计算机研究与发展
Journal of Computer Research and Development
https://crad.ict.ac.cn
63(9): 2338−2354, 2026

PathRoute: 基于全切片图像的主动证据路由的癌症生存预测方法

吕千一[1,2] 罗汉文[1,2] 王 峻[2] 李 真[3] 余国先[1,2]

[1](山东大学软件学院 济南 250101)

[2](山东大学-南洋理工大学人工智能国际联合研究院 济南 250101)

[3](山东大学齐鲁医院 济南 250012)

(qylv@mail.sdu.edu.cn)

PathRoute: Active Evidence Routing for Cancer Survival Prediction on Whole Slide Pathology Images

Lü Qianyi[1,2], Luo Hanwen[1,2], Wang Jun[2], Li Zhen[3], and Yu Guoxian[1,2]

[1] (*School of Software*, *Shandong University*, *Jinan* 250101)

[2] (*Joint SDU-NTU Centre for Artificial Intelligence Research*, *Shandong University*, *Jinan* 250101)

[3] (*Qilu Hospital of Shandong University*, *Jinan* 250012)

Abstract Whole slide images (WSIs) are central to cancer diagnosis and prognosis. However, mainstream survival prediction methods directly map massive tile features into a single global representation for risk readout, failing to explicitly distinguish complex histological composition, ignoring cancer-specific clinical phenotype differences, and leaving the model's internal reasoning process opaque. To address these issues, we propose PathRoute, an active evidence routing framework for interpretable WSI survival prediction. The framework models WSI reasoning as a two-stage directional process that first forms evidence and then reads out risk: four role-specific learnable queries collaborate in a shared latent space, respectively performing tissue-composition sensing, active evidence routing, phenotype-semantic auxiliary supervision, and risk readout; a directional attention mask further constrains forward visibility between queries, so that risk prediction is structurally realized as a downstream readout of pre-formed evidence representations, rather than an end-to-end black-box mapping. On five TCGA solid tumor cohorts under unified inputs and identical patient-level 5-fold cross-validation splits, PathRoute using WSI alone achieves higher average C-index than representative multimodal and WSI-only methods, validating the effectiveness of structured evidence organization. Further analysis shows that routing queries consistently and spontaneously concentrate attention onto tumor regions across all five cancer types without any tissue-type supervision, providing direct interpretability support for the active routing mechanism. Code available at https://github.com/ZephyrGo/pathRoute.

Key words computational pathology; whole slide image (WSI); survival analysis; evidence routing; interpretable reasoning; pan-cancer

摘 要 病理全切片图像(whole slide image, WSI)是癌症诊断与预后判断的核心载体。当前主流的病理全切片生存预测方法将海量的区块特征直接映射为全局表示并输出预后风险,不仅未能在建模中显式区分复杂的组织学构成、忽略了不同癌种特有的临床表型差异,也导致模型的内部推理过程缺乏可解释性。

收稿日期: 2026-06-05; **修回日期**: 2026-07-17

基金项目: 国家自然科学基金重点项目 (62531013, 62432006); 国家自然科学基金面上项目 (62272276)

This work was supported by the Key Program of the National Natural Science Foundation of China (62531013, 62432006) and the General Program of the National Natural Science Foundation of China (62272276).

通信作者: 余国先(gxyu@sdu.edu.cn)

针对上述问题，提出 PathRoute，一个面向可解释生存预测的主动证据路由框架。该框架将 WSI 推理建模为先形成证据、再读出风险的两阶段有向过程：4 类角色明确的可学习查询在统一隐空间中分工协作，分别承担组织组成感知、主动证据路由、表型语义辅助监督与风险读出；方向性注意力掩码进一步约束查询间的前向可见性，使风险预测在架构上成为前序证据表示的下游读出，而非端到端的黑盒映射。在 TCGA 的 5 个常见实体瘤癌种上，采用统一输入特征与相同患者级 5 折交叉验证划分，PathRoute 仅使用 WSI 作为输入，平均生存预测 C-index 优于代表性多模态融合方法与纯 WSI 方法，验证了结构化证据组织的有效性。进一步的分析表明，路由查询在未接受任何组织类型监督的条件下，仍在 5 个癌种上稳定地将注意力自发集中于肿瘤区域，为主动证据路由的可解释性提供了直接支撑。代码目前已在 https://github.com/ZephyrGo/pathRoute 开源。

关键词 计算病理学；全切片图像；生存分析；证据路由；可解释推理；多癌种

中图法分类号 TP391

DOI: 10.7544/issn1000-1239.202660468 **CSTR:** 32373.14.issn1000-1239.202660468

病理全切片图像（whole slide image, WSI）是癌症诊断、分期与预后判断的核心载体，也是计算病理学最具挑战性的数据形态之一。一张切片在20×放大倍率下可被切分为$N \sim 10^4$个区块（tile），但与患者预后真正相关的诊断证据在切片中稀疏分布且事先难以定位；与此同时，临床可获得的生存标签通常仅有数百例，远小于切片本身的视觉复杂度。这一特殊的数据形态决定了 WSI 生存预测无法依靠端到端微调一个超大视觉骨干来解决：模型既需要在大量背景区块中识别出真正具有预后价值的局部证据，又必须在弱监督信号下避免过拟合。

近年来，UNI[1]、UNI2[2]、Virchow2[3]、CONCH[4] 与 GPFM[5] 等病理基础模型在大规模自监督预训练下提供了高质量的通用区块表征，使下游任务无需再从头训练视觉骨干。在这一新范式下，WSI 生存预测的瓶颈从底层视觉特征学习转向在强表征之上的证据组织，即如何从数千个区块特征中识别、聚合并利用与预后相关的诊断信息。然而，主流的多示例学习方法[6-8] 与多模态融合方法[9-12] 大多沿用单步注意力聚合范式：将整张切片压缩为单一向量后直接预测风险。这一范式带来 2 个相互关联的问题。1）证据选择、临床先验注入与风险判断被叠加到同一个表示之中，模型究竟依据哪些区块、哪些临床语义形成最终预后判断难以被审计。2）对全量区块直接执行稠密自注意力的计算代价随区块数N平方增长，对动辄$N \sim 10^4$的 WSI 不可行；硬选择不可微，随机采样亦无法保证命中稀疏的诊断性证据。

针对上述问题，本文提出 PathRoute，一个构建于冻结病理基础模型之上、面向可解释生存预测的主动证据路由框架。其核心思想是将 WSI 推理过程从单步聚合改造为先形成证据、再读出风险的两阶段有方向过程，并由一组角色明确的可学习查询在统一隐空间中分别承担 4 类语义功能：组织组成感知查询可以感知切片的整体组织构成；路由查询以软注意力方式从全量区块中聚合固定数量的诊断证据；表型语义查询通过癌种相关辅助监督接入临床先验；风险查询在前序信息就位的条件下完成最终预后判断。为使该信息流在架构层面具备结构性，PathRoute 进一步引入方向性注意力掩码，从前向计算上约束不同查询之间的可见性方向，使风险预测成为前序证据表示的下游读出，具体框架如图 1。本文的主要贡献如下：

1）面向全切片生存预测的主动证据路由框架。提出了一种面向 WSI 生存预测的主动证据路由框架 PathRoute。该框架打破了主流方法端到端“单步全局聚合”的黑盒范式，首创了“先提取证据、再预测风险”的两阶段定向推理过程。通过在统一隐空间中解耦 4 类功能明确的可学习查询（组织感知、路由聚合、表型语义、风险读出），并引入方向性注意力掩码，实现了对复杂病理特征的结构化建模与信息流的严格约束。

