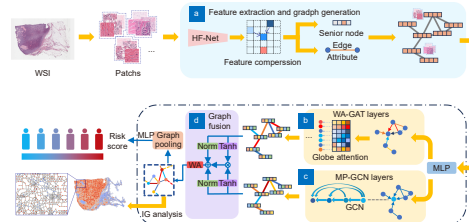


DOI: 10.12086/oe.2024.240011

# 基于图神经网络的 WSI 癌症生存预测方法

叶世杰, 王永雄\*

上海理工大学光电信息与计算机工程学院, 上海 200093



**摘要:** 全切片图像 (Whole slide imaging, WSI) 是癌症诊断和预后的关键依据, 具有尺寸庞大、空间关系复杂以及风格各异等特点。由于其缺乏细节注释, 传统的计算病理学方法难以处理肿瘤组织环境中的空间关系。本文提出了一种新型的基于图神经网络的 WSI 生存预测模型 BC-GraphSurv。首先, 采用迁移学习的预训练策略, 构建 WSI 的病理关系拓扑结构, 实现了对病理学图像特征和空间关系信息的有效提取。然后, 采用 GAT-GCN 双分支结构进行预测, 在图注意力网络中加入边属性和全局连接模块, 同时引入图卷积网络分支补充局部细节, 增强了对 WSI 风格差异的适应能力, 能够有效利用拓扑结构处理空间关系, 区分微病理环境。在 WSI 数据集 TCGA-BRCA 和 TCGA-KIRC 上进行的实验表明, BC-GraphSurv 模型的一致性指数为 0.7950 和 0.7458, 相比于当前先进的生存预测模型提升了 0.0409, 充分证明了模型的有效性。

**关键词:** 生存预测; 图神经网络; 迁移学习; 注意力机制

**中图分类号:** TP18

**文献标志码:** A

叶世杰, 王永雄. 基于图神经网络的 WSI 癌症生存预测方法 [J]. 光电工程, 2024, 51(4): 240011

Ye S J, Wang Y X. Graph neural network-based WSI cancer survival prediction method[J]. *Opto-Electron Eng*, 2024, 51(4): 240011

## Graph neural network-based WSI cancer survival prediction method

Ye Shijie, Wang Yongxiong\*

Institute of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

**Abstract:** Whole slide imaging (WSI) is the main basis for cancer diagnosis and prognosis, characterized by its large size, complex spatial relationships, and diverse styles. Due to its lack of detailed annotations, traditional computational pathology methods are difficult to handle WSI tasks. To address these challenges, this paper proposes a WSI survival prediction model based on graph neural networks, BC-GraphSurv. Specifically, we use transfer learning pre-training to extract features containing spatial relationship information and construct the pathological relationship topology of WSI. Then, the two branch structures of the improved graph attention network (GAT) and graph convolutional network (GCN) are used to predict the extracted features. We combine edge attributes and global perception modules in GAT, while the GCN branch is used to supplement local details, which can achieve adaptability to WSI style differences and effectively utilize topological structures to handle spatial

收稿日期: 2024-01-09; 修回日期: 2024-03-12; 录用日期: 2024-03-12

基金项目: 上海市自然科学基金项目 (22ZR1443700)

\*通信作者: 王永雄, wyxiong@usst.edu.cn。

版权所有©2024 中国科学院光电技术研究所

relationships and distinguish subtle pathological environments. Experimental results on the TCGA-BRCA dataset demonstrate BC-GraphSurv's effectiveness, achieving a C-index of 0.795—a significant improvement of 0.0409 compared to current state-of-the-art survival prediction models. This underscores its robust efficacy in addressing WSI challenges in cancer diagnosis and prognosis.

**Keywords:** survival prediction; graph neural network; transfer learning; attention mechanism

## 1 引言

随着数字病理学近年来的高速发展, 数字化高分辨率的组织病理学全切片图像 (whole slide imaging, WSI)<sup>[1]</sup> 为肿瘤切片提供了充足的影像资料, 已经成为癌症诊断和预后的黄金标准, 基于千兆像素分辨率的 WSI 的生存预测分析是医学图像分析中具有挑战性的任务。该任务旨在对生存持续时间进行建模, 该时间从采样图像到发生特定事件 (例如生物死亡) 为止。由于 WSI 通常缺少密集注释, 使用弱监督深度学习模型进行自动分析是一种合适的方式。

近年来, 深度学习技术在病理学图像处理领域取得了巨大的成功, 在图像分类<sup>[2-3]</sup>、分割<sup>[4-5]</sup>和压缩<sup>[6]</sup>等多个领域应用广泛。然而在 WSI 生存预测方面, 基于深度学习模型的方法主要有以下三个难点: 1) 由于计算资源有限, 使用整个 WSI (典型大小为 60000 像素×60000 像素) 进行模型训练难以实现。因此常常使用采样图像块进行分析的方式获得最终结果, 在没有注释的情况下提取图像块的病理特征是一个有挑战性的问题。2) 与传统的图像任务不同, WSI 包含丰富的细胞和组织结构信息, 空间关系的处理对于生存预测任务至关重要, 然而传统模型受限于感受野, 往往忽略其中的相对关系。3) WSI 的质量和风格因图像采集的设备和方法不同产生异质性。WSI 可能会因染色方式或保存环境的不同而产生风格差异, 且一个患者可能包含不同风格的切片, 对模型的判断有很大影响。

早期的方法<sup>[7]</sup> 广泛采用由病理学家筛选图像块的方式, 然而人工筛选耗时耗力, 因此 Zhu<sup>[8]</sup> 等人提出了随机采样图像块的 WSISA 模型, 并使用卷积神经网络 (convolutional neural network, CNN) 提取特征, 解决了人工选择的限制, 但随机采样往往损失较多关键信息。Yao<sup>[9]</sup> 等人通过两阶段多实例学习框架, 聚类筛选图像块, 并使用 ImageNet 预训练的 VGG16 提取特征, 缓解了关键信息丢失的问题。然而, 自然图像和病理图像之间存在数据分布差异<sup>[10]</sup>, 经过自然图

像预训练的通用骨干网络难以适应病理图像。鉴于医学图像分类任务中迁移学习方法的有效性<sup>[11]</sup>, 本文考虑构建特征迁移模块与 CNN 骨干网络结合, 通过病理学图像迁移学习提升网络的病理特征提取能力, 并基于提取的病理特征构建完整 WSI 的拓扑结构, 可以有效减少计算资源的消耗, 同时保留完整的病理上下文信息。

WSI 中的组织结构往往是非规则的, CNN 善于处理网格状数据, 然而在捕捉空间关系时受限于固定的局部感受野, 无法有效地处理图像中的长距离关系或全局结构。为了更好地感知病理环境, 一些方法<sup>[12-14]</sup> 采用图神经网络 (graph neural network, GNN)。Li<sup>[12]</sup> 等人通过 VGG 主干网络提取采样图像块的特征, 将 WSI 进行图结构建模, 提出了基于 GNN 的 DeepGraphSurv 模型。Chen<sup>[13]</sup> 等人提出 Patch-GCN 模型, 使用 Resnet 提取并构建图结构特征, 最后使用 GNN 预测生存数据。GNN 通过拓扑结构传递信息, 有助于区分不同的组织结构, 但其处理节点之间的空间关系时存在局限, 尤其在肿瘤病理环境中, 节点邻域包含许多异质病理区域。为解决这一问题, 本文引入改进的图注意力网络 (graph attention network, GAT)<sup>[15]</sup>。GAT 是一种 GNN 变体, 在聚合中引入注意力机制赋予不同节点重要性, 更灵活地捕捉邻域节点间的关系, 本文改进 GAT 的聚合过程, 加入边的属性以提供节点的相对空间关系, 使节点所处空间得以纳入考虑范围, 可以更好地理解节点之间的位置关系, 获得对肿瘤微病理环境的细致感知。

