This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

XFibrosis: Explicit Vessel-Fiber Modeling for Fibrosis Staging from Liver Pathology Images

Chong Yin[1], Siqi Liu[2], Fei Lyu[1], Jiahao Lu[3], Sune Darkner[3], Vincent Wai-Sun Wong[4], Pong C. Yuen[1]
[1]Department of Computer Science, Hong Kong Baptist University, Hong Kong
[2]Shenzhen Research Institute of Big Data, Chinese University of Hong Kong, Shenzhen
[3]Department of Computer Science, University of Copenhagen, Denmark
[4]Department of Medicine and Therapeutics, Chinese University of Hong Kong, Hong Kong
{chongyin, feilyu, pcyuen}@comp.hkbu.edu.hk, siqiliu@sribd.cn
{lu, darkner}@di.ku, wongv@cuhk.edu.hk

Abstract

*The increasing prevalence of non-alcoholic fatty liver disease (NAFLD) has caused public concern in recent years. The high prevalence and risk of severe complications make monitoring NAFLD progression a public health priority. Fibrosis staging from liver biopsy images plays a key role in demonstrating the histological progression of NAFLD. Fibrosis mainly involves the deposition of fibers around vessels. Current deep learning-based fibrosis staging methods learn spatial relationships between tissue patches but do not explicitly consider the relationships between vessels and fibers, leading to limited performance and poor interpretability. In this paper, we propose an e**X**plicit vessel-fiber modeling method for **Fibrosis** staging from liver biopsy images, namely XFibrosis. Specifically, we transform vessels and fibers into graph-structured representations, where their micro-structures are depicted by vessel-induced primal graphs and fiber-induced dual graphs, respectively. Moreover, the fiber-induced dual graphs also represent the connectivity information between vessels caused by fiber deposition. A primal-dual graph convolution module is designed to facilitate the learning of spatial relationships between vessels and fibers, allowing for the joint exploration and interaction of their microstructures. Experiments conducted on two datasets have shown that explicitly modeling the relationship between vessels and fibers leads to improved fibrosis staging and enhanced interpretability.*

1. Introduction

Non-alcoholic fatty liver disease (NAFLD) is the most prevalent liver disease globally. It affects around 30% of the population and has raised public concern in recent years

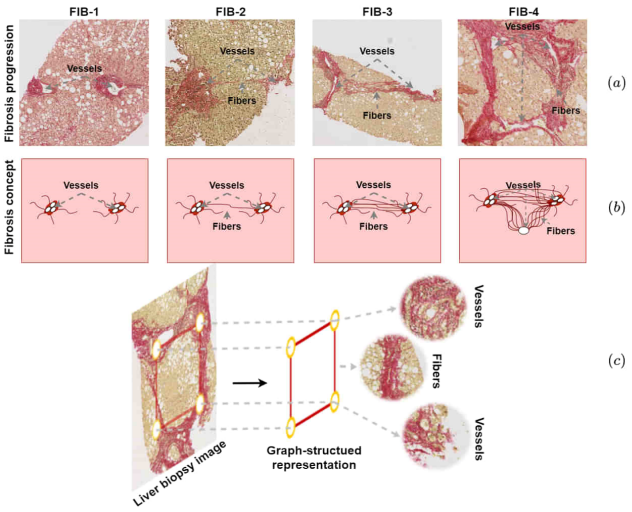

Figure 1. Illustration of fibrosis staging from liver biopsy images and the proposed graph-structured representation. (a) Liver biopsy image samples with different fibrosis stages. (b) Concept illustration of different fibrosis stages. (c) The illustration of a graph-structured representation. It graphically depicts the interaction between vessels and fibers. Vessels are regarded as nodes and the interactions caused by fiber deposition as edges of the graph.

[26]. NAFLD consists of a spectrum of liver damage ranging in severity from simple harmless fatty liver to various degrees of fibrosis. Patients with advanced fibrosis are at an increased risk of developing liver cancer and other complications. Consequently, an accurate diagnosis of liver fibrosis facilitates closer monitoring and enhances clinical care effectiveness. In clinical practice, pathologists will assign a semi-quantitative fibrosis stage (FIB) [31] to determine the extent of fibrosis. Liver biopsy images offer cellular-level visual insights into microstructures, aiding in fibrosis staging. Diagnosing fibrosis is a laborious process with low

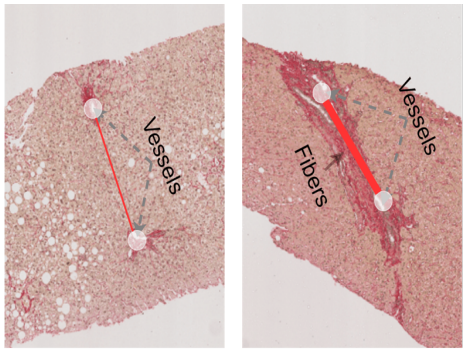

Figure 2. A simple graph is depicted with two white circles representing nodes connected by a red line.

agreement among pathologists. [37]. The rising number of patients would also impose a significant diagnostic burden on the current healthcare system. Developing an automated image analysis model to aid pathologists is crucial.

Deep learning methods have shown promising results in the analysis of whole-slide pathology images (WSIs) [45–47]. Existing pathological image analysis methods can be broadly categorized into two groups. One approach is multi-instance learning (MIL)-based methods [25, 32]. In the MIL framework, the WSI is treated as a collection of instances. However, these methods may not effectively capture the spatial context and relationships between different instances, which are crucial for the diagnosis of the disease. Another approach is graph neural network (GNN)-based methods, which have demonstrated remarkable capabilities in analyzing WSIs [11, 12, 17]. These methods aim to learn topological features between individual objects such as nuclei and tissues, which are represented as graph nodes. [40] estimate fibrosis stages by converting the WSI into a graph, with regularly sliced image patches representing the graph nodes. While the patch-to-patch graph offers a basic representation of spatial relationships between tissue patches, it may not accurately capture fibrosis-related micro-structures, resulting in limited interpretability.

In this work, we propose to explicitly exploit fibrosis-related micro-structures for liver biopsy image analysis. The progression of fibrosis, as defined by semi-quantitative definition [31], is characterized by the deposition of fibers around vessels in liver biopsy images, as depicted in Figure 1. (a) displays liver biopsy images with varying stages of fibrosis (FIB-1 to FIB-4). Figure 1 (b) illustrates the concept of fibrosis, where fibers are deposited around vessels (vessels appear white, fibers appear red). Fibrosis initiates with fiber deposition around the portal tract (FIB-1). As fibers accumulate, they expand and connect portal tracts at different strengths (FIB-2 to FIB-3). Eventually, fibers form bridges, connecting portal tracts and central veins (FIB-4). The relationship between vessels and fibers highlights the inherent advantage of representing images with graphs, as depicted in Figure 1 (c). The vessels in liver biopsy images can be considered as nodes in a graph. Fibers of varying strengths are deposited around the vessels, which can be considered as the edges of the graph. The interactions between vessels and fibers manifest as a graph structure.