2）无监督可解释性机制。赋予了模型高度贴合临床直觉的可解释推理能力。实验证实，在未接受任何病理组织类型直接监督的情况下，PathRoute 中的路由查询能够跨越 5 个解剖部位完全不同的实体瘤，自发且稳定地将注意力精准聚焦于肿瘤区域。这不仅解决了 WSI 中诊断证据高度稀疏的痛点，更为生存预测中“肿瘤感知证据组织过程”提供了直接的视觉解释依据。

3）显著提升的跨癌种泛化与预测性能。在 5 个

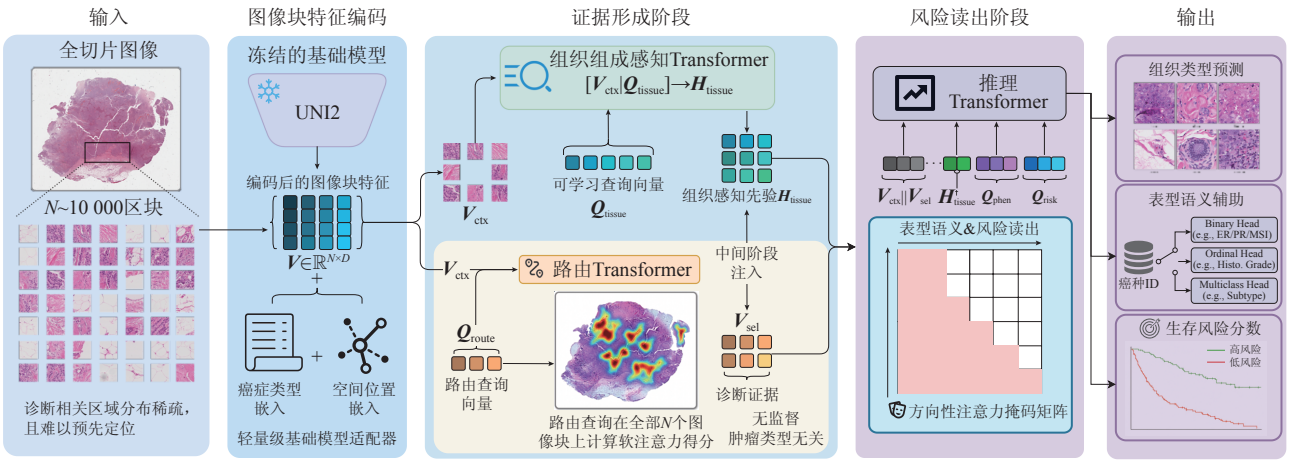


Fig. 1 Overview of the PathRoute framework

图 1 PathRoute 框架总览

代表性的 TCGA 实体瘤数据集(BRCA、STAD、COAD-READ、HNSC、BLCA)的严谨交叉验证中，PathRoute 仅依赖 WSI 单模态输入，其平均生存预测性能便全面超越了当前主流的纯 WSI 基础模型方法与多模态融合方法，充分验证了主动证据路由机制在临床预后评估中的有效性与鲁棒性。

1 相关工作

1.1 病理基础模型与 WSI 表征

大规模自监督预训练为计算病理学提供了高质量的通用区块表征，使下游任务对标注数据与任务特异性骨干训练的依赖显著降低。CONCH[4] 基于对比语言-图像预训练范式，在病理图像-文本配对数据上学习跨模态语义表示。UNI[1] 与其后续版本 UNI2[2] 基于 DINOv2[13] 自蒸馏框架，在超过 10 万张 WSI 上进行大规模预训练，所学表征在下游任务中展现出稳健的零样本与少样本迁移能力。Virchow2[3] 采用掩码自编码器预训练策略，在混合放大倍率的病理图像上训练。GPFM[5] 同时融合多个专家模型与自蒸馏目标，在 72 项病理下游任务的基准评测中取得平均排名第 1。

本文选择 UNI2 作为冻结特征提取器，将视觉表征学习与下游推理建模解耦。这一设计带来 2 点直接好处：其一，无需针对特定癌种或任务微调骨干，规避了在仅有数百例生存标签条件下端到端训练大模型的不稳定性；其二，区块特征可预提取并缓存，训练与推理时无需加载完整骨干，显著降低显存占用。

1.2 全切片生存预测

基于 WSI 的生存预测方法经历了从手工特征到深度学习再到多模态融合的演进。早期工作从病理图像中提取手工设计的形态学特征并结合 Cox 比例风险模型进行生存回归[14-15]。随后 Kather 等人[16] 证明卷积神经网络可从病理图像中学习与多癌种分子改变相关的形态表征，说明深度视觉特征能够捕获具有临床意义的组织学信号。

多示例学习(MIL)框架将 WSI 视为由数千区块组成的“包”，区块为“实例”，从而把千兆像素级图像分析转化为集合级预测问题[17]。Campanella 等人[18] 在大规模临床数据上验证了弱监督 MIL 的可行性[19]。ABMIL[6] 引入参数化注意力机制对每个区块独立评分并加权聚合，成为后续方法的标准基线[7-8,20-21]。为克服 ABMIL 忽略区块间关系的局限，TransMIL[7] 引入 Transformer 自注意力建模区块间的全局交互；DTFD-MIL[8] 通过双层特征蒸馏在效率与精度之间取得平衡。多模态融合方法进一步扩展了输入信息源：MCAT[9] 设计协同注意力机制融合 WSI 形态特征与基因组通路特征；SurvPath[10] 以密集交互方式融合 WSI 与生物通路 token；CIMA[11] 显式建模 WSI 与基因组通路之间的跨模态对齐；而 SurvPGC[12] 在 SurvPath 基础上以双路径跨注意力进一步纳入临床文本信息。

上述方法在 WSI 侧的处理高度同构：大多以单步注意力聚合将整张切片压缩为全局向量后送入风险预测头，使证据选择、临床先验注入与风险判断混叠在同一表示之中。多模态融合虽能提升精度，但依赖额外基因组或临床数据。

1.3 可学习查询与注意力聚合

用一组可学习查询从大量输入 token 中聚合信息，已成为视觉与多模态领域处理大规模或不定长输入的成熟方法学。DETR[22-23] 在目标检测中引入一

组目标查询向量，通过交叉注意力从图像特征中读取目标候选；Perceiver 与 PerceiverIO[24-25] 用一组潜在查询向量把任意大规模、任意模态的输入压缩到固定长度，使后续 Transformer 计算与输入规模解耦；BLIP-2 的 Q-Former[26-27] 通过 32 个可学习查询把视觉表征桥接到语言模型，实现了跨模态信息的紧凑路由。在病理领域，已有部分工作尝试将类似思路引入 WSI 聚合：例如基于可学习提示词元的 MIL 变体[28] 在 ABMIL 框架上引入额外查询以增强表示能力。这些工作共同表明，查询式注意力路由能够同时实现可微的软选择、与输入规模解耦的固定长度推理，以及对稀疏证据的有效聚焦。

然而，已有查询式方法仍难以满足 WSI 生存预测中的结构化推理需求。首先，查询之间缺乏明确的角色分化。无论是 DETR 中的目标查询，还是 Q-Former 中的潜在查询，其查询通常作为一组同构、可互换的表示槽位存在，模型并未显式规定不同查询分别承担组织感知、证据聚合、临床先验注入或风险判断等功能。这会导致不同预后相关信息在同一组查询中混合编码，使查询语义难以稳定解释，也削弱了中间表示与具体临床功能之间的对应关系。其次，已有方法通常允许查询之间进行充分的双向交互，缺乏面向推理流程的方向性约束。这样虽然有利于信息融合，但也会使证据形成与风险读出在前向计算中相互混杂：用于最终风险预测的判别信息可能提前影响中间证据表示，使模型难以区分哪些表示对应上游病理证据、哪些表示服务于下游风险决策。因此，现有查询聚合机制虽能获得紧凑的全局表征，却难以显式刻画“证据先形成、风险后判断”的推理结构。