此外, WSI 通常有不同的染色和保存方式, 如福尔马林固定石蜡包埋和新鲜冷冻组织, 不同的方式会导致 WSI 的风格差异, 放大不同图中节点特征差异<sup>[16]</sup>。GNN 邻域聚合在单层的感受范围有限, 容易受到整体风格的影响, 难以充分感知不同 WSI 中的病理环境。为了解决这一问题, Di<sup>[14]</sup> 等人采用超图结构将 WSI 分解成大面积的语义空间, 考虑更大范围的节点关系, 然而超图虽然拥有更大的感受范围, 但是仍然没有引入全局感知, 缺少对 WSI 整体风格

的把握能力。为了解决这一局限性, 本文在每层 GAT 中加入基于注意力的全局连接模块, 引入 WSI 全局特征的学习, 形成节点、边、全局特征相结合的全面感知, 有助于消除风格差异带来的影响。同时, 为了缓解多层 GAT 产生的过平滑问题, 提高模型的适应能力, 在预测阶段引入基于图卷积网络 (graph convolutional network, GCN)<sup>[17]</sup> 的分支, GCN 通过局部拓扑结构进行信息传递, 可以更好地捕捉局部结构特征, 通过引入额外的图卷积操作, 可以增强对细节病理环境的建模能力, 使网络获取细节信息的补充, 提升模型的鲁棒性。

基于 WSI 生存预测的难点, 本文提出了一种基于图神经网络的乳腺癌生存预测模型 BC-GraphSurv。创新包括:

1) 提出结合了全域连接图注意力网络 (whole association graph attention network, WA-GAT) 和 GCN 的生存预测网络, WA-GAT 融合了节点、边和全局特征的综合注意力感知, 辅以 GCN 分支强化细节特征, 可以有效处理空间信息, 提高了模型的分析能力。

2) 采用了基于 CNN 和自注意力机制的特征提取网络 HF-Net (histopathology transfer network), 使用迁移学习的预训练策略, 使其通过特征迁移模块获得病理学特征识别能力, 并对 WSI 进行图结构压缩建模, 提取了大部分上下文特征。

3) 通过消融实验验证不同模块和改进的作用, 并

通过不同模型的对比实验和可视化分析验证该方法的有效性。

## 2 算法理论推导

提出的生存预测框架主要由四个步骤组成, 图 1 展示了整个工作流程。1) 使用乳腺癌病理图像迁移学习病理学迁移网络 (HF-Net), 从预处理的 WSI 组织区域中提取图像块的特征。2) 将相似图像块特征进行压缩融合成为高级节点, 并对每个 WSI 构建对应的图结构特征。3) 使用全域连接图注意力网络 (WA-GAT), 和多重感知图卷积网络 (multi-perception Graph convolutional network, MP-GCN) 对图结构特征进行学习, 捕获不同组织区域之间的相关性以及不同特征之间的重要性。4) 融合两个分支的特征并获得最终的生存预测结果。

### 2.1 病理学迁移网络 (HF-Net)

由于肿瘤结构和组织纹理特征的复杂性, 使用自然事物图像数据集预训练的主干网络可能不能很好地适应病理图像任务, 迁移学习是使通用骨干网络适应医学图像任务的有效方法<sup>[11]</sup>。为了减少非癌区引起的噪声以及加强对不同癌区的差异感知, 本文基于现有的乳腺癌肿瘤分类数据集训练病理学迁移网络 HF-Net 用于区分不同的组织结构并提取图像块的特征。

HF-Net 由主干网络和基于自注意力机制的特征迁移模块两部分组成, 其主要结构如图 2 所示。首先

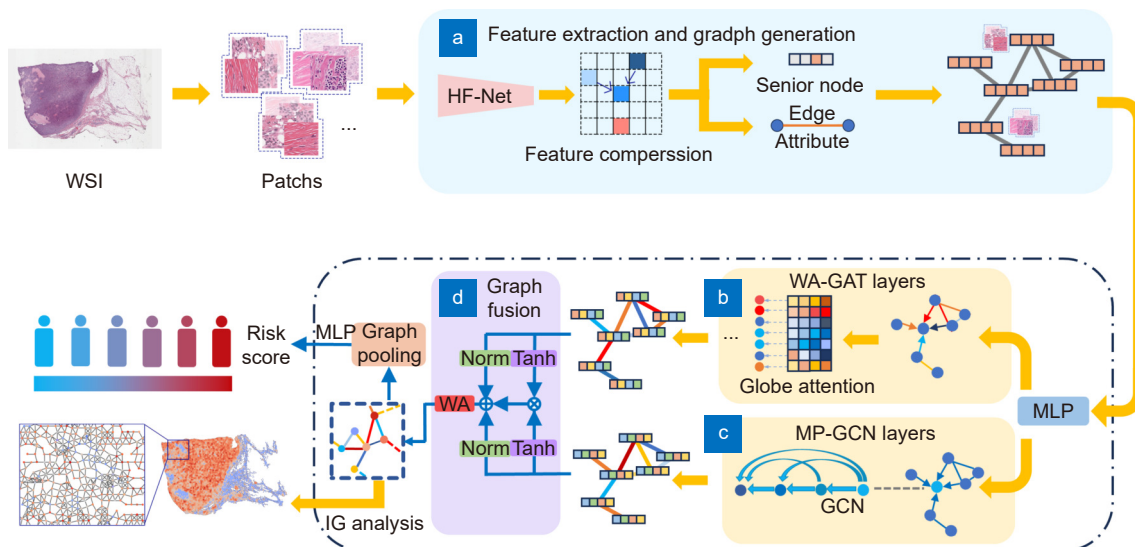


图 1 BC-GraphSurv 模型架构, 主要包括 WSI 预处理, 以及 (a) 特征提取和生成图结构, (b) WA-GAT 分支, (c) MP-GCN 分支、(d) 特征融合等模块

Fig. 1 Architecture of the BC GraphSurv model mainly including modules such as WSI preprocessing, (a) Feature extraction and graph structure generation, (b) WA-GAT branch, (c) MP-GCN branches, and (d) Feature fusion

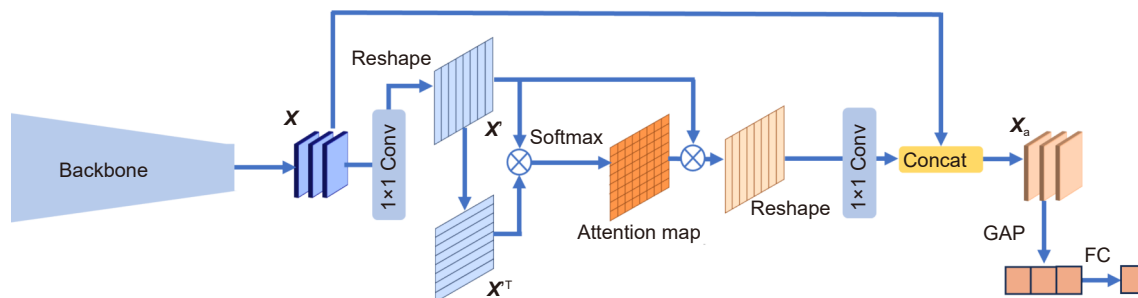


图 2 HF-Net 示意图

Fig. 2 Schematic diagram of HF-Net

将在 ImageNet 预训练的 Resnet50 去除最后的全局平均池化层和全连接层, 并冻结其余层参数, 然后在最后一个残差块输出特征图后, 连接特征迁移模块进行迁移学习并生成对应图像块的特征向量。具体为, Resnet50 的最后一个残差块输出的特征图会先经过一个卷积层调整通道, 再将通道维度的特征图展平, 并与转置的展平特征图做矩阵乘法并经过一个 Softmax 函数得到特征迁移注意力图, 与特征图相乘获得新的特征图, 接着经过通道维度重塑和一个卷积层恢复成原来的形状, 并与原始特征图相加得到迁移后的特征图。具体计算过程为

$$X_a = f(\text{Softmax}(X'X^T)X') + X, \quad (1)$$

其中:  $X$  代表残差块输出的原始特征图,  $X'$  代表经过降维和重塑的特征图,  $X_a$  代表经过特征迁移后的特征图,  $f$  代表通道维度展平或重塑。将获得的特征图进行全局平均池化, 再通过全连接层从而得到分类结果。将全局平均池化层的特征向量与分类层结果进行拼接, 得到对应图像块的特征向量表示。具体计算过程为

$$X_{\text{patch}} = \text{GAP}(X_a) \parallel X_{\text{class}}, \quad (2)$$

其中: GAP 代表全局平均池化处理,  $X_{\text{class}}$  代表全连接层输出的分类结果,  $X_{\text{patch}}$  代表图像块的特征向量。