We propose a mechanism to represent histopathology images as graphs, where the graph nodes denote vessel locations and the graph edges describe fiber features. Figure 2 shows a simple graph comprising two graph nodes (white circle) and one graph edge (red line). The graph nodes indicate the location of vessel regions, and an edge is created when two graph nodes are nearby. In Figure 2 (a), there is no fiber deposition between the two vessels, while in Figure 2 (b), heavy fiber deposition is visible, with the fibers expanding and connecting the two vessels. We represent and encode this fiber information as edge features, denoted by thin and thick lines. This graph definition allows us to model the relationships between vessels and fibers. The ability to jointly explore two micro-structures and their interactions is crucial for graph-structured representation. Based on primal-dual graph theory [18], we can represent vessels and fibers as two graphs, as depicted in Figure 3. The two graphs depict the micro-architectures of vessels and fibers, with graph nodes indicating relevant biological entities. Moreover, the two graphs form a primal-dual relationship. The nodes in the dual graph also serve as the edges in the primal graph, preserving the connection information from fiber deposition between the vessels.

This paper proposes the explicit learning of vessel-fiber relationships for fibrosis staging (XFibrosis) from liver biopsy images. Specifically, XFibrosis represents liver biopsy images as a set of primal and dual graphs induced by vessels and fibers, respectively. These generated graphs mirror the respective micro-architectures of vessels and fibers. The primal-dual relationship enables the model to concurrently examine and interact with the two micro-structures. Additionally, XFibrosis proposes a primal-dual graph convolutional module to learn the relationships between vessels and fibers. The proposed module facilitates information propagation within each graph, thereby enabling micro-structure feature learning. Moreover, by leveraging primal-dual relations, it is possible to induce information propagation between the two graphs, which fosters interaction feature learning. Our contributions are summarized as follows:

- We propose an interpretable method for explicitly learning the relationships between vessels and fibers from liver biopsy images for the staging of fibrosis, termed as XFibrosis.
- XFibrosis converts liver biopsy images into primal-dual graphs induced by vessels and fibers.

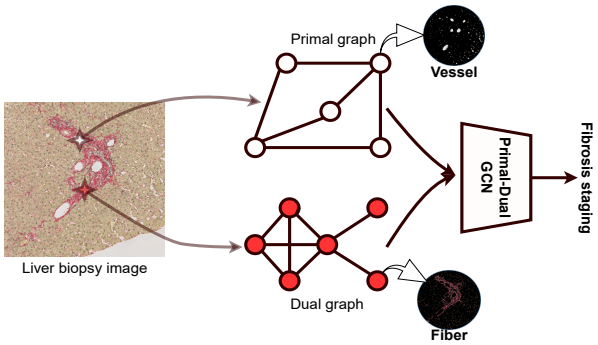

Figure 3. The basic idea of our proposed method. Vessels and fibers are represented by two graphs forming a primal-dual relationship. The primal-dual graphs are further fed into the primal-dual GCN module for spatial relation learning.

- We introduce a primal-dual graph convolution module that learns the relationships between vessels and fibers by concurrently leveraging micro-structures and interactions.

2. Related Works

Multi-instance Learning on Histopathology Images. Multiple-instance learning (MIL) [8, 21, 29] demonstrates a robust capability in handling whole slide images where only slide-level labels are available. In the MIL framework, each Whole Slide Image (WSI) is considered as a bag containing multiple patches cropped from the slide. The slide is diagnosed as positive if at least one of the patches indicates disease. Adhering to a strict binary assumption, MIL-RNN [7] aggregates features from positive instances using a recurrent neural network. AttenMIL [20] introduces an attention-based aggregation method within the MIL framework for diagnosing breast and colon cancer. For tasks involving multiclass classification, ReMix [42] develops a data augmentation method to increase data diversity. IB-MIL [27] further improves feature representation by eliminating confounding factors. However, these patch-based MIL methods do not adequately address fibrosis staging, which depends on the relationships between instances.

Graph Learning on Histopathology Images. Recently, advanced graph convolutional networks have significantly influenced digital histopathological image analysis [3, 12, 17]. Graph representations are apt for modeling relationships between biomedical entities (e.g., cells, nuclei, tissues) observed in histological images. Converting the WSIs into a graph representation is a crucial step in the analysis of histopathological images. Current graph learning-based methods are categorized into cell-graphs [4, 36], patch-graphs [11,30], and tissue-graphs [1,9], based on the choice of graph nodes. The current methods primarily focus on deciphering spatial relationships between single objects, represented by graph nodes in histopathology images. The graph edges only provide limited information, which is derived from these nodes. This approach is insufficient for fibrosis staging, which requires an in-depth examination of two critical micro-structures: vessels and fibers.

3. Proposed Method

As shown in Figure 4, we propose an explicit vessel-fiber modeling method for fibrosis staging from WSIs of the liver. Each WSI is processed through a vessel-fiber segmentor, yielding segments of vessels and fibers. In each image tile containing the segments, two sets of key points are generated and used to form a primal-dual graph, with these key points acting as graph nodes. These graph-structured representations are then inputted into the primal-dual graph convolution module to understand the spatial relationships between vessels and fibers. The tile-level features produced by each image tile are eventually combined to generate WSI-level features for fibrosis staging.

3.1. Primal-Dual Graph Learning-Embedding

Fibrosis staging from liver biopsy images relies on spatial relationships between vessels and fibers. A pivotal aspect of our approach involves generating informative graphs that depict these spatial relationships. To drive the model to focus on micro-structures related to fibrosis (e.g. vessels, fibers), we first develop a vessel-fiber segmentor to extract the vessels and fibers from the WSI I_i. Inspired by [31], vessels (e.g., portal tracts, central veins) appear in the white region. Fibers are colored red or blue in a specific stain color. Following [41], we implement the vessel-fiber segmentor with the mean shift algorithm. Image tiles that contain vessels or fibers are cut out and transformed into graph-structured representations.

In this section, we describe how to transform an image tile into a primal-dual graph. It mainly involves primal-dual graph construction and graph node embedding.