本文的 PathRoute 在这条方法学脉络上的定位是：将“查询式路由”扩展为角色化、有方向的查询式推理。具体来说，PathRoute 为不同认知功能定义独立的查询子集（组织组成感知、证据路由、表型语义、风险读出），并通过方向性注意力掩码在前向计算中显式分离证据形成阶段与风险读出阶段，使最终预测在架构上成为前序证据表示的下游读出。

1.4 计算病理中的可解释性

可解释性是计算病理方法走向临床落地的关键瓶颈之一。现有解释方法大多属于事后归因范畴：注意力可视化将 ABMIL 或 TransMIL 中的区块级注意力权重映射回 WSI 空间，指示模型“关注”了哪些区域[6-7]；Grad-CAM 类方法[29] 通过梯度归因近似输入对输出的贡献；多尺度证据图[30] 进一步将归因从单层注意力扩展至跨尺度证据定位。这些方法的共同优势是不修改原模型即可应用，但其局限也很明确：归因是在模型训练完成后外加的近似估计，而模型的内部信息流本身仍是不可审计的黑箱。

与事后归因互补的另一条路径，是在架构设计阶段就为模型提供具有显式语义角色的中间槽位。概念瓶颈模型（concept bottleneck model, CBM）[31] 强制风险预测必须经由一组人类可读的中间概念，是这一路径的代表性尝试。但 CBM 类方法把全部预测压力施加在唯一的概念瓶颈上，当瓶颈内的概念无法覆盖所有相关信息时，模型性能会显著下降。

本文的 PathRoute 在事后归因与架构层可解释这 2 条路径之间走出了一条折中路线：不依赖事后注意力可视化，而是在架构层面为不同认知功能划分独立的可读语义槽位（组织、证据、表型、风险）；不同于概念瓶颈模型把表型语义设为风险预测的唯一中介路径，本文将其作为多个并列的可监督槽位之一，通过辅助监督向共享表示注入临床先验。这一设计在保留事后归因方法可应用性的同时，使模型的中间表示天然可读、可监督，从而提供“设计时可解释”的额外保障。

2 方　　法

2.1 总览与符号约定

本文提出 PathRoute，一个构建于冻结病理基础模型之上的主动证据路由框架，其核心设计是将全切片推理过程组织为先形成证据、再读出风险的两阶段结构。基础模型输出的区块特征首先被统一映射至后续推理所在的隐空间，并注入空间位置与癌种条件信息（2.2 节）；同时进入证据形成阶段（2.3 节），该阶段并行完成 2 类互补操作：组织组成感知在全局视野下建立切片的组织学先验，主动证据路由通过可学习查询从全量区块中聚合固定数量的诊断证据。在此基础上，风险读出阶段（2.4 节）通过方向性注意力掩码约束不同语义查询之间的前向可见性，使风险预测在架构上成为前序证据表示的下游读出，并借助癌种相关的表型辅助监督向共享表示注入临床先验。训练目标与计算复杂度分析分别在 2.5 节和 2.6 节给出。

本节中，N表示全切片的区块总数，d与D分别表示基础模型输出维度与 PathRoute 隐空间维度，c表示癌种 ID。PathRoute 引入 4 组角色明确的可学习查询，分别由 Q_1、Q_2、Q_3、Q_4个查询构成，记为 $\boldsymbol{Q}_{\text{tissue}}$

(组织组成感知)、$\boldsymbol{Q}_{\text{route}}$(证据路由)、$\boldsymbol{Q}_{\text{phen}}$(表型语义)与$\boldsymbol{Q}_{\text{risk}}$(风险读出)。证据形成阶段使用一个独立的$L_{\text{routing}}$层路由 Transformer 来承担证据路由，并在推理 Transformer 的前L_1层完成组织组成感知；风险读出阶段则在推理 Transformer 的后$L-L_1$层(其中L为推理 Transformer 总层数)展开。生存预测部分将连续生存时间离散为K个风险区间。其余局部符号在各节首次出现时定义。

2.2 区块特征编码

大规模自监督预训练使病理基础模型在海量切片数据上学习到了高质量的通用区块表征。因此，区块特征编码的目标是在不破坏基础模型所学表征的前提下，将其映射至推理模块所在的隐空间，并叠加空间位置与癌种条件，使每个区块的表示在进入证据形成阶段之前即携带完整的上下文信息。

具体地，设全切片经预处理裁剪为N个区块，由冻结的 UNI2 基础模型提取得到区块特征矩阵$\boldsymbol{F}\in\mathbb{R}^{N\times d}$，其中$d=1\,536$为 UNI2 的输出特征维度；同时记录每个区块在切片中的归一化坐标$\boldsymbol{C}\in\mathbb{R}^{N\times 4}$(中心坐标$x,y$与尺寸$w,h$)。为使区块特征与后续模块的查询向量在同一隐空间中进行注意力交互，本节将$\boldsymbol{F}$投影至D维隐空间。线性投影矩阵$\boldsymbol{W}_{\text{proj}}\in\mathbb{R}^{d\times D}$与偏置$\boldsymbol{b}_{\text{proj}}\in\mathbb{R}^{D}$给出：

$$\boldsymbol{V}_0=\boldsymbol{F}\boldsymbol{W}_{\text{proj}}+\boldsymbol{b}_{\text{proj}}。\tag{1}$$

空间位置嵌入$\boldsymbol{E}_{\text{spatial}}\in\mathbb{R}^{N\times D}$由多层感知机以归一化坐标$\boldsymbol{C}$为输入，生成编码区块在切片中的绝对位置与尺寸信息。癌种嵌入$\boldsymbol{e}_{\text{cancer}}(c)\in\mathbb{R}^{D}$为以癌种 ID$c$为索引的可学习向量，加和注入所有查询的初始化，使后续推理以癌种类型为条件。投影与位置嵌入之和经层归一化后得到视觉特征库：

$$\boldsymbol{V}=LayerNorm(\boldsymbol{V}_0+\boldsymbol{E}_{\text{spatial}})\in\mathbb{R}^{N\times D},\tag{2}$$

作为证据形成阶段 2 个子模块的共同输入。

2.3 证据形成阶段

获得全切片的统一区块表示$\boldsymbol{V}$之后，证据形成阶段的任务是从中提炼出 2 类互补的前驱表示，供风险读出阶段使用：其一是反映切片整体组织构成的组织学先验，其二是凝聚诊断相关区块信息的证据向量集合。值得注意的是，一张切片通常包含$N\sim10^4$个区块，对全量区块执行稠密自注意力的计算代价随N平方增长，在实践中不可行。

为此，2 个子模块均以从$\boldsymbol{V}$中均匀采样得到的上下文子集$\boldsymbol{V}_{\text{ctx}}\in\mathbb{R}^{C\times D}$作为共享输入，其中$m\ll N$为预设的上下文 token 数。该子集既保留了全切片的空间代表性，又把后续 Transformer 模块内部的自注意力序列长度从$O(N)$降至常数。2 个子模块以相同的$\boldsymbol{V}_{\text{ctx}}$为输入并行展开，互不依赖。