通过在乳腺癌肿瘤病理学图像分类数据集进行分类预训练, HF-Net 提取的特征可以有效地区分不同的组织成分, 获得对不同组织结构之间差异的感知并提取有代表性的特征, 减少了冗余信息, 提高了图结构压缩特征的质量, 从而提升后续任务的感知效果。

## 2.2 生成图结构特征

由于 WSIs 通常为千亿像素级分辨率且病理组织在相同区域的相似度较高, 提取的特征数量庞大且重复度高, 需要更具有代表性的特征作为高级节点特征并生成 WSI 级的图结构特征, 生成图结构的过程如

图 1(a) 所示。在获取节点特征时应该同时考虑空间上下文关系和特征的内在联系, 因此本文以特征之间的空间距离和语义相似度为依据, 使用  $5 \times 5$  范围的滑动窗口在 WSI 范围滑动, 范围内的图像块与中心图像块的余弦相似度高于设定值, 则认为这个图像块为中心图像块的相似图像块。然后根据每个图像块的相似图像块数对图像块进行排序, 从最相似图像块最多的图像块开始, 融合属于该图像块的相似图像块, 其所有相似图像块的特征平均, 得到高级特征, 将其作为高级图像块。生成高级图像块后, 每个高级图像块被视为图结构上的一个高级节点, 然后根据高级节点之间的空间距离生成边, 在空间距离小于 5 个图像块长度的高级节点之间构建边并存储位置信息。

通过 HF-Net 提取的病理特征, 可以有效地融合病理信息获得高级节点, 生成的 WSI 级图结构特征在有效保留了原本组织结构特征的情况下, 对数据量进行了压缩, 以便后续的计算和存储, 其中边的属性为相邻高节点之间互相的位置信息, 以便提供更细致的邻域感知。

## 2.3 全域连接图注意力网络 (WA-GAT)

在共享的 3 层 MLP 进行特征预处理, 将特征降维到 200 维, 然后将图特征送入 WA-GAT 分支, 该分支使用改进的 GAT 作为主干对特征进行组织病理学的上下文特征学习, GAT 是非谱方法的图神经网络, 具有可移植性, 适合处理不固定结构的有向图的问题。GAT 在邻域聚合过程中学习每条边的注意力得分, 区分不同节点的权重, 学习不同节点的重要性, 从而关注最具有影响能力的输入。但是 GAT 的聚合筛选方式单一, 对上下文的感知力有限, 并且基于邻域的聚合可能会过度关注局部, 而忽视 WSI 的整体异质性。因此本文基于 GAT 提出了一种新的 WA-GAT 网络, 在每层聚合中加入了邻域的特征和全局

特征的感知学习, 如下图 3 所示。具体来说, WA-GAT 每层的特征聚合将边代表的相对位置纳入考虑范围, 将相邻节点间的距离和角度用一维的向量表示, 并通过可学习参数生成距离和角度的特征, 以获得节点间相对空间位置的感知, 并在聚合后的特征上经过全局尺度特征的注意力进行强化, 使用全局特征对每个节点的特征重新加权, 以获得对病理组织全域的整体环境的学习, 强化特征的表达能力。

WA-GAT 每层输入的初始特征为所有高级节点的特征向量, 表示为

$$\beta = \{\beta_1, \beta_2, \dots, \beta_N\}, \beta_i \in \mathbb{R}^{D_n}, \quad (3)$$

其中:  $N$  为高级节点的数量,  $D_n$  为特征经过 MLP 预处理后的维数。除此之外, 在相邻的节点之间距离  $d$  和角度  $a$  作为边的特征一起输入网络。在特征输入到 WA-GAT 后, 为了学习上下文病理特征中周围环境的异质性, 每层网络根据相邻节点之间病理特征和距离特征, 计算注意力互相关系数, 在节点之间获得不同的信息传递权重。式 (4) 中  $d_{ij}$  和  $a_{ij}$  分别代表节点  $i$  和节点  $j$  间的距离和角度特征。  $W_n, W_l$  分别是节点和位置特征在该层共享的线性变换参数。  $\alpha$  代表一个单层前馈神经网络, 用于计算源节点、接收节点、距离和角度特征的注意力系数。  $N_i$  是图中高级节点  $i$  的邻居节点。

$$\varepsilon_{ij} = \frac{\exp(\text{LeakyReLU}(\alpha^T(W_n\beta_i \parallel W_n\beta_j \parallel W_l d_{ij} \parallel W_l a_{ij})))}{\sum_{k \in N_i} \exp(\text{LeakyReLU}(\alpha^T(W_n\beta_i \parallel W_n\beta_k \parallel W_l d_{ik} \parallel W_l a_{ik})))}. \quad (4)$$

得到的归一化的注意力系数将被用于计算相邻特征的线性聚合, 得到基于邻域的上下文病理特征  $\beta'$ , 其中第  $i$  个节点的特征  $\beta'_i$  表示为

$$\beta'_i = \sum_{j \in N_i} (\varepsilon_{ij} W \beta_j). \quad (5)$$

为了获取 WSI 的全域感知, 在全局特征中寻求值得关注的节点特征, 本文提出全域连接模块 (whole

association, WA) 对 GAT 进行改进, 通过全局特征计算每个高级节点的全局注意力系数, 对每个高级节点的特征进行动态调整以适应 WSI 的异质性, 并使用 PReLU 作为非线性激活函数, 使特征得到更充分的表达, 在相邻的每层之间进行残差连接, 得到每层最后的特征输出。计算方式为

$$G_i = \sigma(W_1\beta'_i + b_1) \odot (W_2\beta'_i + b_2), \quad (6)$$

$$\beta_i^{(\text{out})} = \text{PReLU} \left( \frac{\exp(W_g G_i)}{\sum_{k \in N} \exp(W_g G_k)} \beta'_i \right) + \beta_i, \quad (7)$$

其中:  $W_1, W_2, W_g, b_1, b_2$  为可学习参数;  $G$  代表全局特征;  $G_i$  为全局特征的第  $i$  个高级节点特征;  $\sigma$  表示 Sigmoid 激活函数;  $\odot$  表示哈达玛积;  $\beta_i^{(\text{out})}$  为 WA-GAT 的输出特征  $\beta^{(\text{out})}$  的第  $i$  个高级节点的特征。

在这个分支中使用了 3 层 WA-GAT 作为主干进行上下文病理特征的学习, 最终的输出  $Y_1$  为每层 WA-GAT 的输出特征以及原始特征的拼接, 表示为

$$Y_1 = \beta^{(0)} \parallel \beta^{(1)} \parallel \beta^{(2)} \parallel \beta^{(3)}, Y_1 \in \mathbb{R}^{4D_n}. \quad (8)$$

## 2.4 多重感知图卷积网络 (MP-GCN)

为了捕捉更复杂、抽象的多尺度特征, 捕捉局部结构细节, 该分支使用深度为 3 的稠密连接的 GCN 作为主干, GCN 通过局部结构进行信息传递, 可以有效地获取不同组织区域的细节。本文提出基于 GCN 进行结构改进的 MP-GCN, 在不增加深度的情况下最大化网络的信息通路, 形成分层的混合感知, 使网络能够更好地捕捉病理组织之间的关系以及不同层次上的特征, 可以对不同的组织区域精准感知, 为模型补充细节信息, 有效提升模型在大规模图像上的病理组织结构分析的性能和泛化性, 其过程如图 1(c) 所示。

分支网络中的每层 GCN 之间采用邻域聚合的图卷积进行消息传递, 初始的输入特征为

$$\gamma = \{\gamma_1, \gamma_2, \dots, \gamma_N\}, \gamma_i \in \mathbb{R}^{D_n}. \quad (9)$$