Primal-Dual Graph Construction. For each image tile t_{ij} sliced from WSI I_i, the corresponding vessel segments s_{ij}^v and fiber segments s_{ij}^f, we develop a key points generator $\mathcal{P}$ to get the key points set P_{ij} which describes the micro-architectures:

$$P_{ij} = \mathcal{P}(s_{ij}^v, s_{ij}^f) \tag{1}$$

where P_{ij} contains two subsets of points $P_{ij} = \{\mathcal{V}_{ij}, \hat{\mathcal{V}}_{ij}\}$. These two sets of key points carry distinct meanings. $\mathcal{V}_{ij} = \{v_{ij1}, v_{ij2}, \ldots, v_{ijn_{ij}^v}\}$ denotes the centroids or contours of segments, which are called vessel-induced key points. Using these key points as graph nodes, we construct the vessel-

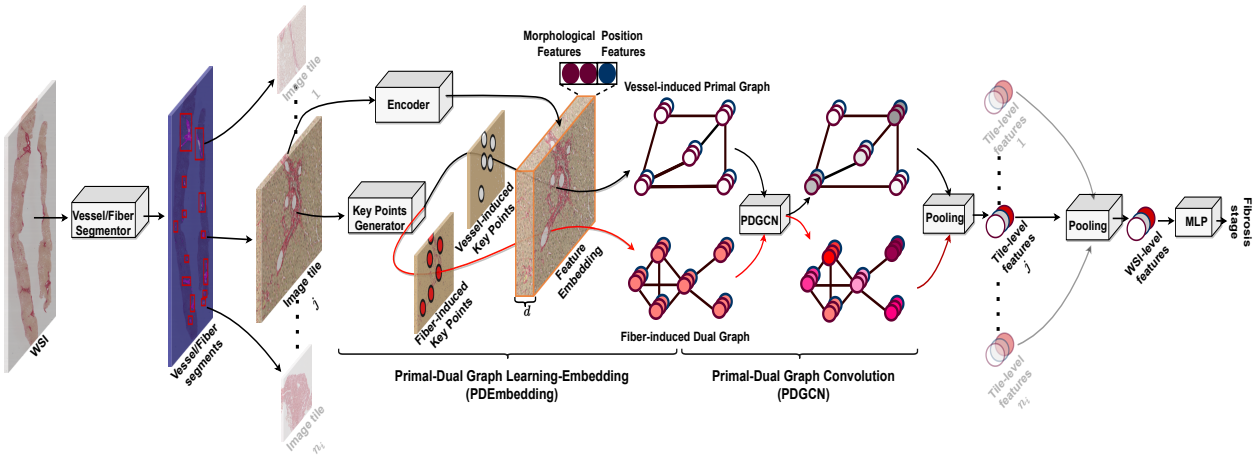

Figure 4. Overview of the proposed method. Each image tile from the WSI of the liver, containing vessel and fiber segments, is transformed into primal-dual graph representations using primal-dual graph learning-embedding (PDEmbedding). The graph-structured data are fed into primal-dual graph convolution (PDGCN) for learning spatial relationships. All tile-level features are aggregated to generate WSI-level features for fibrosis staging.

induced primal graph G_{ij} utilizing the Delaunay Triangulation operator [22] $\mathcal{T}$:

$$G_{ij} = \mathcal{T}(P_{ij}^{v}) \tag{2}$$

where the associated adjacent matrix A_{ij} records the connection information between graph nodes.

The nodes in the primal graph record the information from the vessel. The graph edges should indicate the difference in the microenvironment around vessels, particularly the connectivity caused by fiber deposition. Although general graph convolutional networks (GCNs) are effective in learning interactions among graph nodes, they typically deal with a single object [4]. The edge feature is ignored or reflects the similarity between graph nodes, which fails to represent the fiber information. In order to represent the fiber information that is captured in edges using GCNs, we propose the use of a dual graph. The nodes in the dual graph represent the edges between the nodes of the primal graph, corresponding to the features of the fibers. Once the primal graph is given, its dual graph $\hat{G}_{ij}$ is automatically constructed by altering the roles of graph nodes and edges based on primal-dual graph transform operator $\hat{\mathcal{T}}$ [38]:

$$\hat{G}_{ij} = \hat{\mathcal{T}}(G_{ij}) \tag{3}$$

where $\hat{G}_{ij}$ is referred to as the fiber-induced dual graph. The nodes and edges in the dual graph are determined at the same time. The associated nodes in a dual graph $\hat{G}_{ij}$ are represented as fiber-induced key points $\hat{\mathcal{V}}_{ij} = \{\hat{v}_{ij1}, \hat{v}_{ij2}, \ldots, \hat{v}_{ijn_{ij}^{f}}\}$. They denote the points located in the fiber region.

The sequentially constructed graphs G_{ij} and $\hat{G}_{ij}$ represent two distinct micro-structures and form a primal-dual relationship. The primal graph G_{ij} portrays the topology of the fibers extending around the vessels with graph nodes representing the location of vessels or the end of the fiber extension. The edge of the vessel-induced primal graph traverses the fiber region, while the dual graph captures the intricate morphological characteristics of these edges.

Graph Node Embedding. The image tile containing vessels and fibers is represented by a vessel-induced primal graph and a fiber-induced dual graph. Subsequently, it becomes necessary to learn a feature representation for node embedding. In this context, only the node embedding in the primal graph G_{ij} is illustrated, while the node embedding in the dual graph $\hat{G}_{ij}$ follows a similar pattern.

We learn a central node representation from patches around the centroids of graph nodes. The associated image patch for t^{th} graph node p_{ijt}^{r} is denoted as $p_{ijt}^{r} \in R^{r\times r\times 3}$.

To retain the morphological features, deep neural networks-based feature encoder $\mathcal{F}$ are used to automatically learn a feature representation z^{M} from image patch p_{ijt}^{r}:

$$z_{ijt}^{M} = \mathcal{F}(p_{ijt}^{r}) \tag{4}$$

where $z_{ijt}^{M} \in R^{d_M}$, d_M denotes the dimension of features.

Following the usage of Position Embedding (PE) in the vision transformer [13], the global location of each image tile in WSIs is also important for image recognition. To retain the spatial information in WSI, we encode the position information of each patch via PE:

$$\begin{cases} z_{ijt}^{X} = PE_X(p_{ijt}^{r}) \\ z_{ijt}^{Y} = PE_Y(p_{ijt}^{r}) \end{cases} \tag{5}$$

where $z_{ijt}^{X} \in R^{d_X}$ and $z_{ijt}^{Y} \in R^{d_Y}$ are the position encoding based on the coordinates with respect to X-axis and

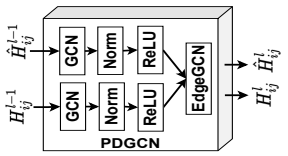

Figure 5. Detailed structure of primal-dual graph convolution module.