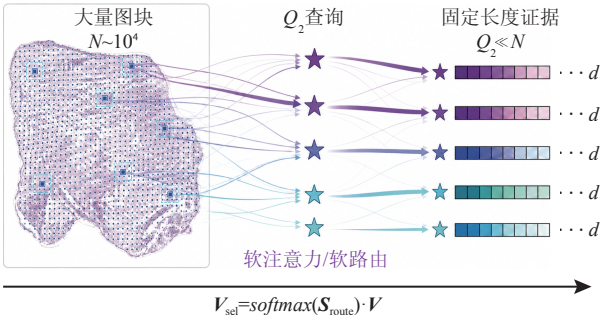


Fig. 2 Schematic diagram of active evidence routing

图 2 主动证据路由示意图

1)组织组成感知

推理 Transformer 由L层注意力块构成，其前L_1层($L_1<L$)专门承担组织组成感知的计算，负责建立切片整体的组织学先验。由Q_1个角色对称的可学习向量构成的组织查询$\boldsymbol{Q}_{\text{tissue}}\in\mathbb{R}^{Q_1\times D}$，与共享上下文拼接形成序列$[\boldsymbol{V}_{\text{ctx}}\|\boldsymbol{Q}_{\text{tissue}}]$后送入$L_1$层注意力块。在自注意力中，视觉 token 之间保持双向交互以充分建模组织间的空间依赖，组织查询单向读取视觉 token，并按位置序排列以诱导查询间的语义分化，避免不同查询坍缩至同一组织类型。该阶段输出组织先验表示$\boldsymbol{H}_{\text{tissue}}\in\mathbb{R}^{Q_1\times D}$，经层归一化后送入组织类型分类头，预测切片的组织类型构成；同时以梯度截断的形式注入下一阶段。

组织类型共设肿瘤、基质、坏死、脂肪、正常上皮、淋巴细胞浸润 6 类，监督标签由 CONCH[4] 零样本分类生成的区块级伪标签提供。由于组织查询预测的是一个无序集合，训练时采用 Hungarian 匹配[32]将查询预测与真值组织类型一一对齐，在匹配对上计算交叉熵损失。

2)主动证据路由

与组织组成感知并行，主动证据路由这一分支负责从全量区块中聚合固定数量的诊断证据，从而使下游推理的序列长度与全量区块数N解耦，如图 2 所示。PathRoute 为此设计了一个独立的L_{routing}层路由 Transformer，与上述组织组成感知所在的推理 Transformer 在参数上完全独立。

由Q_2个可学习向量构成的路由查询$\boldsymbol{Q}_{\text{route}}\in\mathbb{R}^{Q_2\times D}$与上下文$\boldsymbol{V}_{\text{ctx}}$拼接后送入该子模块，其在路由 Transformer 顶层的输出表示记为$\boldsymbol{Q}'_{\text{routing}}\in\mathbb{R}^{Q_2\times D}$。$\boldsymbol{Q}'_{\text{routing}}$经线性投影矩阵$\boldsymbol{W}_{\text{routing}}\in\mathbb{R}^{D\times D}$与全量区块特征$\boldsymbol{V}$做缩放点积，

得到路由得分：

$$\boldsymbol{S}_{\text{route}}=\frac{\boldsymbol{Q}'_{\text{routing}}\boldsymbol{W}_{\text{routing}}\boldsymbol{V}^{\text{T}}}{\sqrt{D}}\in\mathbb{R}^{Q_2\times N}。\tag{3}$$

在路由得分上施加$softmax$归一化（令$\boldsymbol{A}=softmax(\boldsymbol{S}_{\text{route}})$）后对全量区块进行连续加权聚合：

$$\boldsymbol{V}_{\text{sel}}=\boldsymbol{A}\cdot\boldsymbol{V}\in\mathbb{R}^{Q_2\times D}。\tag{4}$$

每个路由查询生成一个聚合证据向量，共同构成诊断证据$\boldsymbol{V}_{\text{sel}}$。该设计使路由器可在生存损失监督下联合优化，同时对低得分区块中潜在的弱预后信号保留梯度通道。路由查询自身的表示在生成$\boldsymbol{V}_{\text{sel}}$后即被舍弃。

3）阶段间衔接

至此，2 条并行的分支分别产出了诊断证据集合$\boldsymbol{V}_{\text{sel}}$与组织先验$\boldsymbol{H}_{\text{tissue}}$，它们将汇入风险读出阶段。聚合证据$\boldsymbol{V}_{\text{sel}}$与上下文$\boldsymbol{V}_{\text{ctx}}$拼接，构成增强的视觉序列；而组织先验$\boldsymbol{H}_{\text{tissue}}$则以梯度截断方式注入：

$$\boldsymbol{H}^{\dagger}_{\text{tissue}}=stopgrad(\boldsymbol{H}_{\text{tissue}})。\tag{5}$$

截断操作确保了组织组成感知与风险读出阶段在优化目标上保持严格隔离，防止生存损失回传导致组织查询的语义纯度受损。

2.4 风险读出阶段

风险读出阶段直接建立在证据形成阶段所提供的两路前驱表示之上。推理 Transformer 的后$L-L_1$层在组装好的综合序列上展开计算：

$$\boldsymbol{X}_{\text{read}}=\left[\boldsymbol{V}_{\text{ctx}}\|\boldsymbol{V}_{\text{sel}}\|\boldsymbol{H}^{\dagger}_{\text{tissue}}\|\boldsymbol{Q}_{\text{phen}}\|\boldsymbol{Q}_{\text{risk}}\right]，\tag{6}$$

并在自注意力运算中施加方向性注意力掩码M。

1）方向性注意力掩码

如图 3 所示，方向性注意力掩码M旨在约束序列$\boldsymbol{X}_{\text{read}}$的前向信息流向，规则如下。

底层证据双向互见，即视觉证据 token（$\boldsymbol{V}_{\text{ctx}}$与$\boldsymbol{V}_{\text{sel}}$）之间保持双向交互；

高层查询单向读取，即所有查询 token 可单向读取视觉证据，但查询对视觉证据 token 不可见；

查询角色严格排序，即查询 token 之间按角色逻辑先后施加下三角约束：$\boldsymbol{H}^{\dagger}_{\text{tissue}}\rightarrow\boldsymbol{Q}_{\text{phen}}\rightarrow\boldsymbol{Q}_{\text{risk}}$。

这一掩码机制保障了风险查询必须在视觉证据、组织先验与表型语义均就位的条件下才能完成预后判断；同时，表型查询在计算时无法窥视风险查询的中间表征，保证了其作为独立临床监督接口的可读性。

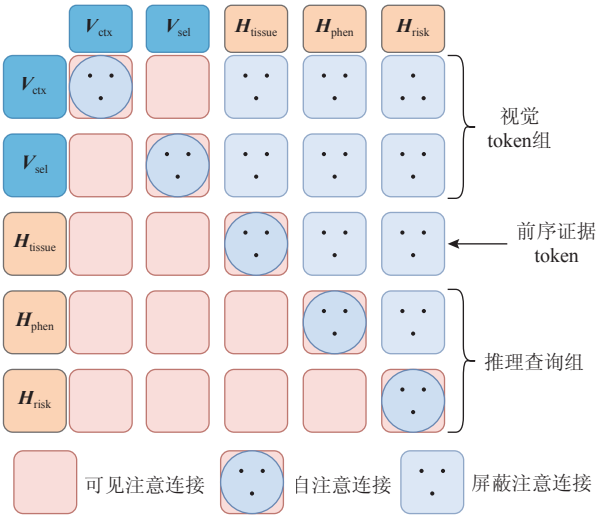


Fig. 3 Illustration of the directional attention mask

图 3 方向性注意力掩码示意图

2）表型语义辅助监督

在完成序列交互后，由Q_3个向量构成的表型查询$\boldsymbol{Q}_{\text{phen}}$输出表示$\boldsymbol{H}_{\text{phen}}\in\mathbb{R}^{Q_3\times D}$，经平均池化送入癌种相关的辅助监督头，用于预测受体状态、组织学分级等形态学相关的临床表型。该设计旨在为共享推理表示注入临床先验知识，其角色定位为辅助监督而非风险预测的强制中介。

