the five-fold cross-validation for model evaluation. Furthermore, to ensure a fair comparison, all the compared weakly supervised deep learning methods followed the identical data splitting strategy. The concordance index (c-index), which can quantify the model's ability to correctly rank the survival outcomes of different patients, is employed in this study for performance evaluation.

4.3. Comparison with state-of-the-art methods

This subsection provides the performance comparison between our proposed DM-GNN with other weakly supervised learning methods for WSI-based cancer prognosis. Here, these baseline methods can be categorized as attention-based (Ilse et al., 2018; Lu et al., 2021), graph-based (Li et al., 2018; Chen et al., 2021; Carmichael et al., 2022; Liu et al., 2023; Hou et al., 2023), transformer-based (Shao et al., 2021) and their mixture (Zhao et al., 2023). As we can see from Table 2, DM-GNN outperformed all previous techniques on the four cancer types except on the UCEC dataset. The overall performance of DM-GNN reaches 0.687, surpassing the second-best methods by 1.3%. On the TCGA-UCEC dataset, TransMIL (Shao et al., 2021) outperformed our DM-GNN by 0.6%; however, DM-GNN achieved the surpassing on the other four datasets by at least 1%. Additionally, one significant limitation of TransMIL (Shao et al., 2021) lies in the dramatically increasing computation on the large-size WSIs. DeepGraphConv (Li et al., 2018) tried to enhance the patch-level features by aggregating similar patches from the global view. Nevertheless, using the PCA-selected features to model the graph could lead to unrepresentative edges. Additionally, the high-similarity nature of instance features is not considered in this process. Patch-GCN (Chen et al., 2021) particularly considered the microenvironment information by extensively exploring the spatial contextual information of WSIs; however, they ignored both the correlations within morphologically similar patches and the communication between informative ones, resulting in at least

4.5% worse performance than ours. In terms of CoADS (Zhao et al., 2023), GraphLSurv (Liu et al., 2023), and HGT (Hou et al., 2023), they additionally considered the long-range patch correlation while modeling the WSIs. They either use the learnable relationships or the patch characteristics. Based on the experiment results, we can easily conclude that our learning strategy utilizing two branches to model the morphological and latent correlations is superior to them for a more profound understanding of the patients. In conclusion, our proposed DM-GNN represents an advanced and effective approach for WSI-based cancer prognosis, holding significant potential as a powerful tool to improve our understanding of the interdependencies among patches, enhance the effectiveness of WSI-based survival prediction, and contribute to the fields of digital pathology and precision medicine.

4.4. Ablation study

Analysis of the threshold for the Affinity Matrix. As discussed in Section 3.3, we employed a pre-defined threshold to binarize the affinity matrix and established the correlations between patches with similar morphological characteristics. Here, to examine the effect of the thresholds on model performance, we specifically tested the varying thresholds ranging from 0.75 to 0.95 with a stepsize of 0.05. In this testing phase, we only utilized the global analysis branch to model the WSIs and mean operation as the readout function. As shown in Table 3, our model performed best on TCGA-BLCA, TCGA-BRCA, and TCGA-LUAD datasets with a threshold of 0.90, achieving c-index values of 0.583, 0.649, and 0.631, respectively. When tested on the TCGA-GBMLGG and TCGA-UCEC datasets, the model achieved the best performance under the thresholds of 0.85 and 0.95, respectively. The model underwent a substantial decline in performance with the threshold setting as 0.75. This might be attributed to the established correlations and message passing between dissimilar patches, which impeded node-level feature learning and consequently affected the discriminative power of the learned representations. On the other hand, when the threshold reached 0.95, there was also a performance decline across the four datasets. This suggests that strict restrictions might disrupt the learning of the potential correlations between patches, thus negatively impacting the message-passing process. Moreover, we observed that different kinds of tumors have distinct optimal thresholds, which are potentially reflective of the fact that each cancer type possesses unique morphological characteristics. All the experiments in Table 4 were conducted using the optimal threshold of each dataset as the default setting.