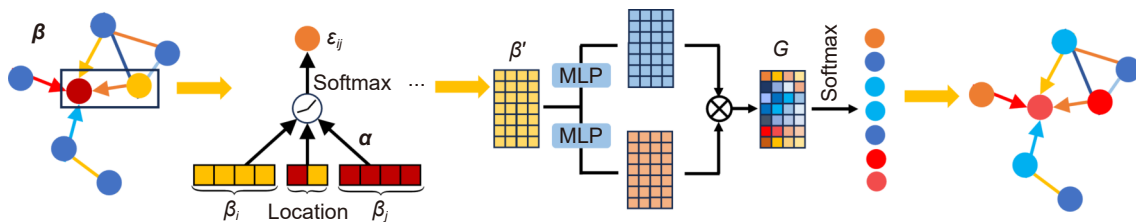


图 3 WA-GAT 示意图

Fig. 3 Schematic diagram of WA-GAT

图卷积过程中会根据不同节点连接度进行动态的特征传递, 同时考虑高级节点自身的特征, 当输入为特征矩阵和邻接矩阵时, 图卷积的过程可以表示为

$$F^{\text{GCN}}(\gamma^{(L)}, \mathbf{A}, \boldsymbol{\theta}) = \text{PReLU}\left(\overline{\mathbf{D}}^{-\frac{1}{2}} \overline{\mathbf{A}} \overline{\mathbf{D}}^{-\frac{1}{2}} \gamma^{(L)} \boldsymbol{\theta}\right), \quad (10)$$

$$\overline{\mathbf{A}} = \mathbf{A} + \mathbf{I}, \quad (11)$$

其中:  $\gamma^{(L)}$ 代表第  $L$  层 GCN 输出的特征;  $\overline{\mathbf{A}}$ 为加入了自连接的邻接矩阵;  $\overline{\mathbf{D}}$ 代表  $\overline{\mathbf{A}}$ 的连接度矩阵;  $\boldsymbol{\theta}$ 为该层 GCN 的可学习参数。在该分支中, 各层之间稠密连接, 即每一个 MP-GCN 输出的特征为之前所有的输出特征和原始输入特征在进行 GCN 操作之后的和, 包含之前所有深度的病理特征, 且每两个输出之间的 GCN 层相互独立。其前向传播过程中的第  $i$  个输出特征可以表示为

$$\gamma^{(i)} = \sum_{k < i} (F^{\text{GCN}}(\gamma^{(k)}, \mathbf{A}, \boldsymbol{\theta}^{(ki)}) + F^{\text{GCN}}(\gamma^{(0)}, \mathbf{A}, \boldsymbol{\theta}^{(0i)})), \quad (12)$$

其中:  $\gamma^{(0)}$ 表示初始特征,  $\gamma^{(i)}$ 表示 MP-GCN 的第  $i$  个输出特征,  $\boldsymbol{\theta}^{(ki)}$ 表示连接第  $k$  个到第  $i$  个输出之间的 GCN 的可学习参数。该分支最终得到的输出为各阶段输出和原始输入特征的拼接, 表示为

$$\mathbf{Y}_2 = \gamma^{(0)} \parallel \gamma^{(1)} \parallel \gamma^{(2)} \parallel \gamma^{(3)}, \mathbf{Y}_2 \in \mathbb{R}^{4D_n}. \quad (13)$$

## 2.5 特征融合与生存预测

WA-GAT 分支通过邻域交叉注意力和全局特征的整体感知对组织结构进行全面的融合感受, MP-GCN 分支通过基于结构的聚合方式和不同深度信息交互获取不同层面的细节补充。双分支融合结构可以结合不同分支的信息, 提高模型的鲁棒性<sup>[18]</sup>, 使其对于组织图像中的噪声和变形更具有适应性, 有助于对抗数据中的变化, 在没有图像细节标注的生存预测任务中有助于提高模型的性能。

在特征融合阶段, 使用分组交叉融合的方法进一步对优越的特征进行选择, 减少特征中的冗余信息。具体由  $\tanh$  函数、最大最小值归一化以及乘法器和累加器组成, 其过程如图 1(d) 所示。两个分支得到的特征矩阵  $\mathbf{Y}_1$  和  $\mathbf{Y}_2$  先进行初步融合获得  $\mathbf{Y}'$ , 具体过程为

$$\mathbf{Y}' = \tanh(\mathbf{Y}_1) \odot \tanh(\mathbf{Y}_2), \quad (14)$$

其中:  $\odot$ 代表哈达玛积操作。多层 GNN 往往会产生平滑问题, 由于 GNN 中节点表示是通过聚合其邻居节点的信息更新, 多层聚合可能导致模型将相邻节点的信息过度平均化, 从而使得节点之间的差异变得模糊, 特别是存在明显的结构差异或节点的局部特

征差异较大时, 可能影响模型对节点的判断。在理想的状态下, WA-GAT 分支为主要分支, MP-GCN 分支作为辅助分支, 两者独立的学习类似结构的特征, 使用哈达玛积将两个分支对应的特征相乘, 可以使对应的特征得到一定程度的强化, 能够更好地保留节点之间的差异。接着  $\mathbf{Y}_1$  和  $\mathbf{Y}_2$  经过最大最小值归一化并与  $\mathbf{Y}'$  进行元素层面的累加, 经过一个 WA 模块进行特征优化, 获得最终的融合特征  $\mathbf{Y}_f$ , 再通过节点维度的全局平均池化和两层 800 和 400 维的 MLP 作为后处理, 得到输出特征  $\mathbf{y}_{\text{out}}$ , 其过程可以表示为

$$\mathbf{Y}_f = f^{\text{WA}}(\mathbf{Y}' + \varphi(\mathbf{Y}_1) + \varphi(\mathbf{Y}_2)), \quad (15)$$

$$\mathbf{y}_{\text{out}} = \text{MLP}(\text{GAP}(\mathbf{Y}_f)), \quad (16)$$

其中:  $\varphi$ 表示最大最小值归一化,  $f^{\text{WA}}$ 表示全局联系模块。通过将这两个分支网络结合在一个端到端的融合框架中, 可以更容易地进行训练和优化, 而不需要分开调整两种不同方法的参数, 有助于简化整个模型的设计和管理。

## 2.6 损失函数

在生存预测分析的任务中患者删失是普遍存在的情况, 患者删失表示截至最近的随访日期患者仍然存活, 针对删失数据建模不能使用传统的分类和回归损失, 否则会对网络可学习的数据量产生影响, 因此采用在生存预测分析领域常用的 cox 回归损失函数<sup>[19]</sup>来优化模型, 该损失函数通过最小化负对数似然来优化模型, 具体为

$$l(\boldsymbol{\eta}, \mathbf{y}_{\text{out}}) = - \sum_{i \in \mu} \delta_i (\boldsymbol{\eta}^T \mathbf{y}_{\text{out}}^{(i)} - \log \sum_{j \in \theta} e^{\boldsymbol{\eta}^T \mathbf{y}_{\text{out}}^{(j)}}), \quad (17)$$

其中:  $\boldsymbol{\eta} \in \mathbb{R}$ 是最后一层前馈神经网络的参数;  $\boldsymbol{\eta}^T \mathbf{y}_{\text{out}}^{(i)}$ 代表模型输出的第  $i$  个患者的预测风险;  $\delta_i$ 是患者删失的指示变量, 若患者删失则值为 0, 否则为 1;  $\mu$ 表示所有病人的集合,  $\theta$ 表示存活时间比患者  $i$  长的患者的集合。

## 3 实验结果与分析

### 3.1 数据集及数据预处理

在本研究中, 提出的 BC-GraphSurv 在癌症基因组图谱计划<sup>[20]</sup> (the cancer genome atlas, TCGA) 的乳腺癌数据集 (breast cancer, BRCA) 和肾透明细胞癌数据集 (kidney clear cell carcinoma, KIRC), 上进行了评估。BRCA 数据集是最大的乳腺癌 WSI 公开数据集, 包含来自 1062 位患者病例的 1133 幅 WSI 图像, 平