3）风险读出

作为整个模型的最终输出端，由Q_4个向量构成的风险查询$\boldsymbol{Q}_{\text{risk}}$输出表示$\boldsymbol{H}_{\text{risk}}\in\mathbb{R}^{Q_4\times D}$，经平均池化后送入生存预测头。模型采用离散时间生存预测框架[10]，预测头输出K个风险区间的 hazard 率$h_1,\cdots,h_K$，对应的离散生存函数为：

$$S(t_j)=\prod_{k=1}^{j}(1-h_k)。\tag{7}$$

最终计算得出的风险分数为$r=-\sum_{j=1}^{K}S(t_j)$。风险区间边界由训练集中未删失样本生存时间的K分位数事先确定。

2.5 训练目标

PathRoute 在训练阶段同时优化 3 类监督损失与 1 项查询多样性正则。各项损失分别对应不同的设计目标，而非简单的多任务加权：组织监督约束组织组成感知查询学习稳定的全局组织学先验；表型监督为共享推理表示注入癌种相关的临床先验；生存监督直接优化最终风险输出；多样性正则约束同组查询不发生表示坍缩。总损失$\mathcal{L}$定义为：

$$\mathcal{L}=\lambda_{\text{tissue}}\mathcal{L}_{\text{tissue}}+\lambda_{\text{phen}}\mathcal{L}_{\text{phen}}+\lambda_{\text{surv}}\mathcal{L}_{\text{surv}}+\lambda_{\text{div}}\mathcal{L}_{\text{div}}。\tag{8}$$

1）组织监督$\mathcal{L}_{\text{tissue}}$。组织组成感知查询的预测与无序的真值组织集合通过 Hungarian 匹配[32]一一对齐后计算交叉熵。组织类型标签由 CONCH[4]零样本分类生成的区块级伪标签提供，5 个癌种共用同一套

6类组织定义。

2)表型监督$\mathcal{L}_{\text{phen}}$。由通用调度器根据癌种注册表自动派发。调度器扫描当前癌种注册的所有监督任务，根据预测logit维度自动推断损失类型：输出维度为1时使用二元交叉熵(BCE)；目标为整数序数时使用累积BCE；其余情况使用多类交叉熵。

3)生存监督$\mathcal{L}_{\text{surv}}$。采用离散时间负对数似然损失[33]约束风险输出。设B为当前训练批次的样本数，$c_i\in\{0,1\}$为样本i的删失指示变量($c_i=1$表示删失)，$y_i\in\{1,\cdots,K\}$为该样本所属的风险区间索引(全文以$k\in\{1,\cdots,K\}$表示风险区间索引，以$i\in\{1,\cdots,B\}$表示批内样本索引，二者不混用；式中$S(\cdot)$依式(7)定义，h_k为第k个区间的hazard率)，则生存损失定义为：

$$\mathcal{L}_{\text{surv}}=-\frac{1}{B}\sum_{i=1}^{B}[(1-c_i)(\log S(t_{y_i-1})+\log h_{y_i})+c_i\log S(t_{y_i})]。\tag{9}$$

4)多样性正则$\mathcal{L}_{\text{div}}$。为了促进不同查询编码不同的语义，模型对组织、表型、风险这3组查询的输出向量分别计算归一化余弦相似度矩阵，并惩罚同组查询之间的非对角相似度均值。设$s\in\{\text{tissue},\text{phen},\text{risk}\}$为查询组索引，$Q_s$为该组内的查询数量，$\boldsymbol{q}_i^{(s)}$为该组的第$i$个查询向量表示，正则项计算为：

$$\mathcal{L}_{\text{div}}=\frac{1}{3}\sum_{s}\frac{1}{Q_s(Q_s-1)}\sum_{i\neq j}|\cos(\boldsymbol{q}_i^{(s)},\boldsymbol{q}_j^{(s)})|。\tag{10}$$

需要特别指出的是，路由查询$\boldsymbol{Q}_{\text{route}}$不参与此项正则，因为它们的“多样性”已经由其对全切片不同区域的注意力分布间接体现和保障。

式(8)的4项损失并非在所有训练步骤中同等激活。$\mathcal{L}_{\text{div}}$在每次前向传播中无条件计算；$\mathcal{L}_{\text{tissue}}$仅在当前切片具有CONCH伪标签时激活；$\mathcal{L}_{\text{phen}}$仅在当前癌种已注册表型监督任务且批次中存在对应标签时生效；$\mathcal{L}_{\text{surv}}$则在所有具有生存时间与事件标注的样本上均参与计算，在本文5个癌种数据集中无缺失。

算法1给出了PathRoute单次前向传播的完整流程，与上述各节的方法描述保持一致，可作为实现参考。

算法 1. PathRoute端到端推理与损失计算流程。

输入：基础模型提取的区块特征$\boldsymbol{F}\in\mathbb{R}^{N\times d}$，归一化坐标$\boldsymbol{C}\in\mathbb{R}^{N\times 4}$，癌种ID$c$；

输出：风险分数r，联合损失$\mathcal{L}$。

/* 1. 区块特征编码*/

① $\boldsymbol{V}_0\leftarrow\boldsymbol{F}\boldsymbol{W}_{\text{proj}}+\boldsymbol{b}_{\text{proj}}$；

② $\boldsymbol{V}\leftarrow LayerNorm(\boldsymbol{V}_0+MLP(\boldsymbol{C}))$；/*癌种嵌入加入各查询初始化*/

③ $\boldsymbol{V}_{\text{ctx}}\leftarrow UniformSample(\boldsymbol{V},m)$；/*保留$m$个token*/

/* 2. 证据形成阶段：双分支并行*/

/* % 分支2a：组织组成感知*/

④ $\boldsymbol{H}_{\text{tissue}}\leftarrow Transformer_{1\to L_1}([\boldsymbol{V}_{\text{ctx}}\|\boldsymbol{Q}_{\text{tissue}}])$；

⑤ $\boldsymbol{H}^{\dagger}_{\text{tissue}}\leftarrow stop_gradient(\boldsymbol{H}_{\text{tissue}})$；/*维持语义纯度的梯度截断*/

/* % 分支2b：主动证据路由*/

⑥ $\boldsymbol{Q}'_{\text{routing}}\leftarrow RoutingTransformer_{1\to L_{\text{routing}}}([\boldsymbol{V}_{\text{ctx}}\|\boldsymbol{Q}_{\text{route}}])$；

⑦ $\boldsymbol{S}_{\text{route}}\leftarrow\dfrac{\boldsymbol{Q}'_{\text{routing}}\boldsymbol{W}_{\text{routing}}\boldsymbol{V}^{\mathrm{T}}}{\sqrt{D}}$，$\boldsymbol{A}\leftarrow softmax(\boldsymbol{S}_{\text{route}})$；

⑧ $\boldsymbol{V}_{\text{sel}}\leftarrow\boldsymbol{A}\boldsymbol{V}$(加权聚合得到诊断证据集合)；

/* 3. 风险读出阶段*/

⑨ $\boldsymbol{X}_{\text{read}}\leftarrow[\boldsymbol{V}_{\text{ctx}}\|\boldsymbol{V}_{\text{sel}}\|\boldsymbol{H}^{\dagger}_{\text{tissue}}\|\boldsymbol{Q}_{\text{phen}}\|\boldsymbol{Q}_{\text{risk}}]$；/*构造方向性注意力掩码$M$，约束前向可见性*/