Ablation study of each proposed module. In this subsection, we conducted the ablation study to examine the effectiveness of the proposed branches and modules through benchmarking experiments on five TCGA datasets. The experimental results are provided in Table 4. $Model_0$ utilized the FUB for feature extraction and the thresholds for matrix binarization were referred to the optimal option of each dataset in Table 3. $Model_1$ and $Model_2$ constructed the graph of GAB based on ATM and its momentum one (MAM), respectively, for feature engineering. $Model_3$ is the combination of $Model_0$ and $Model_2$. In

Table 4

Ablation study on the TCGA-GBMLGG, TCGA-LUAD and TCGA-UCEC datasets. Specifically, we evaluate the significance of the proposed FUB, GAB, and AARM in terms of c-index with detail.

		BLCA	BRCA	GBMLGG	LUAD	UCEC	Overall
0	FUB + mean	0.583 ± 0.050	0.649 ± 0.068	0.842 ± 0.024	0.631 ± 0.070	0.605 ± 0.110	0.662
1	GAB (w/ ATM) + mean	0.568 ± 0.052	0.622 ± 0.058	0.841 ± 0.041	0.626 ± 0.088	0.617 ± 0.098	0.655
2	GAB (w/ MAM) + mean	0.571 ± 0.047	0.633 ± 0.070	0.852 ± 0.023	0.626 ± 0.064	0.629 ± 0.049	0.660
3	FUB + GAB (w/ MAM) + mean	0.576 ± 0.049	0.645 ± 0.078	0.855 ± 0.026	0.642 ± 0.060	0.661 ± 0.033	0.676
4	FUB + GAB (w/ MAM) + AARM	0.587 ± 0.056	0.649 ± 0.082	0.857 ± 0.025	0.661 ± 0.067	0.661 ± 0.058	0.683
5	FUB + GAB (w/ MCM) + AARM	0.597 ± 0.045	0.651 ± 0.076	0.862 ± 0.019	0.664 ± 0.065	0.663 ± 0.031	0.687
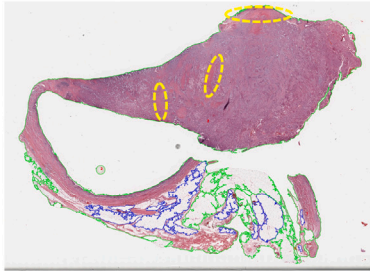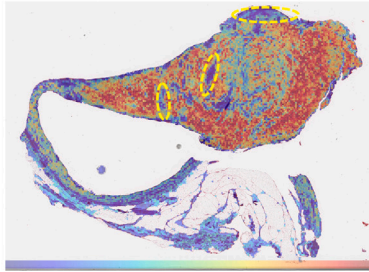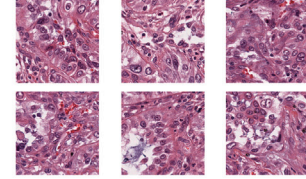
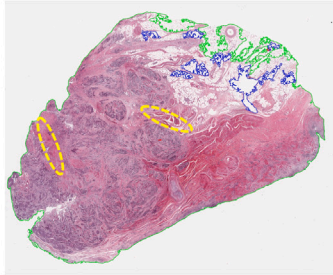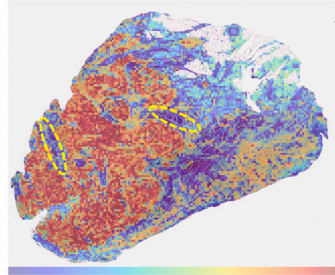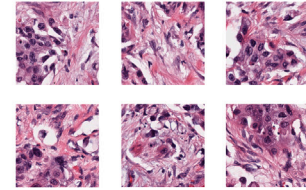**Fig. 3.** Attention visualization of DM-GNN on two WSIs in high-risk and low-risk cohorts from the TCGA-BLCA dataset. (a,d), (b,e) and (c,f) present the segmented WSIs, attention-mapped WSIs, and high-attention patches, respectively. Particularly, experienced pathologists depict the morphological features of the attention patches, revealing the prediction insights of the trained model.