Y-axis. d_X and d_Y are the dimensions of position feature embedding. The (x, y) is the coordinates centered at the image patch p^r_{ijt}.

Following [4, 19, 28] which mainly focus on small-scale histopathology image datasets, the position embeddings $\{z^X_{ijt}, z^Y_{ijt}\}$ and morphological features z^M_{ijt} are stacked together to generate the node feature embedding f_{ijt}:

$$f_{ijt} = z^M_{ijt} \oplus z^X_{ijt} \oplus z^Y_{ijt} \tag{6}$$

where $\oplus$ is the concatenation operator. The dimension of the feature embedding f_{ijt} is the sum of these three components: $d = d_M + d_X + d_Y$. Hence f_{ijt} records the information from both morphological features and locations.

Given an image tile t_{ij}, the associated vessel-induced primal graph is $G_{ij}(\mathcal{V}, \mathcal{E})$, the corresponding node feature is $F_{ij} \in \mathbb{R}^{|\mathcal{V}| \times d}$. The fiber-induced dual graph is $\hat{G}_{ij}(\hat{\mathcal{V}}, \hat{\mathcal{E}})$, the corresponding node feature is $\hat{F}_{ij} \in \mathbb{R}^{|\hat{\mathcal{V}}| \times d}$, which also represents the edge features for primal graph.

3.2. Primal-Dual Graph Convolution

Each image tile is converted into a graph-structured representation. We propose a primal-dual graph convolution module to comprehend the spatial relationships between vessels and fibers. Figure 5 illustrates the detailed structure of the proposed primal-dual graph convolution module.

Since vessels and fibers are two distinct micro-structures, we first develop two separate graph convolutional layers, $\mathcal{G}_p$ and $\mathcal{G}_d$, to learn their micro-architectures:

$$\begin{cases} H^l_{ij} = \sigma \circ \mathcal{N} \circ \mathcal{G}_p(A_{ij}, H^{l-1}_{ij}) \\ \hat{H}^l_{ij} = \sigma \circ \mathcal{N} \circ \mathcal{G}_d(\hat{A}_{ij}, \hat{H}^{l-1}_{ij}) \end{cases} \tag{7}$$

where $H^0_{ij} = F_{ij}$, $\hat{H}^0_{ij} = \hat{F}_{ij}$ are the hidden representation. Each graph convolutional layer is followed by a graph feature normalization layer $\mathcal{N}$ [14] and a non-linear activation layer σ [2]. $\circ$ denotes the function composition. The graph node features are updated by exchanging information with its neighbors.

In addition to the micro-architectures, the interactions between vessels and fibers are also important for fibrosis staging. Leveraging the primal-dual graph property, the nodes in the dual graph function as edges of the primal graph. Consequently, we employ the edge-enhanced graph convolution layer G_e. This layer utilizes multidimensional features from nodes and edges for information propagation, leading to the following updates in the hidden representations:

$$[H^l_{ij}; \hat{H}^l_{ij}] = \mathcal{G}_e(A_{ij}, [H^{l-1}_{ij}; \hat{H}^{l-1}_{ij}]) \tag{8}$$

The tile-level features for each image tile emerge from the aggregation of node features, derived from both the vessel-induced primal graph and the fiber-induced dual graph. We utilize the sum-pooling operation $\mathcal{P}ool_{sum}$ to amalgamate graph representations, subsequently generating the tile-level features f_{ij}. The power of this method comes from its ability to capture the whole structure of the full graph representation.

$$f_{ij} = \mathcal{P}ool_{sum}(H^l_{i,j}, \hat{H}^l_{ij}) \tag{9}$$

where $f_{ij} \in R^{d_h}$, and d_h denotes the dimension of hidden feature space.

For each WSI x_i, we obtain a set of tile-level representations $\{f_{i1}, \ldots, f_{in_i}\}$. We then aggregate these tile-level features to create a WSI-level representation. Since fibrosis only occurs in tissue regions that have experienced damage or an immune response, not all regions contribute to it. Therefore, our goal is to select the most significant, typically damaged or responsive areas, to create WSI-level features. Inspired by TopK pooling introduced in [16], for a given tile-level representation $f_{ij} \in \mathcal{R}^{d_h}$, we calculate an importance score s_{ij} using the scorer function $\mathcal{S}_\theta$:

$$s_{ij} = \mathcal{S}_\theta(f_{ij}) \tag{10}$$

where the importance score s_{ij} indicates the importance of image tile t_{ij}. Tiles in the image with a higher score are more likely to be fibrosis. WSI-level features f_i are generated by applying mean pooling on selected top k_i tile-level representations via $\mathcal{P}ool^{topk}_{mean}$:

$$\begin{cases} f_i = \mathcal{P}ool^{topk}_{mean}(f_{i1}, \ldots, f_{in_i}; k_i) \\ k_i = \underset{k}{\operatorname{argmin}} \sum_{j=0}^{k} s_{ij} > s_\delta \end{cases} \tag{11}$$

where s_δ is the saturation threshold that indicates the level of contribution for these selected k image tiles. It aims to select enough image tiles whose accumulated contribution is above a threshold s_δ. In this way, k_i is adaptively calculated in each image sample.

Multi-layer perceptrons (MLPs) are employed as the classifier to estimate fibrosis stage $\hat{y}_i$:

$$\hat{y}_i = \mathcal{M}LP(f_i) \tag{12}$$

Given the ground truth fibrosis stage y, the objective function to minimize for learning the model is defined as cross-entropy loss $\mathcal{L}_{base}$:

$$\mathcal{L}_{base} = \frac{1}{N} \sum_{i=0}^{N} y_i log \hat{y}_i \tag{13}$$

Table 1. Fibrosis staging with different methods on Liver-Fibrosis-SR dataset

Methods		$AUC_{FIB\geq 1}$	$AUC_{FIB\geq 2}$	$AUC_{FIB\geq 3}$	$AUC_{FIB\geq 4}$	AUC	Accuracy
MIL	ReMix [42]	42.50±14.80	59.86±9.44	56.29±10.49	65.70±13.36	56.09±8.53	36.73±7.74
	DTFD [48]	**83.70**±8.15	77.23±8.42	82.57±7.64	76.79±10.18	80.07±3.09	42.22±8.16
	IBMIL [27]	83.18±10.92	78.17±8.25	81.75±7.49	86.51±6.89	82.40±2.99	45.76±8.81
General GNNs	GATv2 [6]	62.03±16.12	63.88±10.00	66.74±9.45	69.64±11.17	65.57±2.89	37.48±8.15
	ARMA [5]	46.33±16.99	68.48±8.78	63.69±11.27	58.56±14.35	59.26±8.25	37.75±8.43
	HGNN [9]	70.17±14.29	67.31±9.2	72.32±8.7	72.87±10.5	70.67±2.19	41.15±8.54
GNNs for Fibrosis	GNN-Fibrosis [40]	58.23±14.48	69.53±9.20	69.93±8.45	76.30±14.00	68.50±6.51	43.46±8.14
	Ours	83.58±13.77	**78.39**±8.29	**86.11**±7.36	**87.93**±6.50	**84.00**±3.59	**53.63**±7.97

where N is the number of WSIs.