⑩ $\boldsymbol{H}_{\text{out}}\leftarrow Transformer_{L_1+1\to L}(\boldsymbol{X}_{\text{read}},mask=M)$；

⑪ 从$\boldsymbol{H}_{\text{out}}$中按序拆分出表型输出$\boldsymbol{H}_{\text{phen}}$与风险输出$\boldsymbol{H}_{\text{risk}}$；

/* 4. 读出与监督计算*/

⑫ $\hat{y}_{\text{tissue}}\leftarrow TissueHead(\boldsymbol{H}_{\text{tissue}})$；

⑬ $\hat{y}_{\text{phen}}\leftarrow PhenotypeHead(MeanPool(\boldsymbol{H}_{\text{phen}}))$；

⑭ $h_1,\cdots,h_K\leftarrow SurvivalHead(MeanPool(\boldsymbol{H}_{\text{risk}}))$；

⑮ $S(t_j)\leftarrow\prod_{k=1}^{j}(1-h_k)$，其中$j\in\{1,\cdots,K\}$；

⑯ $r\leftarrow-\sum_{j=1}^{K}S(t_j)$；

⑰ $\mathcal{L}\leftarrow\lambda_{\text{tissue}}\mathcal{L}_{\text{tissue}}+\lambda_{\text{phen}}\mathcal{L}_{\text{phen}}+\lambda_{\text{surv}}\mathcal{L}_{\text{surv}}+\lambda_{\text{div}}\mathcal{L}_{\text{div}}$；

⑱ 返回$r,\mathcal{L}$。

2.6 计算复杂度

WSI生存预测的现实瓶颈在于：对$N\sim10^4$的全量区块直接执行稠密自注意力计算在显存和算力上均不可行。PathRoute通过两阶段设计将主要的注意力计算转化为固定长度的交互，使风险读出阶段的计算量与切片物理大小彻底解耦。

具体而言，各阶段的计算复杂度如下：

1)主动证据路由。该阶段的输入序列长度为$C+Q_2$；此外，路由查询还需与全量N个区块进行缩放点积以计算路由得分。因此，其计算复杂度为$O((C+Q_2)^2+Q_2N)$。

2)组织组成感知。参与计算的序列长度为$C+Q_1$，其注意力计算复杂度为$O((C+Q_1)^2)$，与区块总数N无关。

3)风险读出。参与前向交互的综合序列长度为$C+Q_2+Q_1+Q_3+Q_4$，其复杂度为$O((C+Q_2+Q_1+Q_3+$

$Q_4)^2$)，同样与区块总数N无关。

综合来看，得益于主动证据路由的降维作用，PathRoute 的端到端注意力计算复杂度仅线性依赖于区块总数N（主要来源为路由得分计算项Q_2N），其余所有的深层语义交互计算均与切片的区块总数N严格解耦。

3 实　　验

3.1 数据集与评估指标

实验在 TCGA[34]PanCanAtlas 的 5 个癌种（BRCA、STAD、COADREAD、HNSC、BLCA）上进行，覆盖乳腺癌、胃癌、结直肠癌、头颈癌和膀胱癌等常见实体瘤；样本量分别为 1 111、398、565、467、455。所有实验采用相同的患者级 5 折交叉验证划分，确保与现有方法的公平比较。各癌种样本量见本节说明。

生存预测的评估指标为 Harrell's C-index[35]，对每个癌种以 5 折交叉验证给出验证集的 C-index 均值±标准差，表 1 中 Overall 列为 5 个癌种均值的算术平均。C-index 衡量模型预测的风险排序与实际生存时间排序的一致性，取值范围为[0,1]，0.5 对应随机预测。

3.2 实现细节

1）特征提取。所有 WSI 方法均使用 UNI2[2] 作为冻结骨干网络，输出特征维度$d = 1\,536$；WSI 经 20 倍放大裁剪为256×256像素的区块，每张切片预提取得到的切片级特征以 HDF5 格式缓存为统一输入。

2）模型架构。PathRoute 的架构超参数在所有癌种和所有实验中统一固定：隐藏维度$D = 1\,024$；4 类查询数$Q_1/Q_2/Q_3/Q_4 = 8/16/16/8$；上下文采样数$C = 200$；路由 Transformer 层$L_{\text{routing}} = 4$，推理 Transformer 层$L = 12$（前 4 层用于组织组成感知）；注意力头数为 16，FFN 维度为 4 096；组织类型数$T = 6$，生存风险区间数$K = 4$。所有实验采用统一训练设置。查询数（Q_3）、组织组成感知层数（L_1）与表型损失权重（λ_{phen}）等核心超参数的敏感性分析见附录 C，模型对其在合理范围内不敏感（C-index 最大波动小于 0.02）。

3.3 基线方法

本文选取 12 种代表性相关方法进行比较：

1）纯基因组方法。SNN[36] 为自归一化神经网络，仅以基因组通路特征为输入，代表基因组单模态基线的性能上界。MLP-Omics 为基因组通路特征上的多层感知机基线，作为相对于 SNN 的简单参照模型，用以衡量基因组单模态在最简结构下可达的水平。

2）纯 WSI 方法。ABMIL[6] 通过参数化注意力权重对区块特征做加权池化，将 WSI 压缩为单一向量后预测生存；DeepMISL[37] 在注意力聚合的基础上引入深度多示例学习框架，对区块进行分层聚合以捕获更丰富的切片级表示；TransMIL[7] 引入 Transformer 自注意力机制建模区块间的全局交互关系。上述方法均为单步聚合范式，缺乏显式的中间推理阶段。

3）多模态方法（WSI + 基因组）。上述 3 种纯 WSI 方法均有对应的"+Path"多模态变体，通过拼接 WSI 聚合特征与基因组通路特征实现早期融合。MCAT[9] 设计了协同注意力机制，实现 WSI 形态特征与基因组通路特征的深度交互融合。SurvPath[10] 以密集交互方式融合 WSI token 与生物通路 token 进行生存预测。CIMA[11] 面向多模态不完整场景，通过循环模态重建模块相互填补缺失的病理与基因组数据，并以多粒度对齐显式建模 WSI 区块与生物通路之间的粗粒度与细粒度交互。SurvPGC[12] 在 SurvPath 基础上引入双路径跨注意力融合结构，原文同时纳入临床信息文本提示词；本文为对齐统一的 WSI 与基因组输入条件，仅采用其病理与基因组双模态变体进行公平比较。

需要指出的是，PathRoute 仅使用 WSI 图像输入，不依赖任何基因组数据，因此与纯 WSI 方法（ABMIL、DeepMISL、TransMIL）一样为同类可比方法，而与多模态方法的比较进一步刻画了结构化证据组织的相对作用。

3.4 主要结果

为保证公平比较，所有 WSI 方法均使用相同的 UNI2 编码特征作为输入，所有多模态方法使用相同的基因组通路输入，并采用相同的患者级 5 折交叉验证划分。在此条件下，表 1 比较的是不同模型结构对同一输入表征的利用能力。