terms of $Model_4$, we added the AARM module on top of $Model_3$ for feature aggregation. The final $Model_5$ represents our proposed DM-GNN model with spatial contextual information. In comparison with other state-of-the-art methods in Table 2, $Model_0$ and $Model_2$ showed superior overall performance than DeepGraphConv (Li et al., 2018), Patch-GCN (Chen et al., 2021), CoADS (Zhao et al., 2023), and GraphLSurv (Liu et al., 2023), and competitive performance than TransMIL (Shao et al., 2021) and HGT (Hou et al., 2023). The c-index results from $Model_0$, $Model_1$, and $Model_2$ demonstrate that establishing correlations between the morphologically similar patches in FUB and latent communications within informative patches in GAB are both powerful feature extractors for WSIs. Based on the performance of $Model_0$ and $Model_2$, GAB did not perform as effectively as FUB on the TCGA-BLCA and TCGA-BRCA datasets. Apart from the intrinsic factors caused by different cancer types, we found that the size of WSIs could also potentially influence the optimization of the neural network. The detailed results and discussion can be found in the Supplementary Material.

By comparing the performance of $Model_1$ and $Model_2$, we observed a performance increase and a decrease in terms of mean and standard deviation, respectively. For instance, on the TCGA-UCEC dataset, the mean c-index increased by 1.2% and the standard deviation decreased by 4.9%. This suggests that momentum operation in GAB can help the model to converge and improve the training stability. Compared with $Model_2$ and $Model_0$, $Model_3$ achieved significant performance improvement in terms of c-index value on the TCGA-LUAD and TCGA-UCEC datasets by at most 6%. Such improvement indicates that two

branches are complementary with each other and accordingly, combining them together can considerably improve the representations from different views. In Supplementary Material Table 4, we also combine TransMIL (Shao et al., 2021) with the GAB and FUB for cancer analysis. However, the performance is not consistently incremental with some having a big drop, which indirectly demonstrates the complementary nature of the FUB and GAB in cancer analysis. After replacing the mean readout function with our proposed AARM, $Model_4$ also achieved performance improvement to a certain degree compared to $Model_3$. Especially on the TCGA-LUAD datasets, the improvement reached 1.9%. In the case of $Model_5$, we added the spatial contextual correlations during the graph construction. Although the performance improvements may not be pronounced, the standard deviation was decreased, indicating that the model's performance on the five-fold cross-validation became more stable. In conclusion, each proposed module or branch is effective and they can work well both independently and cooperatively.

4.5. Kaplan–Meier curve analysis

The Kaplan–Meier estimator, alternatively referred to as the product limit estimator, is a non-parametric statistic employed for estimating the survival function from lifetime data (Kaplan and Meier, 1958). The Kaplan–Meier estimator is instrumental in areas where an understanding of the time until an event occurs is paramount, such as in medical research, biology, engineering, and economics. In this subsection, we applied the Kaplan–Meier curves to evaluate the distinction and separation between the model-predicted high-risk and low-risk patients
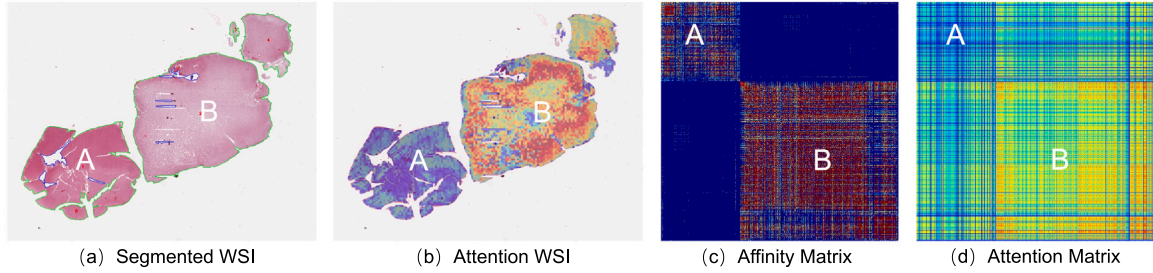
Fig. 4. Case study of the WSI from TCGA-GBMLGG dataset in terms of segmented WSI (a), attention WSI (b), affinity matrix (c), and attention matrix (d).