Considering that the FIB reflects the degree of disease severity in fibrosis. Inspired by [15], we apply more substantial penalties to estimated fibrosis stages that significantly deviate from the actual stages. This process is described by the following regularization loss $\mathcal{L}_{reg}$:

$$\mathcal{L}_{reg} = \frac{1}{N}\sum_{i=0}^{N}\sum_{c=0}^{C}\hat{y}_i^c||c - y_i||_2^2 \tag{14}$$

where $\hat{y}_i^c$ is the estimated probability on class c. C is the number of classes. $||\cdot||_2^2$ denotes the $L2\text{-}Norm$ distance.

So, the final objective function $\mathcal{L}$ is the combination of $\mathcal{L}_{reg}$ and $\mathcal{L}_{reg}$:

$$\mathcal{L} = \mathcal{L}_{base} + \lambda\mathcal{L}_{reg} \tag{15}$$

where λ is the balance weights between two loss terms.

4. Experimental Results

In this section, we first describe the datasets that are used to carry out experiments and then report the experimental settings and present results. More experimental results are presented in the supplementary materials.

Datasets and Evaluation Metrics. **Liver-Fibrosis-SR** [49] contains 132 liver WSI collected from 132 patients with NAFLD. The liver tissue sections are stained with Sirius Red (SR). **FibrosisMT-v1.0** contains 47 liver WSIs from two age groups [33] [49]. These WSIs are scanned under a lens 40×. The liver tissue sections are stained with Masson's and Triise (MT). Each liver biopsy image has been assigned a fibrosis stage (FIB 0-4) by an expert pathologist. Detailed information on datasets is provided in supplementary materials. **Evaluation metrics** For each dataset, we randomly split all patients into 3 groups and validated the method using three-fold cross-validation. Following [40, 43, 44], we choose the accuracy and area under the receiver operating characteristic values (AUC) as the evaluation matrix. Comprehensive details regarding the evaluation metrics used in our implementation can be found in the supplementary materials.

Implementation details. The model is trained using the Adam optimizer for 100 epochs. The initial learning rate is set to $1e^{-4}$. We adopt a learning rate schedule of exponential decay with power 0.9. The batch size is set to 2. Following [15], we set the balance weights to $\lambda = 0.05$. The saturation threshold $s_\delta = 0.65$ is determined by cross-validation. More implementations are described in the Supplementary Materials.

4.1. Comparison With State-of-the-Art Methods

We compare our method against the SOTA MIL-based methods ReMix [42], DTFD [48], IBMIL [27] and the GNN-based methods GATv2 [6] ARMA [5] GNN-Fibrosis [40], HGNN [9].

Table 1 shows the fibrosis staging results on the Liver-Fibrosis-SR dataset. It demonstrates that the proposed method is effective in modeling vessel-fiber relations for fibrosis staging. Compared to general GNN-based methods, MIL-based methods show a better performance. IBMIL achieves the second-best results with an AUC of 82.40% and an accuracy of 45.76%. The regular grid image tiles do not accurately represent the spatial relationships between vessels and fibers. Rather than building graphs with regular grid image tiles, our method builds vessel and fiber-induced graphs which enable the model to focus on fibrosis-related regions. The graph-structured representations more accurately depict the micro-architectures associated with the fibrosis stage. We observe similar performance when conducting experiments on the smaller FibrosisMT-v1.0 dataset, which consists of patients from two age groups, as shown in Table 2. Our method consistently improves fibrosis staging performance independent of datasets, achieving the best performance with a mean AUC of 92.82% and an accuracy of 64.40%.

4.2. Graph Interpretation On WSIs

The procedure for generating vessel-induced primal graphs and fiber-induced dual graphs on the WSI is first demonstrated in this section, as depicted in Figure 6. The benefits of introducing the primal-dual graph to comprehend vessel-fiber relations are demonstrated by utilizing K-Means [23] to cluster the tile-level features represented by

Table 2. Fibrosis staging with different methods on FibrosisMT-v1.0 dataset

Methods		$AUC_{FIB\geq 1}$	$AUC_{FIB\geq 2}$	$AUC_{FIB\geq 3}$	$AUC_{FIB\geq 4}$	AUC	Accuracy
MIL	ReMix [42]	67.76±22.77	71.18±16.18	72.41±16.69	95.27±8.30	76.66±10.88	50.11±12.61
	DTFD [48]	67.18±20.20	88.90±10.75	86.72±10.67	91.11±11.74	83.48±9.54	58.73±12.95
	IBMIL [27]	75.48±15.21	85.27±13.57	85.29±16.72	91.77±14.51	84.45±5.82	60.21±14.38
General GNNs	GATv2 [6]	73.30±18.53	79.46±14.87	72.93±17.71	33.57±17.19	64.82±18.22	35.94±13.99
	ARMA [5]	73.62±20.18	72.38±16.64	65.12±18.82	65.13±35.13	69.06±3.96	41.39±14.56
	HGNN [9]	73.86±17.80	73.63±15.78	65.01±18.90	59.99±23.75	68.12±5.90	53.89±14.31
GNNs for Fibrosis	GNN-Fibrosis [40]	76.01±17.31	80.86±14.20	69.48±17.52	67.08±17.56	73.36±5.43	51.16±14.57
	Ours	**80.09**±13.71	**98.53**±3.43	**95.72**±6.05	**96.94**±5.83	**92.82**±7.42	**64.40**±14.40