表 2 展示了 PathRoute 与基线方法在 5 个癌种上的生存预测 C-index 对比。PathRoute 在 BRCA（0.680）、STAD（0.665）、COADREAD（0.717）、HNSC（0.646）四个癌种上取得最高 C-index，5 个癌种 Overall 为 0.666。与最强的多模态基线 SurvPGC（0.643）相比，Overall 提升 0.023，在 HNSC、STAD、COADREAD 上优势尤为明显，分别提升 0.048, 0.037, 0.026；与最强的纯 WSI 基线 TransMIL（0.604）相比，Overall 提升 0.062。值得注意的是，BRCA 上 PathRoute（0.680）仅小幅优于该癌种上最强的多模态基线 MCAT（0.677），差距 0.003，两者差距已在标准差范围内，这一结果与 BRCA 已有较成熟的形态学表型监督信号、信息增益空间有限相一致。上述结果表明，在统一输入

Table 1 C-index for Survival Prediction on Five Cancer Types

表 1 五癌种生存预测 C-index

方法	Patho.	Geno.	BRCA	STAD	COADREAD	HNSC	BLCA	Overall
SNN		✓	0.634±0.110	0.643±0.056	0.649±0.062	0.572±0.020	0.612±0.061	0.622
MLP-Omics		✓	0.609±0.140	0.539±0.094	0.687±0.059	0.602±0.064	0.647±0.069	0.617
ABMIL	✓		0.531±0.071	0.506±0.092	0.692±0.152	0.583±0.025	0.639±0.025	0.590
DeepMISL	✓		0.585±0.086	0.580±0.041	0.684±0.118	0.554±0.047	0.522±0.014	0.585
TransMIL	✓		0.575±0.079	0.507±0.105	0.673±0.147	0.624±0.055	0.643±0.016	0.604
ABMIL+Path	✓	✓	0.589±0.084	0.531±0.097	0.685±0.129	0.561±0.037	0.620±0.047	0.597
DeepMISL+Path	✓	✓	0.593±0.050	0.548±0.060	0.592±0.117	0.555±0.059	0.597±0.052	0.577
TransMIL+Path	✓	✓	0.578±0.105	0.577±0.120	0.709±0.132	0.623±0.046	0.630±0.029	0.623
MCAT	✓	✓	0.677±0.087	0.576±0.077	0.692±0.084	0.617±0.039	0.649±0.054	0.642
SurvPath	✓	✓	0.633±0.123	0.536±0.088	0.618±0.161	0.629±0.058	0.629±0.027	0.609
CIMA	✓	✓	0.664±0.090	0.592±0.017	0.620±0.097	0.626±0.026	0.595±0.050	0.619
SurvPGC	✓	✓	0.638±0.038	0.628±0.061	0.691±0.138	0.598±0.023	**0.658±0.036**	0.643
PathRoute	✓		**0.680±0.050**	**0.665±0.075**	**0.717±0.143**	**0.646±0.075**	0.621±0.073	**0.666**

注：黑体数值为最优结果，下划线数值为次优结果。“±”前的数值为 5 折均值，“±”后的数值为标准差。Overall 为 5 个癌种均值的算术平均值。

Table 2 Component and Structural Ablations

表 2 组件与结构消融

分类	配置	BRCA	STAD	COADREAD	HNSC	BLCA	Overall
组件消融（关闭单个语义组件）	完整 PathRoute	**0.680±0.050**	0.665±0.075	**0.717±0.143**	**0.646±0.075**	0.621±0.073	**0.666**
	w/o Tissue Census	0.666±0.040	**0.668±0.090**	0.686±0.148	0.619±0.031	0.612±0.083	0.650
	Random Routing	0.677±0.047	**0.668±0.060**	0.712±0.119	0.610±0.053	**0.624±0.055**	0.658
	w/o Phenotype Supervision	0.658±0.063	0.653±0.100	0.711±0.130	0.603±0.054	0.609±0.058	0.647
	Survival-Only	0.634±0.060	0.660±0.066	0.704±0.157	0.606±0.053	0.595±0.098	0.639
结构消融（保留语义组件、扰动结构约束）	完整 PathRoute	**0.680±0.050**	0.665±0.075	0.717±0.143	**0.646±0.075**	0.621±0.073	**0.666**
	Block-Directional Mask	0.673±0.034	0.665±0.099	0.718±0.148	0.607±0.044	0.624±0.072	0.657
	Fully Bidirectional Mask	0.666±0.040	**0.671±0.096**	0.720±0.150	0.604±0.041	0.620±0.074	0.656
	Reordered Roles	0.673±0.040	0.661±0.095	**0.727±0.144**	0.596±0.041	0.626±0.077	0.657
	w/o Tissue Stop-Gradient	**0.680±0.039**	0.658±0.102	0.723±0.138	0.602±0.045	**0.627±0.064**	0.658

注：黑体数值为该列最高值，下划线数值为次高值。“±”前的数值为 5 折 C-index 均值，“±”后的数值为标准差。Overall 为 5 个癌种均值的算术平均值。

条件下，结构化证据路由与方向性风险读出能够更有效地利用 WSI 表征完成生存预测；同时也说明，在强视觉基础表征（UNI2）支撑下，仅凭 WSI 本身即可提供较强的预后信号，不必需要额外的基因组通路输入作为补充。

在 BLCA 上，PathRoute（0.621）低于多模态方法 SurvPGC（0.658），降低 0.037。这与膀胱癌预后对基因组通路信息（如 FGFR 通路突变状态）的依赖较强结论一致，纯 WSI 输入难以完全恢复其分子层面的预后线索。

就结果方差而言，COADREAD 的标准差在各方法中普遍偏高（PathRoute 为±0.143），这与结直肠癌队列较小、5 折划分时各折的终点事件分布不均衡有关：个别折的验证集中终点事件高度集中，风险排序难度显著高于其他折，从而拉大了折间 C-index 的波动幅度。其他癌种标准差相对较小（BRCA 为±0.050、STAD 为 ±0.075、HNSC 为 ±0.075、BLCA 为 ±0.073），整体结果可信。

3.5 消融实验

为系统评估 PathRoute 不同设计对最终生存预测的贡献，本节将消融实验分为 2 组，如表 2 所示。组件消融关闭一个语义模块，但不改动其余模块的结构与监督，回答“哪些结构化语义模块是有效的”；结构消融保持 4 类查询及其监督全部就位，仅替换或扰动方向性注意力掩码、查询顺序、梯度截断等结构机制，回答“为何采用当前的信息流组织方式”。

所有消融均沿用主实验的输入特征、训练超参数与5折交叉验证划分。下文按组件消融、结构消融、综合分析依次展开。

3.5.1 组件消融

组件消融逐一关闭一个语义模块。w/o Tissue Census移除组织组成感知分支，下游查询直接从视觉token中获取信息，用以检验全局组织学先验对共享表示的塑形作用。Random Routing用均匀随机采样替代主动证据路由，但保持下游推理阶段的视觉序列长度不变，用以检验软注意力路由相对于无偏采样的增益。w/o Phenotype Supervision保留表型查询及其位置，仅去除表型监督损失（受体状态、组织学分级、MSI等标签均不使用），用以检验表型语义辅助监督对共享表示学习的帮助。Survival-Only是合并消融，同时关闭组织组成感知、将证据路由替换为随机采样、并取消所有表型监督，使风险查询成为唯一受监督信号驱动的组件，估计上述3类辅助模块作为整体的贡献区间。

表2显示，移除任一组件均会造成五癌种Overall平均C-index下降，说明组织组成感知、主动证据路由与表型辅助监督均对最终预测产生贡献，且贡献来源各不相同。w/o Tissue Census对C-index带来平均1.6个百分点的下降，验证了全局组织学先验作为共享上下文的塑形作用。w/o Phenotype Supervision带来最大的单组件下降（1.9个百分点），且在乳腺癌（BRCA）这一组织异质性最高的癌种上下降最为显著（从0.680降至0.658）。这与表型监督的设计目标相符：通过受体状态等癌种相关标签为共享推理表示注入临床先验，使学习信号不仅来自稀疏的生存事件。需要强调的是，表型监督的作用是训练阶段的多任务正则，而非风险预测的强制中介路径，这与方法节中关于方向性掩码“不强制风险经由表型语义中转”的论述一致。