均大小为 1.05 GB, KIRC 数据集则包含来自 513 位患者病例的 519 幅 WSI 肾透明细胞癌图像, 平均大小为 0.85 GB。每幅 WSI 包括 4 倍、10 倍、20 倍和 40 倍四种放大倍率图层。

特征提取阶段 HF-Net 的预训练数据集为乳腺癌组织病理学数据库<sup>[21]</sup> (breast cancer histopathological database, BraKHis), 包含来自 82 位患者的 7909 幅 700 像素×460 像素大小的组织病理图像, 分为良性和恶性两个大类, 每个大类下包含 4 个小类共 8 个小类。

本文选择了 WSI 的 40 倍图层进行实验, 该倍率的组织结构和形态最清晰, 有利于深度神经网络的训练。WSI 中含有大量的空白玻璃背景, 为了排除背景部分, 使用 Otsu 算法<sup>[22]</sup> 对 WSI 的背景和组织进行分割处理, 然后将 40 放大倍率的图层切割成 256×256 像素的图像块, 并过滤掉包含组织像素小于 75% 的图像块。经过处理后的图像块不包括不必要的空白背景。

### 3.2 评价指标与模型训练

实验中采用一致性指数<sup>[23]</sup> (consistency index, C-index)。作为主要评价指标为 C-index 被广泛用于衡量逻辑回归模型中二元结果的拟合准确度。当有预测的风险值  $P = \{r_1^p, r_2^p, \dots, r_N^p\}$  和真实的风险值  $G = \{r_1^g, r_2^g, \dots, r_N^g\}$  时, C-index 的计算公式如下:

$$c = \frac{1}{M} \sum_{i:\delta_i=1} \sum_{j:r_i^g < r_j^g} \mathbb{I}[r_i^p < r_j^p], \quad (18)$$

其中:  $M$  为可比较的样本对的总数,  $\delta_i$  是患者删失的指示变量,  $\mathbb{I}[\cdot]$  表示指示函数。C-index 的取值范围为 0~1, 值越大意味着模型的预测性能越好, 其中 1 表示完全预测, 0.5 表示模型性能等同于随机预测, 小于 0.5 表示性能差于随机。

实验所用的硬件环境为: GPU 为 GeForce RTX 4060 Ti, 处理器为 Intel(R)Core i5-12600KF, 内存为 64 GB。软件环境为: CUDA 11.6+Python 3.8.7+Pytorch 1.11.2。训练共设置 100 个轮次, batch size 设为 16, 选用 AdamW 优化器<sup>[24]</sup>, OneCycleLR<sup>[25]</sup> 最大学习率设置为 0.0001。数据集的划分比例为训练集 (70%), 验证集 (10%), 测试集 (20%)。在实验时将数据集划分成 5 个子集, 并进行 5 折交叉验证, 每次实验将其中一个作为测试集, 其余划分为训练集和验证集, 使用不同的测试集分别进行 5 次实验, C-index

平均值和标准差 (SD) 作为最终结果进行比较。

### 3.3 消融实验

为了验证改进的模型有效性, 在 BRCA 数据集上进行实验。首先将 HF-Net 与 ResNet50 以及 EfficientNet-b5 作为特征提取阶段的骨干网络, 分别进行实验。然后在 BraKHis 数据集预训练的 HF-Net 作为特征提取网络的基础上, 增加 GAT、GCN、WA-GAT、MP-GCN 以及图特征融合 (graph fusion, GF) 模块进行预测算法的消融实验。实验结果对比如表 1 所示, 实验结果表明: HF-Net 相比于 EfficientNet-b5 以及 ResNet50 在 C-index 指标下分别提升了 0.0644 和 0.0725, 相比于 ImageNet 的通用 CNN 骨干网络相比, HF-Net 在没有增加参数量的同时取得了更好的结果, 证明了针对病理图像进行迁移学习可以有效地提升模型骨干网络的性能, 在病理图像任务中的性能优于通用网络。

表 1 特征提取网络实验结果对比

Table 1 Comparison of experimental results of feature extraction network

	C-index	SD	Parameters/MB
ResNet50	0.7225	0.011	97.49
EfficientNet-b5	0.7306	0.022	115.93
HF-Net	<b>0.7950</b>	0.013	113.72

消融实验结果指标对比如表 2 所示, 可以看出, 在 C-index 指标下, MP-GCN 和 WA-GAT 在 GCN 和 GAT 的基础上分别提升了 0.0595 和 0.0336, 稠密连接的层级结构可以提升 GCN 的多尺度特征表达能力, 增加边和全局特征可以有效提升 GAT 模型的肿瘤环境学习能力, 对预测模型的性能提升贡献较大。WAGAT+MPGCN 双分支结构相比于单独网络分别提升了 0.0068 和 0.0098, 增加融合模块 GF 后提升了 0.004, 说明双分支融合架构可以有效提升模型的预

表 2 消融实验结果对比

Table 2 Comparison of ablation experiment results

WA-GAT	MP-GCN	GAT	GCN	GF	C-index	SD
		√			0.7506	0.021
			√		0.7217	0.019
√					0.7842	0.007
	√				0.7812	0.008
√	√				0.7910	0.003
√	√			√	<b>0.7950</b>	0.013

测性能。整合了所有模块和分支的模型 C-index 达到 0.7950, 取得了最好的预测效果。

### 3.4 对比实验

为了验证本文提出的 BC-GraphSurv 模型在生存预测任务中的有效性, 本文选择了几种生存预测常用模型在 BRCA 数据集和 KIRC 数据集进行对比试验, 分别为 3 层的 MLP 基准网络、Attention-MIL<sup>[26]</sup>、WSISA<sup>[8]</sup>, 基于 GNN 的 DeepGraphSurv<sup>[12]</sup>、Patch GCN<sup>[13]</sup>、H<sup>2</sup>-MIL<sup>[27]</sup>、Tea-Graph<sup>[28]</sup>、HEAT<sup>[3]</sup>, 以及基于 Transformer 的 MIL-Transformer<sup>[29]</sup> 和 SurvTRACE<sup>[30]</sup>。所有模型均采用通过 BreakHis 数据集预训练的 HF-Net 进行 WSI 图像块的特征提取。

与基于 GNN 及 CNN 等算法进行对比, 实验结果如表 3 所示。由实验结果可以看出, 基于 GNN 的算法 (DeepGraphSurv<sup>[12]</sup>、Patch GCN<sup>[13]</sup>、H<sup>2</sup>-MIL<sup>[27]</sup>、Tea-Graph<sup>[28]</sup>、HEAT<sup>[3]</sup>、BC-GraphSurv) 普遍优于基于 CNN 的算法 (WSISA<sup>[8]</sup>) 和基于注意力的多实例学习算法 (Attention-MIL<sup>[26]</sup>), 表明利用图结构可以捕获 WSI 的拓扑信息, 提高生存预测的性能。此外, 引入注意力机制可以提升模型的学习能力, Attention-MIL<sup>[26]</sup> 基于自注意力机制进行多实例学习, C-index 相较 MLP 网络在两个数据集上提升了 0.0915 和 0.0728; 基于注意力的 GAT 预测模型 Tea-Graph<sup>[28]</sup> 的 C-index 相对于基于 GCN 的 DeepGraphSurv<sup>[12]</sup> 和 Patch GCN<sup>[13]</sup> 在 BRCA 数据集上分别高 0.0139 和 0.0027, 在 KIRC 数据集上分别高 0.0164 和 0.0334, 取得了较好的成绩; 基于异构图的算法 H<sup>2</sup>-MIL<sup>[27]</sup> 和 HEAT<sup>[3]</sup> 的 C-index 分别为 0.7338 和 0.7529。本文算法在 GAT 的基础上进一步扩大了注意力的范围和形