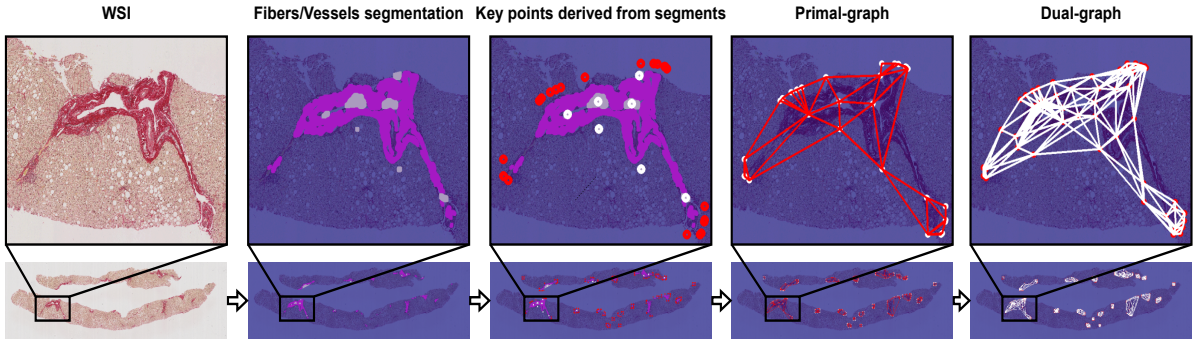

Figure 6. Visualization of the primal-dual graph-based representation. The WSI depicts detected vessel and fiber segments highlighted in white and red, respectively. Key points derived from these segments indicate central vessel locations and fiber region contours, represented as white points. The nodes in the primal graph are used as points, which are then directly converted into the corresponding dual graph.

the graph, with the clusters visualized using t-stochastic neighbor embedding (t-SNE) [10], as illustrated in Figure 7. The distribution of the importance score estimated by TopK on WSI is also plotted and shown in Figure 8. Selected samples are used to illustrate the topology represented by primal graphs in supplementary materials.

Primal-Dual Graph Generation. The primal-dual graph generation procedure is demonstrated in Figure 6. Starting with a WSI, vessel and fiber segments (rendered in white and red colors) are first detected. Key points, which indicate the central locations of vessels or the contour of the fiber regions (rendered as white points), are derived from these segments. These derived key points are utilized as graph node sets. Using Delaunay Triangulation [22], the primal graph is constructed. The primal graph is directly converted into the dual graph. Finally, a collection of primal and dual graphs represents a WSI. These graphs are constructed on image tiles that contain vessels or fibers.

Graph Embedding Cluster. The benefits of introducing graph-structured representations are demonstrated by using K-Means to cluster the tile-level representations. After experimenting with different numbers of clusters, it was found that setting $clusters = 15$ leads to easier interpretation. We then plot the feature embedding using t-SNE [39]. As shown in Figure 7, graph representations show better discriminability. Each feature cluster has a clear boundary with each other. Some points are also sampled from the clusters, and their associated image tiles and graphs are visualized. From a geometric perspective, graph-structured representations enable the model to cluster fiber regions. The samples in each cluster correspond to the specific characteristics of the graphs. The image tiles on the left depict a large area of dense fibers, which are represented by complex graphs. Conversely, simple graphs show the smaller fibers in the image tiles from the other clusters. The fiber density and structure are well reflected by the graphical representation.

Importance Score On WSI. The estimated importance score on WSIs of the liver is visualized to demonstrate that our proposed model indeed learns to focus on the fibrosis region. Regions with higher scores (rendered in red) are more likely to have fibrosis. The distribution of importance scores on slides diagnosed as FIB-1 to FIB-4 is demonstrated in Figure 8. The concentration of the importance score mainly in regions containing fibers or vessels aligns with clinical practice. This alignment occurs because our approach transforms only those image tiles containing vessels or fibers into graph-structured representations, rather than building graphs on whole-slide images. The creation of accurate graphical representations through this method allows for explicit modeling of fibrosis, thereby enhancing interpretability. Furthermore, our model is capable of geomet-

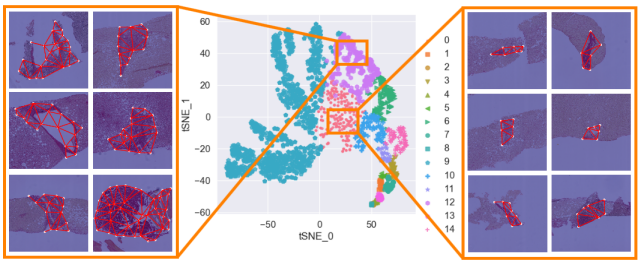

Figure 7. t-SNE plot of image tiles containing vessels or fibers clustered by k-means clustering method using tile-level features.

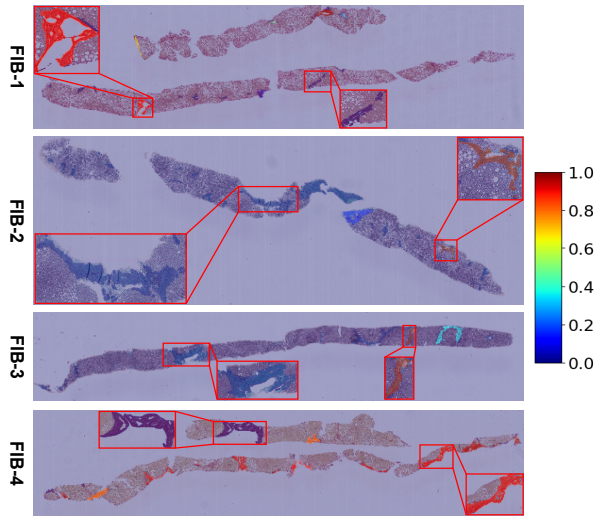

Figure 8. Distribution of the importance score on WSIs. The importance scores are in the range from 0 to 1 (blue to red). Regions with higher scores are more likely to be fibrosis regions.

rically distinguishing between two fiber regions, even when they have similar densities. Despite some areas demonstrating a higher fiber density than the actual fibrosis area (FIB-1 to FIB-4), our model assigns them a lower importance score. This discrepancy is primarily due to the variations in typologies and positions of these areas, as illustrated in the graphical representation.

4.3. Ablation Study On Graph Convolution

Effectiveness of PDGCN. To evaluate the effectiveness of our proposed PDGCN module, we initially exclude the PDGCN and employ solely the PDEmbedding for fibrosis staging. As demonstrated in Table 3, utilizing only PDEmbedding, the model attains an AUC of 80.14% and an accuracy of 48.77%. Each image tile comprises vessels and fibers, facilitating the model's focus on the disease region inducing fibrosis. The incorporation of the PDGCN results in an increase of 3.96% and 4.86%. The enhancements are attributable to the spatial relationships between vessels and fibers discerned by the PDGCN. The PDGCN encodes morphological features in graph nodes, allowing the model to exchange messages between graphs and extract multi-dimensional information from graph nodes and edges.

Table 3. Effectiveness of primal-dual graph convolutional block.

Methods	AUC	Accuracy
PDEmbedding	80.14±3.15	48.77±4.63
PDEmbedding + PDGCN	**84.00**±3.59	**53.63**±7.97

Table 4. Variants of GCN architecture choice.