on the TCGA-BLCA, TCGA-BRCA, TCGA-GBMLGG, TCGA-LUAD, and TCGA-UCEC datasets. As illustrated in Fig. 2, we visualized both the ground truth (GT) and DM-GNN predicted distributions across the five-fold testing data for a more comprehensive comparison. Here, the ground truth referred to the patient's living status at the timepoint t_2. Regarding the testing data, patients were categorized according to the optimal thresholds which referred to receiver operating characteristic (ROC) curves of the training data. The high-risk and low-risk patients were represented by red and blue curves, respectively. The *P-Values* on the TCGA-BLCA, TCGA-BRCA, TCGA-GBMLGG, TCGA-LUAD, and TCGA-UCEC datasets were 0.02346, 0.00042, 0.0000, 0.00419, and 0.04409, respectively. All of the *P-Values* were less than 0.05, indicating a strong statistical significance of the observed difference derived by our proposed DM-GNN. By comparing the distribution between our predicted and ground truth, we observed that the predicted low-risk curves closely resembled the ground truth. With respect to the high-risk patients, our model exhibited a certain degree of hysteresis, indicating that some patients might experience worse outcomes than the model predictions. Overall, we can conclude that the proposed DM-GNN method has a strong capability of distinguishing and stratifying high-risk and low-risk patients.

4.6. Model interpretation

In this subsection, we interpreted our proposed DM-GNN by visualizing the attention distribution, as well as retrieving and depicting the high-attention patches. The interpretations were conducted in high-risk and low-risk cohorts from the TCGA-BLCA dataset. Fig. 3 shows the original WSIs after segmentation (a, d), attention-mapped WSIs (b, e), and high-attention patches (c, f). Generally, cancer cells, characterized by larger nuclei than normal cells, tend to appear darker after undergoing the Hematoxylin and Eosin (HE) staining process (Luna, 1968). This characteristic in appearance aids in the identification and differentiation of cancerous cells from healthy cells during histopathological analysis. By examining the original WSIs and attention-mapped images, it is evident that the model focuses more on cancer cells and their spatial contextual neighbors rather than normal cells (we randomly marked some normal tissue regions with yellow dashed lines), highlighting its superior capability in tissue identification. Furthermore, we retrieved the high-attention patches for each slide (Fig. 3 c, f) and depicted their morphological features by experienced pathologists. The WSI of the low-risk patient indicates a significant presence of immune cells and showed minimal signs of malignant or abnormal attributes, along with a high differentiation degree. On the contrary, the WSI of the high-risk patient displays a high malignancy rate. The nuclei of these tumor cells come in a variety of shapes, some of which are featured by multiple nuclei. Moreover, the chromatin within these cells is tightly compacted, for which there is a lower level of cell differentiation. The depiction of the high-attention patches from pathologists indicated the useful prediction insights provided by the well-trained model, which enhanced the model's interpretation capability. Overall, our proposed DM-GNN represents a useful tool to assist pathologists in cancer patient survival analysis.

4.7. Case study for AFM and MAM visualization

In this subsection, we performed a case study by visualizing the dependencies during the two graph construction processes. In Fig. 4, we specifically showcased the segmented WSI, attention-mapped WSI, AFM, and MAM. To enhance clarity in the visualization, we selected a WSI from the TCGA-GBMLGG dataset that exhibits two distinct morphological tissues (denoted as A and B in Fig. 4) located at different regions of the image. Upon a close inspection of the AFM, we observed that the two tissue regions updated their features independently without exploring any cross-correlations. Conversely, the MAM allowed each patch to establish correlations from a global perspective. Notably, the intra-correlations within tissue A were weak and the intra-correlations within tissue B were very strong. Furthermore, connections between tissues A and B can be observed, indicating inter-correlations between the two regions. Based on the analysis and observations presented, we can conclude that the two graphs constructed in the two branches of the model are theoretically distinct and have the potential to be complementary to each other. MAM emphasized global correlations and established connections between patches based on a global viewpoint. On the other hand, AFM allows independent feature updates within specific tissue regions without considering inter-region correlations. This distinct construction of the two graphs enables the model to capture different aspects of the data, enhancing its overall understanding and representation of the underlying information. By combining the information from both branches, the model can leverage the complementary nature of the two graphs to improve its performance and predictive capabilities.