表 3 不同算法对比试验 1

Table 3 First comparison experiments of different methods

Method	BRCA		KIRC	
	C-index	SD	C-index	SD
MLP	0.6176	0.027	0.5862	0.019
Attention-MIL <sup>[26]</sup> (2018)	0.7091	0.052	0.6590	0.044
WSISA <sup>[8]</sup> (2017)	0.6802	0.083	0.6151	0.057
DeepGraphSurv <sup>[12]</sup> (2018)	0.7402	0.012	0.6945	0.045
Patch GCN <sup>[13]</sup> (2021)	0.7514	0.036	0.6775	0.067
H <sup>2</sup> -MIL <sup>[27]</sup> (2022)	0.7338	0.055	0.6915	0.024
Tea-Graph <sup>[28]</sup> (2022)	0.7541	0.021	0.7109	0.023
HEAT <sup>[3]</sup> (2023)	0.7529	0.009	0.7011	0.018
BC-GraphSurv	<b>0.7950</b>	0.013	<b>0.7458</b>	0.020

式, 并与 GCN 的细节信息补充融合, 相比最先进的算法最高提升了 0.0409, 在两个数据集分别取得了 0.7950 和 0.7458 的最高指标。

与基于 Transformer 的对比试验结果如表 4 所示。MIL-Transformer<sup>[29]</sup> 是基于 Transformer 对采样图像块进行多实例学习的生存预测算法, 在两个数据集上相对于 MLP 基准网络的 C-index 分别提高了 0.0729 和 0.078; SurvTRACE<sup>[30]</sup> 对特征表示中的多种混杂因素综合考虑, 基于 Transformer 进行特征的优化和预测, 在两个数据集上的 C-index 分别为 0.7382 和 0.6974, 本文算法相对最高值提升了 0.0568, 说明基于 GNN 的方法在处理大规模的复杂病理结构上具有优势。

表 4 不同算法对比试验 2

Table 4 Second comparison experiments of different methods

Method	BRCA		KIRC	
	C-index	SD	C-index	SD
MLP	0.6176	0.027	0.5862	0.019
MIL-Transformer <sup>[29]</sup> (2021)	0.6905	0.046	0.6642	0.029
SurvTRACE <sup>[30]</sup> (2022)	0.7382	0.019	0.6974	0.027
BC-GraphSurv	<b>0.7950</b>	0.013	<b>0.7458</b>	0.020

### 3.5 实验结果与分析

为了验证 BC-GraphSurv 的风险区分能力, 将 BC-GraphSurv 与几种常用的生存模型在 TCGA-BRCA 数据集上进行生存预测实验。根据模型所有的预测风险值大小将预测样本均分成高危和低危两组, 通过 log-rank 检验<sup>[31]</sup> 来验证不同模型的分层性能, 使用单变量 Kaplan-Meier (KM) 估计, 其高危组和低危组之间的间距越明显、P 值越小, 说明模型预测的高危组和低危组患者的生存概率差别越大, 对患者风险区分的性能越好。实验结果如图 4 所示。

可以看出, BC-GraphSurv 的 KM 曲线中高危组和低危组有最显著的区分度, 以及最小的 P 值, 能清楚的区分患者的风险等级, 相比于其他的算法具有较高的生存预测性能, 证明该模型在针对大量患者的风险排序上性能优越。

此外, 在病理学分析中, 能否区分出肿瘤的病理环境的异质性是风险值判断的基础, 模型的结果通常用于辅助临床决策, 如疾病诊断、治疗规划等。医生需要了解模型如何得出结论, 以便更好地理解并接受模型的建议。因此本文通过高级节点的 IG 值分析可视化每个高级节点所代表区域的病理环境, 使用可解



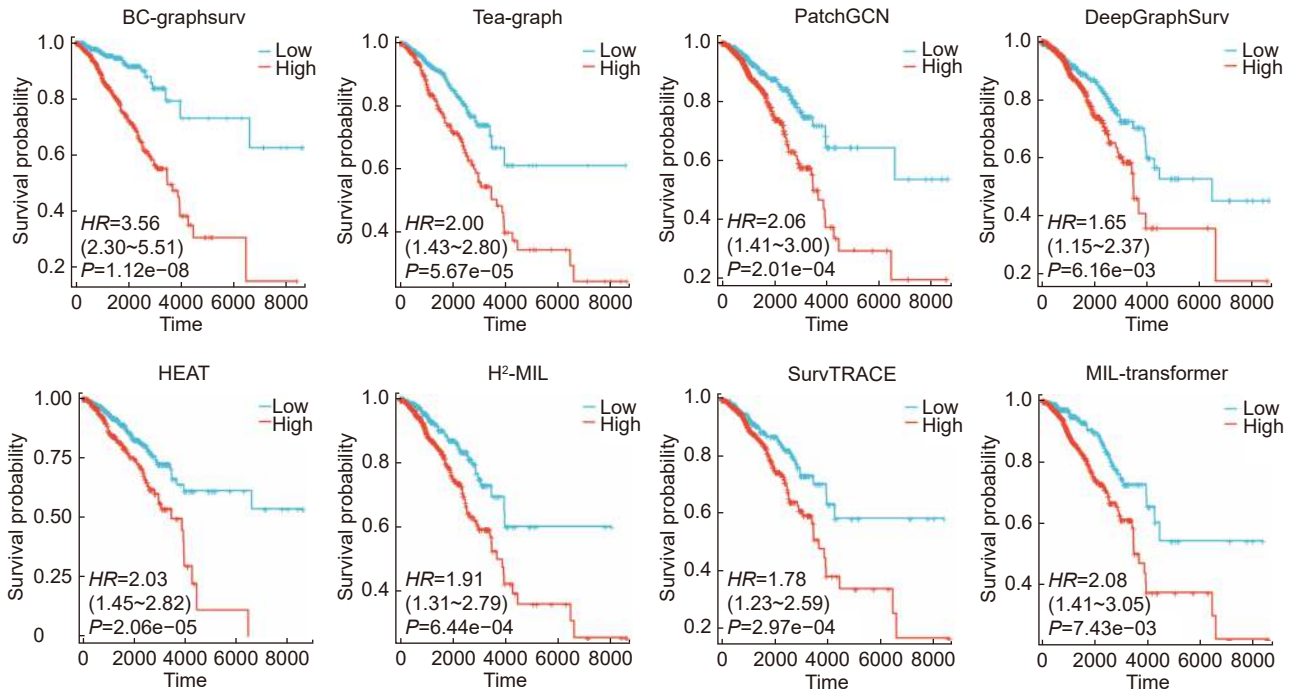


图 4 几种常用方法的 KM 曲线及 P 值对比

Fig. 4 Comparison of KM curves and P-values of several commonly used methods

释性工具 captum(0.4.0) 模块中的 IntegratedGradients 函数实现。节点的 IG 值为正则说明节点区域对该 WSI 是模型预测结果影响倾向于风险值增高, IG 值为负则说明节点区域对模型预测结果的影响倾向于风险值降低, IG 值绝对值越高则表示该节点区域对预测的影响越大。不同 IG 值的图像块如图 5 所示。

图 5 展示了模型对于肿瘤环境背景下不同细胞作用的敏感性。通过高 IG 值图像块和低 IG 值图像块的对比, 我们可以观察到模型成功识别了肿瘤组织细胞的不同作用。这种精细的识别和反映使得模型在风险预测中能够展现出高效的病理分析能力。

图 6 为 WSI 整体的 IG 值可视化分析, 相比于只在 ImageNet 预训练 ResNet50, 增加了病理迁移学习模块的 HF-Net 可以更准确地区分不同病理环境, 特别是在肿瘤细胞聚集扩散等高风险组织区域和正常组织区域之间具有更高的区分能力, 对病理图像任务的

适应能力更强, 从而在不同病理环境基础上做出合理的区域风险估计。这表明模型对于病理学图像的理解和分析能力较强, 能够有效地识别和区分不同组织类型, 为风险预测提供了可靠的基础。通过 IG 值可视化分析模型能够关注到对生存预测具有重要影响的区域, 提高了模型对于关键病理特征的敏感性, 从而增强了其在个体生存预测中的准确性和可解释性, 这一点对于医学图像处理中的临床应用尤为重要。