GCN layers	AUC	Accuracy
GEN [24]	82.85±2.39	44.08± 3.74
GATv2 [6]	82.98±1.75	47.15±3.61
PDN [34]	82.14±3.14	43.73±1.88
Transformer [35]	81.10±2.44	42.53±5.62
PDGCN (Ours)	**84.00**±3.59	**53.63**±7.97

Variances of GCN. We compare the proposed PDGCN with other edge-enhanced graph convolutional networks [6, 24, 34, 35]. Keeping the PDEmbedding the same, we replace the PDGCN with these modules to enable spatial learning of fiber and vessel connections. As shown in Table 4, compared with other edge-enhanced graph convolution methods, the proposed PDGCN demonstrates leading performance. As observable, our proposed PDGCN surpasses the edge-enhanced graph convolution methods in performance. All edge-enhanced graph convolution methods concentrate on learning node features while overlooking the abundant edge information furnished by fibers. The proposed PDGCN comprises two distinct graph convolutional layers, capable of learning features from vessels and fibers independently. Additionally, the proposed PDGCN can merge features of vessels and fibers to enhance understanding of their relationship.

5. Conclusion

In this paper, our proposed model utilizes a primal-dual graph to explicitly interpret vessel-fiber interactions for fibrosis staging in liver biopsy images. We use a vessel-induced primal graph and a fiber-induced dual graph to capture topological details of fibrosis-associated structures identified in WSIs. Furthermore, a primal-dual graph convolution module is introduced to independently learn the features of both vessels and fibers while concurrently understanding their interactions. Experimental results from two datasets demonstrate the effectiveness of our proposed method in terms of classification performance and interpretability. Future research may involve studying the modeling of multiple tissue or cellular components by analyzing multiple subgraphs.

Acknowledgement. This work was supported by the Hong Kong Research Grants Council General Research Fund under Grant RGC/HKBU12200122.

References

[1] Mohammed Adnan, Shivam Kalra, and Hamid R Tizhoosh. Representation learning of histopathology images using graph neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 988–989, 2020. 3

[2] Abien Fred Agarap. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*, 2018. 5

[3] David Ahmedt-Aristizabal, Mohammad Ali Armin, Simon Denman, Clinton Fookes, and Lars Petersson. A survey on graph-based deep learning for computational histopathology. *Computerized Medical Imaging and Graphics*, page 102027, 2021. 3

[4] Deepak Anand, Shrey Gadiya, and Amit Sethi. Histographs: graphs in histopathology. In *Medical Imaging 2020: Digital Pathology*, volume 11320, pages 150–155. SPIE, 2020. 3, 4, 5

[5] Filippo Maria Bianchi, Daniele Grattarola, Lorenzo Livi, and Cesare Alippi. Graph neural networks with convolutional arma filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. 6, 7

[6] Shaked Brody, Uri Alon, and Eran Yahav. How attentive are graph attention networks? In *International Conference on Learning Representations*, 2022. 6, 7, 8

[7] Gabriele Campanella, Matthew G Hanna, Luke Geneslaw, Allen Miraflor, Vitor Werneck Krauss Silva, Klaus J Busam, Edi Brogi, Victor E Reuter, David S Klimstra, and Thomas J Fuchs. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature medicine*, 25(8):1301–1309, 2019. 3

[8] Marc-André Carbonneau, Veronika Cheplygina, Eric Granger, and Ghyslain Gagnon. Multiple instance learning: A survey of problem characteristics and applications. *Pattern Recognition*, 77:329–353, 2018. 3

[9] Tsai Hor Chan, Fernando Julio Cendra, Lan Ma, Guosheng Yin, and Lequan Yu. Histopathology whole slide image analysis with heterogeneous graph representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15661–15670, 2023. 3, 6, 7

[10] Matthew C Cieslak, Ann M Castelfranco, Vittoria Roncalli, Petra H Lenz, and Daniel K Hartline. t-distributed stochastic neighbor embedding (t-sne): A tool for eco-physiological transcriptomic analysis. *Marine genomics*, 51:100723, 2020. 7

[11] Kexin Ding, Qiao Liu, Edward Lee, Mu Zhou, Aidong Lu, and Shaoting Zhang. Feature-enhanced graph networks for genetic mutational prediction using histopathological images in colon cancer. In *MICCAI*, pages 294–304. Springer, 2020. 2, 3

[12] Saisai Ding, Zhiyang Gao, Jun Wang, Minhua Lu, and Jun Shi. Fractal graph convolutional network with mlp-mixer based multi-path feature fusion for classification of histopathological images. *Expert Systems with Applications*, 212:118793, 2023. 2, 3

[13] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Dehghani, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020. 4

[14] Vijay Prakash Dwivedi, Chaitanya K Joshi, Thomas Laurent, Yoshua Bengio, and Xavier Bresson. Benchmarking graph neural networks. *arXiv preprint arXiv:2003.00982*, 2020. 5

[15] Adrian Galdran, Jose Dolz, Hadi Chakor, Hervé Lombaert, and Ismail Ben Ayed. Cost-sensitive regularization for diabetic retinopathy grading from eye fundus images. In *MICCAI*, pages 665–674. Springer, 2020. 6

[16] Hongyang Gao and Shuiwang Ji. Graph u-nets. In *ICML*, pages 2083–2092. PMLR, 2019. 5

[17] Zhiyang Gao, Zhiyang Lu, Jun Wang, Shihui Ying, and Jun Shi. A convolutional neural network and graph convolutional network based framework for classification of breast histopathological images. *IEEE Journal of Biomedical and Health Informatics*, 26(7):3163–3173, 2022. 2, 3

[18] Frank Harary and Robert Z Norman. Some properties of line digraphs. *Rendiconti del circolo matematico di palermo*, 9(2):161–168, 1960. 2

[19] Ziwang Huang, Hua Chai, Ruoqi Wang, Haitao Wang, Yuedong Yang, and Hejun Wu. 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*, pages 561–570. Springer, 2021. 5

[20] Maximilian Ilse, Jakub Tomczak, and Max Welling. Attention-based deep multiple instance learning. In *International conference on machine learning*, pages 2127–2136. PMLR, 2018. 3

[21] Maximilian Ilse, Jakub M Tomczak, and Max Welling. Deep multiple instance learning for digital histopathology. In *Handbook of Medical Image Computing and Computer Assisted Intervention*, pages 521–546. Elsevier, 2020. 3

[22] Stephen J Keenan, James Diamond, W Glenn McCluggage, Hoshang Bharucha, Deborah Thompson, Peter H Bartels, and Peter W Hamilton. An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (cin). *The Journal of pathology*, 192(3):351–362, 2000. 4, 7

[23] K Krishna and M Narasimha Murty. Genetic k-means algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 29(3):433–439, 1999. 6