5. Conclusion

In this study, we have developed a novel dual-stream multi-dependency graph neural network, referred to as DM-GNN to improve histopathology image-based cancer patient survival analysis. Importantly, DM-GNN is capable of modeling the complex correlations between numerous morphology-diverse patches in each WSI, thereby enabling a more profound understanding and inference of patients' survival status. More specifically, DM-GNN models the original WSI as two independent graphs with theoretically distinct dependencies, which focus on the morphological similarities and global co-activating correlations, respectively. Leveraging such a strategy, DM-GNN can successfully establish the deep correlations between patches from the global viewpoint, thereby being able to conduct a comprehensive analysis of each WSI. Moreover, we also propose a new affinity-guided attention recalibration module to enable more robust node-level feature aggregation against noise from multiple perspectives. To assess the performance and utility of the proposed DM-GNN framework, we have performed extensive benchmarking experiments on five TCGA benchmark datasets. We envision that the development and availability of the data-driven deep learning-based tools, such as DM-GNN proposed in this study, can be explored as powerful tools to facilitate community-wide efforts and inform clinical decision-making underpinning digital pathology and precision oncology.

CRediT authorship contribution statement

Zhikang Wang: Writing – original draft, Methodology. **Jiani Ma:** Visualization, Software. **Qian Gao:** Visualization, Data curation. **Chris Bain:** Supervision, Resources. **Seiya Imoto:** Writing – review & editing, Formal analysis. **Pietro Liò:** Resources, Methodology, Formal analysis. **Hongmin Cai:** Writing – review & editing, Software. **Hao Chen:** Writing – review & editing, Validation, Supervision. **Jiangning Song:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.media.2024.103252.

References

Abels, E., Pantanowitz, L., Aeffner, F., Zarella, M.D., van der Laak, J., Bui, M.M., Vemuri, V.N., Parwani, A.V., Gibbs, J., Agosto-Arroyo, E., et al., 2019. Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. J. Pathol. 249 (3), 286–294.

Ba, J.L., Kiros, J.R., Hinton, G.E., 2016. Layer normalization. arXiv preprint arXiv:1607.06450.

Barker, J., Hoogi, A., Depeursinge, A., Rubin, D.L., 2016. Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles. Med. Image Anal. 30, 60–71.

Carmichael, I., Song, A.H., Chen, R.J., Williamson, D.F., Chen, T.Y., Mahmood, F., 2022. Incorporating intratumoral heterogeneity into weakly-supervised deep learning models via variance pooling. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part II. Springer, pp. 387–397.

Chen, R.J., Chen, C., Li, Y., Chen, T.Y., Trister, A.D., Krishnan, R.G., Mahmood, F., 2022. Scaling vision transformers to gigapixel images via hierarchical self-supervised learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 16144–16155.

Chen, R.J., Lu, M.Y., Shaban, M., Chen, C., Chen, T.Y., Williamson, D.F., Mahmood, F., 2021. Whole slide images are 2d point clouds: Context-aware survival prediction using patch-based graph convolutional networks. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VIII 24. Springer, pp. 339–349.

Cheng, J., Mo, X., Wang, X., Parwani, A., Feng, Q., Huang, K., 2018. Identification of topological features in renal tumor microenvironment associated with patient survival. Bioinformatics 34 (6), 1024–1030.

Deschoolmeester, V., Baay, M., Van Marck, E., Weyler, J., Vermeulen, P., Lardon, F., Vermorken, J.B., 2010. Tumor infiltrating lymphocytes: an intriguing player in the survival of colorectal cancer patients. BMC Immunol. 11 (1), 1–12.

Di, D., Zou, C., Feng, Y., Zhou, H., Ji, R., Dai, Q., Gao, Y., 2022. Generating hypergraph-based high-order representations of whole-slide histopathological images for survival prediction. IEEE Trans. Pattern Anal. Mach. Intell. 45 (5), 5800–5815.

Dunnwald, L.K., Rossing, M.A., Li, C.I., 2007. Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients. Breast Cancer Res. 9, 1–10.

Elston, C.W., Ellis, I.O., 1991. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 19 (5), 403–410.