## 4 结论

为了实现大规模病理图像的自动分析, 提出了一种新型的乳腺癌 WSI 生存预测算法 BC-GraphSurv。通过采用基于 CNN 和自注意力机制的特征提取网络, 结合迁移学习的预训练策略, 成功实现了对病理学图像特征和空间关系信息的有效提取。通过改进的 WA-GAT 的预测网络辅以 MP-GCN 的预测网络分支, BC-

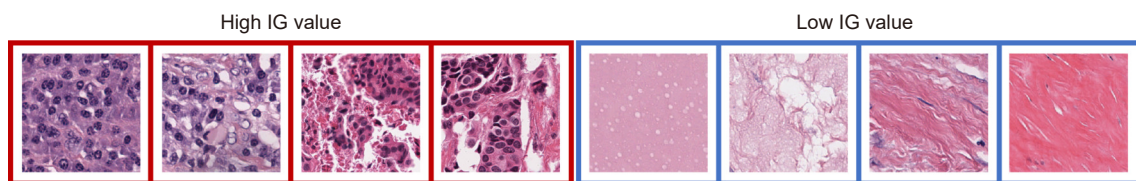


图 5 不同 IG 值图像块对比

Fig. 5 Comparison of patches with different IG values

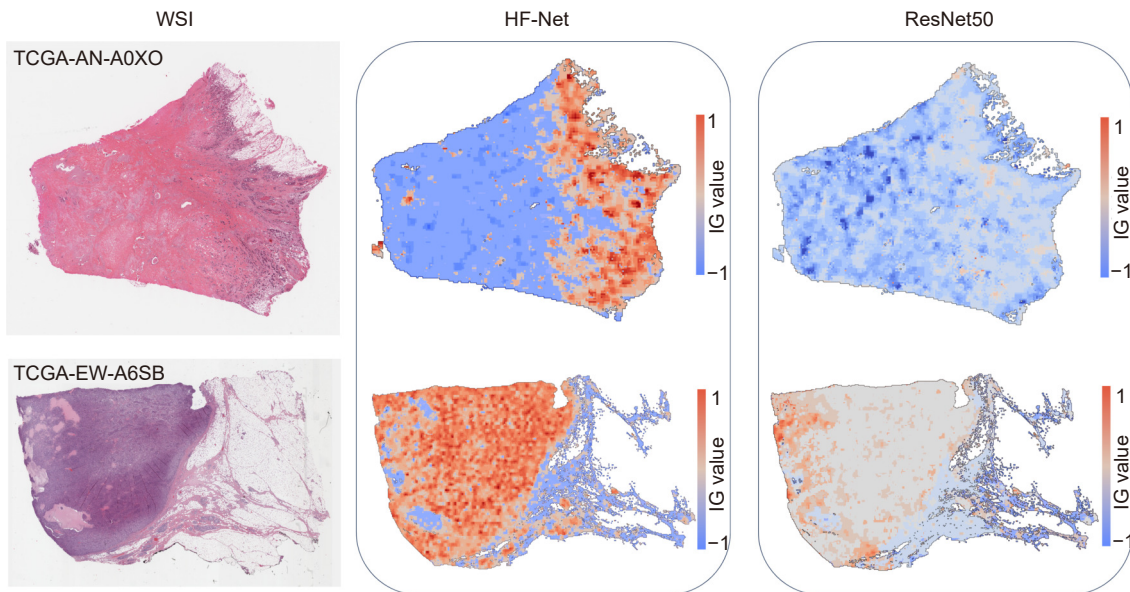


图 6 部分 WSI 病理环境可视化

Fig. 6 Visualization of WSI pathological environment

GraphSurv 成功融合了多种空间结构信息并能适应不同 WSI 的整体差异, 可以精确区分出肿瘤病理环境的异质性, 提高了模型的分析 and 预测能力。实验结果表明, 相较于此前最先进的算法, BC-GraphSurv 的 C-index 提升了 0.0409, 获得了精确的风险预测, 具备高效的病理分析能力。通过对模型关注区域的可视化分析, 可以辅助病理学家发现潜在的预后标志物, 提高模型的临床安全性。为了进一步提升算法性能, 未来的工作中可以适当地增加患者临床病历信息和基因组等信息, 进行多模态结合的分析预测, 提高模型预测的准确性。

## 参考文献

- [1] Li X T, Li C, Rahaman M M, et al. A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches[J]. *Artif Intell Rev*, 2022, **55**(6): 4809–4878.
- [2] Huang P, He P, Yang X, et al. Breast tumor grading network based on adaptive fusion and microscopic imaging[J]. *Opto-Electron Eng*, 2023, **50**(1): 220158.  
黄盼, 何鹏, 杨兴, 等. 基于自适应融合和显微成像的乳腺肿瘤分级网络[J]. *光电工程*, 2023, **50**(1): 220158.
- [3] Chan T H, Cendra F J, Ma L, et al. Histopathology whole slide image analysis with heterogeneous graph representation learning[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, Canada, 2023: 15661–15670. <https://doi.org/10.1109/CVPR52729.2023.01503>.
- [4] Liang L M, Lu B H, Long P W, et al. Adaptive feature fusion cascade Transformer retinal vessel segmentation algorithm[J]. *Opto-Electron Eng*, 2023, **50**(10): 230161.  
梁礼明, 卢宝贺, 龙鹏威, 等. 自适应特征融合级联Transformer视网膜血管分割算法[J]. *光电工程*, 2023, **50**(10): 230161.
- [5] Lv J, Wang Z Y, Liang H C. Boundary attention assisted dynamic graph convolution for retinal vascular segmentation[J]. *Opto-Electron Eng*, 2023, **50**(1): 220116.  
吕佳, 王泽宇, 梁浩城. 边界注意力辅助的动态图卷积视网膜血管分割[J]. *光电工程*, 2023, **50**(1): 220116.
- [6] Tellez D, Litjens G, van der Laak J, et al. Neural image compression for gigapixel histopathology image analysis[J]. *IEEE Trans Pattern Anal Mach Intell*, 2021, **43**(2): 567–578.
- [7] Zhu X L, Yao J W, Huang J Z. Deep convolutional neural network for survival analysis with pathological images[C]//*2016 IEEE International Conference on Bioinformatics and Biomedicine*, Shenzhen, China, 2016: 544–547. <https://doi.org/10.1109/BIBM.2016.7822579>.
- [8] Zhu X L, Yao J W, Zhu F Y, et al. WSISA: making survival prediction from whole slide histopathological images[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 2017: 6855–6863. <https://doi.org/10.1109/CVPR.2017.725>.
- [9] Yao J W, Zhu X L, Jonnagaddala J, et al. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks[J]. *Med Image Anal*, 2020, **65**: 101789.
- [10] Fan L, Sowmya A, Meijering E, et al. Cancer survival prediction from whole slide images with self-supervised learning and slide consistency[J]. *IEEE Trans Med Imaging*, 2023, **42**(5): 1401–1412.
- [11] Kim H E, Cosa-Linan A, Santhanam N, et al. Transfer learning for medical image classification: a literature review[J]. *BMC Med Imaging*, 2022, **22**(1): 69.
- [12] Li R Y, Yao J W, Zhu X L, et al. Graph CNN for survival analysis on whole slide pathological images[C]//*21st International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, Spain, 2018: 174–182. [https://doi.org/10.1007/978-3-030-00934-2\\_20](https://doi.org/10.1007/978-3-030-00934-2_20).
- [13] Chen R J, Lu M Y, Shaban M, et al. Whole slide images are