[24] Guohao Li, Chenxin Xiong, Ali Thabet, and Bernard Ghanem. Deepergcn: All you need to train deeper gcns. *arXiv preprint arXiv:2006.07739*, 2020. 8

[25] Honglin Li, Chenglu Zhu, Yunlong Zhang, Yuxuan Sun, Zhongyi Shui, Wenwei Kuang, Sunyi Zheng, and Lin Yang. Task-specific fine-tuning via variational information bottleneck for weakly-supervised pathology whole slide image classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7454–7463, 2023. 2

[26] Huapeng Lin, Xinrong Zhang, Guanlin Li, Grace Lai-Hung Wong, and Vincent Wai-Sun Wong. Epidemiology and clinical outcomes of metabolic (dysfunction)-associated fatty

liver disease. *Journal of Clinical and Translational Hepatology*, 9(6):972, 2021. 1

[27] Tiancheng Lin, Zhimiao Yu, Hongyu Hu, Yi Xu, and Chang-Wen Chen. Interventional bag multi-instance learning on whole-slide pathological images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19830–19839, 2023. 3, 6, 7

[28] Wei Lou, Xiang Wan, Guanbin Li, Xiaoying Lou, Chenghang Li, Feng Gao, and Haofeng Li. Structure embedded nucleus classification for histopathology images. *arXiv preprint arXiv:2302.11416*, 2023. 5

[29] Ming Y Lu, Tiffany Y Chen, Drew FK Williamson, Melissa Zhao, Maha Shady, Jana Lipkova, and Faisal Mahmood. Ai-based pathology predicts origins for cancers of unknown primary. *Nature*, 594(7861):106–110, 2021. 3

[30] Wenqi Lu, Simon Graham, Mohsin Bilal, Nasir Rajpoot, and Fayyaz Minhas. Capturing cellular topology in multi-gigapixel pathology images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 260–261, 2020. 3

[31] Puneet Puri and Arun J. Sanyal. Nonalcoholic fatty liver disease: Definitions, risk factors, and workup. *Clinical Liver Disease*, 1, 2012. 1, 2, 3

[32] Linhao Qu, Zhiwei Yang, Minghong Duan, Yingfan Ma, Shuo Wang, Manning Wang, and Zhijian Song. Boosting whole slide image classification from the perspectives of distribution, correlation and magnification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 21463–21473, 2023. 2

[33] Mousumi Roy, Fusheng Wang, Hoang Vo, Dejun Teng, George Teodoro, Alton B Farris, Eduardo Castillo-Leon, Miriam B Vos, and Jun Kong. Deep-learning-based accurate hepatic steatosis quantification for histological assessment of liver biopsies. *Laboratory Investigation*, 100(10):1367–1383, 2020. 6

[34] Benedek Rozemberczki, Peter Englert, Amol Kapoor, Martin Blais, and Bryan Perozzi. Pathfinder discovery networks for neural message passing. In *Proceedings of the Web Conference 2021*, pages 2547–2558, 2021. 8

[35] Yunsheng Shi, Zhengjie Huang, Shikun Feng, Hui Zhong, Wenjin Wang, and Yu Sun. Masked label prediction: Unified message passing model for semi-supervised classification. *IJCAI*, 2021. 8

[36] Mookund Sureka, Abhijeet Patil, Deepak Anand, and Amit Sethi. Visualization for histopathology images using graph convolutional neural networks. In *BIBE*, pages 331–335. IEEE, 2020. 3

[37] Yoshihisa Takahashi and Toshio Fukusato. Histopathology of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *World journal of gastroenterology: WJG*, 20(42):15539, 2014. 2

[38] William Thomas Tutte and William Thomas Tutte. *Graph theory*, volume 21. Cambridge university press, 2001. 4

[39] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008. 7

[40] Marta Wojciechowska, Stefano Malacrino, Natalia Garcia Martin, Hamid Fehri, and Jens Rittscher. Early detection of liver fibrosis using graph convolutional networks. In *MICCAI*, pages 217–226. Springer, 2021. 2, 6, 7

[41] Geming Wu, Xinyan Zhao, Shuqian Luo, and Hongli Shi. Histological image segmentation using fast mean shift clustering method. *Biomedical engineering online*, 14(1):1–12, 2015. 3

[42] Jiawei Yang, Hanbo Chen, Yu Zhao, Fan Yang, Yao Zhang, Lei He, and Jianhua Yao. Remix: A general and efficient framework for multiple instance learning based whole slide image classification. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part II*, pages 35–45. Springer, 2022. 3, 6, 7

[43] Koichiro Yasaka, Hiroyuki Akai, Akira Kunimatsu, Osamu Abe, and Shigeru Kiryu. Deep learning for staging liver fibrosis on ct: a pilot study. *European radiology*, 28:4578–4585, 2018. 6

[44] Koichiro Yasaka, Hiroyuki Akai, Akira Kunimatsu, Osamu Abe, and Shigeru Kiryu. Liver fibrosis: deep convolutional neural network for staging by using gadoxetic acid–enhanced hepatobiliary phase mr images. *Radiology*, 287(1):146–155, 2018. 6

[45] Chong Yin, Siqi Liu, Rui Shao, and Pong C Yuen. Focusing on clinically interpretable features: selective attention regularization for liver biopsy image classification. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24*, pages 153–162. Springer, 2021. 2

[46] Chong Yin, Siqi Liu, Vincent Wai-Sun Wong, and Pong C Yuen. Learning sparse interpretable features for nas scoring from liver biopsy images. In *IJCAI*, pages 1580–1586, 2022. 2

[47] Marina Yusoff, Toto Haryanto, Heru Suhartanto, Wan Azani Mustafa, Jasni Mohamad Zain, and Kusmardi Kusmardi. Accuracy analysis of deep learning methods in breast cancer classification: A structured review. *Diagnostics*, 13(4):683, 2023. 2

[48] Hongrun Zhang, Yanda Meng, Yitian Zhao, Yihong Qiao, Xiaoyun Yang, Sarah E Coupland, and Yalin Zheng. Dtfd-mil: Double-tier feature distillation multiple instance learning for histopathology whole slide image classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18802–18812, 2022. 6, 7

[49] Yu-Jie Zhou, Feng Gao, Wen-Yue Liu, Grace Lai-Hung Wong, Sanjiv Mahadeva, Nik Raihan Nik Mustapha, Xiao-Dong Wang, Wah-Kheong Chan, Vincent Wai-Sun Wong, and Ming-Hua Zheng. Screening for compensated advanced chronic liver disease using refined baveno vi elastography cutoffs in asian patients with nonalcoholic fatty liver disease. *Alimentary pharmacology & therapeutics*, 54(4):470–480, 2021. 6