Gao, Z., Hong, B., Li, Y., Zhang, X., Wu, J., Wang, C., Zhang, X., Gong, T., Zheng, Y., Meng, D., et al., 2023. A semi-supervised multi-task learning framework for cancer classification with weak annotation in whole-slide images. Med. Image Anal. 83, 102652.

Gooden, M.J., de Bock, G.H., Leffers, N., Daemen, T., Nijman, H.W., 2011. The prognostic influence of tumour-infiltrating lymphocytes in cancer: a systematic review with meta-analysis. Br. J. Cancer 105 (1), 93–103.

Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B., 2009. Histopathological image analysis: A review. IEEE Rev. Biomed. Eng. 2, 147–171.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.

Hou, W., He, Y., Yao, B., Yu, L., Yu, R., Gao, F., Wang, L., 2023. Multi-scope analysis driven hierarchical graph transformer for whole slide image based cancer survival prediction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 745–754.

Huang, Z., Chai, H., Wang, R., Wang, H., Yang, Y., Wu, H., 2021. Integration of patch features through self-supervised learning and transformer for survival analysis on whole slide images. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VIII 24. Springer, pp. 561–570.

Ilse, M., Tomczak, J., Welling, M., 2018. Attention-based deep multiple instance learning. In: International Conference on Machine Learning. PMLR, pp. 2127–2136.

Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53 (282), 457–481.

Kipf, T.N., Welling, M., 2017. Semi-supervised classification with graph convolutional networks. In: International Conference on Learning Representations. ICLR.

Kumar, V., Abbas, A.K., Fausto, N., Aster, J.C., 2014. Robbins and Cotran Pathologic Basis of Disease, Professional Edition E-Book. Elsevier health sciences.

Li, B., Li, Y., Eliceiri, K.W., 2021. 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. pp. 14318–14328.

Li, R., Yao, J., Zhu, X., Li, Y., Huang, J., 2018. Graph CNN for survival analysis on whole slide pathological images. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II. Springer, pp. 174–182.

Liu, P., Ji, L., Ye, F., Fu, B., 2023. GraphLSurv: A scalable survival prediction network with adaptive and sparse structure learning for histopathological whole-slide images. Comput. Methods Programs Biomed. 231, 107433.

Louis, D.N., Feldman, M., Carter, A.B., Dighe, A.S., Pfeifer, J.D., Bry, L., Almeida, J.S., Saltz, J., Braun, J., Tomaszewski, J.E., et al., 2016. Computational pathology: a path ahead. Arch. Pathol. Lab. Med. 140 (1), 41–50.

Lu, M.Y., Williamson, D.F., Chen, T.Y., Chen, R.J., Barbieri, M., Mahmood, F., 2021. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat. Biomed. Eng. 5 (6), 555–570.

Luna, L.G., 1968. Manual of Histologic Staining Methods of the Armed Forces Institute of Pathology. Blakiston Division, McGraw-Hill.

Maron, O., Lozano-Pérez, T., 1997. A framework for multiple-instance learning. Adv. Neural Inf. Process. Syst. 10.

Mobadersany, P., Yousefi, S., Amgad, M., Gutman, D.A., Barnholtz-Sloan, J.S., Velázquez Vega, J.E., Brat, D.J., Cooper, L.A., 2018. Predicting cancer outcomes from histology and genomics using convolutional networks. Proc. Natl. Acad. Sci. 115 (13), E2970–E2979.

Narod, S., 2012. Tumour size predicts long-term survival among women with lymph node-positive breast cancer. Curr. Oncol. 19 (5), 249–253.

Richardson, S.M., Wurtz, L.D., Collier, C.D., 2022. Ninety percent or greater tumor necrosis is associated with survival and social determinants of health in patients with osteosarcoma in the National Cancer Database. Clin. Orthop. Relat. Res.® 10–1097.

Shao, Z., Bian, H., Chen, Y., Wang, Y., Zhang, J., Ji, X., et al., 2021. Transmil: Transformer based correlated multiple instance learning for whole slide image classification. Adv. Neural Inf. Process. Syst. 34, 2136–2147.

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F., 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: Cancer J. Clin. 71 (3), 209–249.