- 2D point clouds: context-aware survival prediction using patch-based graph convolutional networks[C]//24th International Conference on Medical Image Computing and Computer Assisted Intervention, Strasbourg, France, 2021: 339–349. [https://doi.org/10.1007/978-3-030-87237-3\\_33](https://doi.org/10.1007/978-3-030-87237-3_33).
- [14] Di D L, Zhang J, Lei F Q, et al. Big-Hypergraph factorization neural network for survival prediction from whole slide image[J]. *IEEE Trans Image Process*, 2022, **31**: 1149–1160.
- [15] Velickovic P, Cucurull G, Casanova A, et al. Graph attention networks[C]//6th International Conference on Learning Representations, Vancouver, Canada, 2018: 1–12.
- [16] Ozyoruk K B, Can S, Darbaz B, et al. A deep-learning model for transforming the style of tissue images from cryosectioned to formalin-fixed and paraffin-embedded[J]. *Nat Biomed Eng*, 2022, **6**(12): 1407–1419.
- [17] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[C]//5th International Conference on Learning Representations, Toulon, France, 2017: 1–14.
- [18] Gao Z Y, Lu Z Y, Wang J, et al. A convolutional neural network and graph convolutional network based framework for classification of breast histopathological images[J]. *IEEE J Biomed Health Inform*, 2022, **26**(7): 3163–3173.
- [19] Katzman J L, Shaham U, Cloninger A, et al. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network[J]. *BMC Med Res Methodol*, 2018, **18**(1): 24.
- [20] Kandath C, McLellan M D, Vandin F, et al. Mutational landscape and significance across 12 major cancer types[J]. *Nature*, 2013, **502**(7471): 333–339.
- [21] Spanhol F A, Oliveira L S, Petitjean C, et al. A dataset for breast cancer histopathological image classification[J]. *IEEE Trans Biomed Eng*, 2016, **63**(7): 1455–1462.
- [22] Otsu N. A threshold selection method from gray-level histograms[J]. *IEEE Trans Syst Man Cybern*, 1979, **9**(1): 62–66.
- [23] Heagerty P J, Zheng Y Y. Survival model predictive accuracy and ROC curves[J]. *Biometrics*, 2005, **61**(1): 92–105.
- [24] Loshchilov I, Hutter F. Decoupled weight decay regularization[C]//7th International Conference on Learning Representations, New Orleans, LA, USA, 2019: 1–8.
- [25] Smith L N, Topin N. Super-convergence: very fast training of neural networks using large learning rates[J]. *Proc SPIE*, 2019, **11006**: 1100612.
- [26] Ilse M, Tomczak J, Welling M. Attention-based deep multiple instance learning[C]//35th International Conference on Machine Learning, Stockholm, Sweden, 2018: 2127–2136.
- [27] Hou W T, Yu L Q, Lin C X, et al. H<sup>2</sup>-MIL: exploring hierarchical representation with heterogeneous multiple instance learning for whole slide image analysis[C]//Proceedings of the 36th AAAI Conference on Artificial Intelligence, 2022: 933–941. <https://doi.org/10.1609/aaai.v36i1.19976>.
- [28] Lee Y, Park J H, Oh S, et al. Derivation of prognostic contextual histopathological features from whole-slide images of tumours via graph deep learning[J]. *Nat Biomed Eng*, 2022. <https://doi.org/10.1038/s41551-022-00923-0>.
- [29] Chen R J, Lu M Y, Weng W H, et al. Multimodal co-attention transformer for survival prediction in gigapixel whole slide images[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 2021: 3995–4005. <https://doi.org/10.1109/ICCV48922.2021.00398>.
- [30] Wang Z F, Sun J M. SurvTRACE: transformers for survival analysis with competing events[C]//Proceedings of the 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, Northbrook, IL, USA, 2022: 49. <https://doi.org/10.1145/3535508.3545521>.
- [31] Lakatos E. Sample sizes based on the log-rank statistic in complex clinical trials[J]. *Biometrics*, 1988, **44**(1): 229–241.

## 作者简介



叶世杰 (2000-), 男, 硕士研究生, 研究方向为机器视觉。

E-mail: 223330843@qq.com



【通信作者】王永雄 (1970-), 男, 博士, 教授, 研究方向为智能机器人与机器视觉。

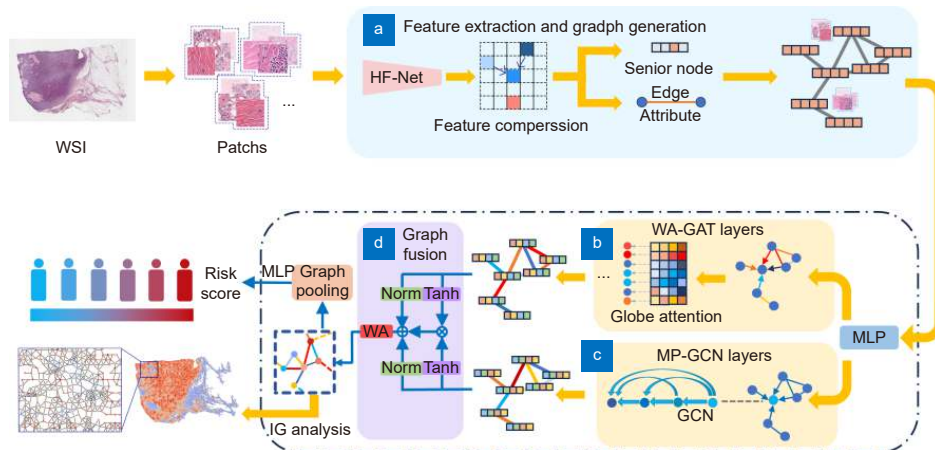
E-mail: wyxiong@usst.edu.cn



扫描二维码, 获取PDF全文

# Graph neural network-based WSI cancer survival prediction method

Ye Shijie, Wang Yongxiong\*



The model architecture proposed in this article mainly includes modules such as WSI preprocessing, (a) feature extraction and graph generation, (b) WA-GAT layers, (c) MP-GCN layers, and (d) graph fusion

**Overview:** In this study, we present BC-GraphSurv, an innovative model for breast cancer survival prediction utilizing Whole Slide Imaging (WSI). Given the challenges of large size, complex spatial relationships, and diverse styles in WSIs, BC-GraphSurv addresses these issues through a novel approach that integrates transfer learning and feature extraction using the HF-Net. The model consists of four steps: transfer learning with HF-Net, compression and fusion of similar features, construction of graph structure features, and learning with WA-GAT and MP-GCN. The model commences with a transfer learning pre-training strategy, utilizing HF-Net to construct the pathological relationship topology of WSIs. This strategy facilitates the effective extraction of features and spatial relationship information. HF-Net, trained on a breast cancer tumor classification dataset, is crucial for adapting a general backbone network to the complexity of tumor structures and tissue texture features. This network reduces noise in non-cancerous regions and enhances differentiation between cancerous and non-cancerous areas. The feature extraction network, combining Convolutional Neural Networks (CNN) and self-attention mechanisms, benefits from transfer learning to enhance pathology feature recognition via a feature transfer module. This module, coupled with spatial correlation and semantic similarity integration, enables compressed graph modeling and extraction of crucial contextual features for survival prediction. To overcome specific challenges in WSI tasks, BC-GraphSurv introduces improvements to the Graph Attention Network (GAT) in the form of the Whole Association Graph Attention Network (WA-GAT). This prediction branch employs cross-attention on node and edge features, a global perception module, and a Dense Graph Convolutional Network (GCN) for fine-grained details. The integration of WA-GAT and GCN enhances the model's adaptability to diverse WSI styles and spatial differences, effectively processing spatial information and improving analytical capabilities. Experimental validation involves ablation experiments assessing the impact of different modules and improvements. Comparative experiments with various models and visual analyses confirm the effectiveness of BC-GraphSurv. In conclusion, BC-GraphSurv provides a comprehensive solution for breast cancer survival prediction using WSIs. Experimental results on the TCGA-BRCA dataset showcase its effectiveness, with a consistency index of 0.795, surpassing current state-of-the-art models. The model's innovations effectively tackle the challenges inherent in WSI survival prediction, demonstrating robustness and superiority.

Ye S J, Wang Y X. Graph neural network-based WSI cancer survival prediction method[J]. *Opto-Electron Eng*, 2024, 51(4): 240011; DOI: 10.12086/oe.2024.240011

Foundation item: Project supported by Natural Science Foundation of Shanghai (22ZR1443700)

Institute of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

\* E-mail: wyxiong@usst.edu.cn