Wang, Z., Bi, Y., Pan, T., Wang, X., Bain, C., Bassed, R., Imoto, S., Yao, J., Daly, R.J., Song, J., 2023a. Targeting tumor heterogeneity: multiplex-detection-based multiple instance learning for whole slide image classification. Bioinformatics 39 (3), btad114.

Wang, Z., Gao, Q., Yi, X., Zhang, X., Zhang, Y., Zhang, D., Liò, P., Bain, C., Bassed, R., Li, S., et al., 2023b. Surformer: An interpretable pattern-perceptive survival transformer for cancer survival prediction from histopathology whole slide images. Comput. Methods Programs Biomed. 241, 107733.

Wang, Z., Li, J., Pan, Z., Li, W., Sisk, A., Ye, H., Speier, W., Arnold, C.W., 2021. Hierarchical graph pathomic network for progression free survival prediction. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VIII 24. Springer, pp. 227–237.

Wang, Y., Wang, Y.G., Hu, C., Li, M., Fan, Y., Otter, N., Sam, I., Gou, H., Hu, Y., Kwok, T., et al., 2022. Cell graph neural networks enable the precise prediction of patient survival in gastric cancer. NPJ Precis. Oncol. 6 (1), 45.

Wang, H., Xing, F., Su, H., Stromberg, A., Yang, L., 2014. Novel image markers for non-small cell lung cancer classification and survival prediction. BMC Bioinformatics 15, 1–12.

Wang, Z., Zhang, Y., Xu, Y., Imoto, S., Chen, H., Song, J., 2024. Histo-genomic knowledge distillation for cancer prognosis from histopathology whole slide images. arXiv preprint arXiv:2403.10040.

Yao, J., Wang, S., Zhu, X., Huang, J., 2016. Imaging biomarker discovery for lung cancer survival prediction. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19. Springer, pp. 649–657.

Yu, J.-G., Wu, Z., Ming, Y., Deng, S., Li, Y., Ou, C., He, C., Wang, B., Zhang, P., Wang, Y., 2023. Prototypical multiple instance learning for predicting lymph node metastasis of breast cancer from whole-slide pathological images. Med. Image Anal. 85, 102748.

Yu, K.-H., Zhang, C., Berry, G.J., Altman, R.B., Ré, C., Rubin, D.L., Snyder, M., 2016. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nature Commun. 7 (1), 12474.

Yuan, Y., 2016. Spatial heterogeneity in the tumor microenvironment. Cold Spring Harb. Perspect. Med. 6 (8), a026583.

Yuan, Y., Failmezger, H., Rueda, O.M., Ali, H.R., Gräf, S., Chin, S.-F., Schwarz, R.F., Curtis, C., Dunning, M.J., Bardwell, H., et al., 2012. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. Sci. Transl. Med. 4 (157), 157ra143.

Zadeh, S.G., Schmid, M., 2020. Bias in cross-entropy-based training of deep survival networks. IEEE Trans. Pattern Anal. Mach. Intell. 43 (9), 3126–3137.

Zaheer, M., Kottur, S., Ravanbakhsh, S., Poczos, B., Salakhutdinov, R.R., Smola, A.J., 2017. Deep sets. Adv. Neural Inf. Process. Syst. 30.

Zhang, J., Gold, K.A., Lin, H.Y., Swisher, S.G., Xing, Y., Lee, J.J., Kim, E.S., William, Jr., W.N., 2015. Relationship between tumor size and survival in non–small-cell lung cancer (NSCLC): an analysis of the surveillance, epidemiology, and end results (SEER) registry. J. Thorac. Oncol. 10 (4), 682–690.

Zhao, L., Hou, R., Teng, H., Fu, X., Han, Y., Zhao, J., 2023. CoADS: Cross attention based dual-space graph network for survival prediction of lung cancer using whole slide images. Comput. Methods Programs Biomed. 236, 107559.

Zhu, X., Yao, J., Huang, J., 2016. Deep convolutional neural network for survival analysis with pathological images. In: 2016 IEEE International Conference on Bioinformatics and Biomedicine. BIBM, IEEE, pp. 544–547.

Zhu, X., Yao, J., Zhu, F., Huang, J., 2017. Wsisa: Making survival prediction from whole slide histopathological images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7234–7242.