相比之下，Random Routing在五癌种C-index的Overall上仅造成平均0.8个百分点的下降，是3个单组件消融中降幅最小的，这一数字需结合3点解读。其一，Overall平均掩盖了单癌种上的显著差异：在头颈癌（HNSC）上主动路由（0.646）较随机路由（0.610）高3.6个百分点，难癌种上的稳定性收益不可忽略。其二，在UNI2强表征下随机采样已能给出有竞争力的排序，故主动路由的必要性主要体现在可解释的证据组织：3.6节展示主动路由能在2 391张切片上稳定将注意力质量自发集中于肿瘤（集中度中位数为2.30，Wilcoxon $p<10^{-48}$），而随机路由集中度恒为1且不产生任何可对齐肿瘤的证据聚合；一旦移除主动路由，本文3.6节与3.7节所依托的整个可解释性分析将不复成立。其三，如实而言，纯C-index维度上主动路由相对随机采样增益有限；PathRoute将其定位为“在几乎不损失预测精度的前提下换取可审计的证据组织”，而非仅仅提升C-index的机制。

Survival-Only同时关闭3类辅助模块，带来平均2.7个百分点的下降，幅度大于任一单组件消融。这一结果表明3类辅助模块的贡献并非相互冗余，而是共同在共享表示上施加了来自不同语义层面的正则信号。

3.5.2 结构消融

结构消融保留全部4类查询不变，仅扰动其实现机制中的某一结构约束。Block-Directional Mask与Fully Bidirectional Mask分别将方向性注意力掩码松弛为“组内双向、跨组单向”和“完全双向”，对应文献中2类常见的多步推理掩码方案。Reordered Roles保持掩码形式不变，但将表型查询物理上前置于组织组成感知查询之前，破坏组织先验在前向计算中的优先性。w/o Tissue Stop-Gradient移除组织先验的梯度截断，允许生存损失沿组织先验回传至组织组成感知查询，用以检验梯度截断对维持组织组成感知作为可读组织学先验槽位是否必要。

表2显示，结构消融的整体降幅小于组件消融，且不同癌种之间存在波动。Overall从0.666下降至0.656~0.658，绝对差距处于跨折标准差范围内。这一观察与本文对方向性掩码的方法定位一致：方向性掩码的核心作用并非作为单纯的性能放大技巧，而是在前向计算层面规定不同语义角色的可见性方向，使风险查询成为前序证据表示的下游读出。

Reordered Roles与w/o Tissue Stop-Gradient进一步表明，结构化查询的语义角色需要与其在推理流中的位置和监督边界保持一致。当组织先验被允许在生存损失下被反向更新，或被排在表型查询之后时，模型仍能学习到具有竞争力的风险预测，但其内部表示与“组织先验—诊断证据—表型语义—风险读出”的设计对应关系被削弱，组织组成感知作为可读组织学先验的槽位意义随之减弱。

3.5.3 综合分析

综合2类消融可以看到，组件消融表明组织组成感知、主动证据路由与表型辅助监督均对最终性能产生独立而互补的贡献；结构消融表明方向性注意力掩码、查询顺序与梯度截断等机制主要承担方向性结构约束的功能，而非单纯的性能放大手段。2类消融共同支持了本文将WSI生存预测组织为“组织

先验形成—诊断证据路由—表型语义注入—风险读出”两阶段结构化推理的设计思路。

3.6 主动证据路由的语义对齐分析

本节定量考察主动证据路由学到的注意力分布与组织语义之间的对齐情况。值得强调的是，PathRoute中的路由查询$\boldsymbol{Q}_{\text{route}}$本身并不接受任何直接的组织类型监督：尽管模型整体保留了组织组成感知的监督，但路由查询的优化信号仅来自下游生存损失。因此，路由分布是否对齐到诊断相关组织，可作为对所学证据语义的间接观察。

具体地，对每张测试切片，将路由 Transformer 输出的Q_2个查询的注意力权重按查询维度平均，得到每个区块的归一化注意力权重a_i；同时利用 CONCH 零样本伪标签$\ell_i \in \{1,\cdots,6\}$得到每个区块的组织类型。对每种组织类型$t$，定义注意力集中度为

$$conc_t = \frac{\sum_i a_i \cdot \mathbb{I}[\ell_i = t]}{\sum_i \mathbb{I}[\ell_i = t]/N_{\text{tiles}}} = \frac{attention_share_t}{tile_share_t}, \quad (11)$$

其中$attention_share_t$为注意力质量在组织类型t上的占比，$tile_share_t$为该类型区块的数目占比。$conc_t = 1$对应“路由对组织t无偏好”，与均匀随机路由相同；$conc_t > 1$表明路由主动将注意力集中到该组织。

表 3 汇总了 5 个癌种共 2 391 张切片上的每个切片的集中度中位数统计。所有 5 个癌种的肿瘤组织集中度中位数均显著大于 1(1.90~2.84)，其余 5 种组织的集中度中位数均小于 1，且超过 90% 的切片其肿瘤集中度中位数大于 1。配对 Wilcoxon 符号秩检验拒绝“$conc_{\text{tumor}} \leqslant 1$”零假设的$p$值在所有癌种上均小于$10^{-48}$。跨癌种合并 2 391 张切片：肿瘤集中度中位数为 2.30，93.8% 的切片肿瘤集中度大于 1；其余 5 种组织的集中度中位数均小于 1，路由主动冷落了非诊断性组织(如脂肪、淋巴细胞)。

这一现象在 5 个解剖部位完全不同的癌种上保持一致，表明主动证据路由在生存监督下学到了与肿瘤分布稳定对齐的注意力偏好，并非某一癌种的偶然产物。结合 3.5 节的消融结果可以看到，主动路由相对随机路由的 C-index 增益虽然有限，但其在证据选择层面的语义对齐为模型提供了端到端 C-index 之外的可解释价值。需要说明的是，本节分析依赖 CONCH 生成的组织伪标签而非人工精标注的肿瘤区域。主动证据路由与传统注意力池化(如 ABMIL)在信息利用与关键区域筛选上的机制差异及定量对照分析见附录 D。CONCH 组织伪标签相对人工标注的可靠性验证见附录 E。

3.7 案例研究：4 类查询的可解释性

3.6 节在 2 391 张切片上定量考察了主动证据路由的语义对齐情况。本节进一步以一例 BRCA 切片(TCGA-3C-AALI)为代表，展示 PathRoute 中 4 类语义查询在单次前向传播中分别产出的可读中间输出，从而具体说明“结构化证据组织”在临床推理层面的可解释性体现。

图 4 整合了该切片的可解释性快照。图 4(a)(b)呈现：每个区块按 CONCH 伪标签上色(颜色映射见右上“组织类型”图例)，上下文采样区块以白框标记，路由 top-20 区块以阶梯化标记叠加在叠加图上，反映了不同排名查询所聚焦的区域(标记说明见右上“路由标记”图例)。图 4(c)~(f)呈现 4 类查询的中间输出。

1)组织组成感知。给出 8 个组织查询各自最确定的组织类型预测及其置信度。在该切片上，多数查询同时识别出肿瘤、坏死、基质、淋巴细胞等典型组织类型，且组织类型预测在查询q_0~q_7之间相互分化、未发生坍缩，与多样性正则的设计目标一致。

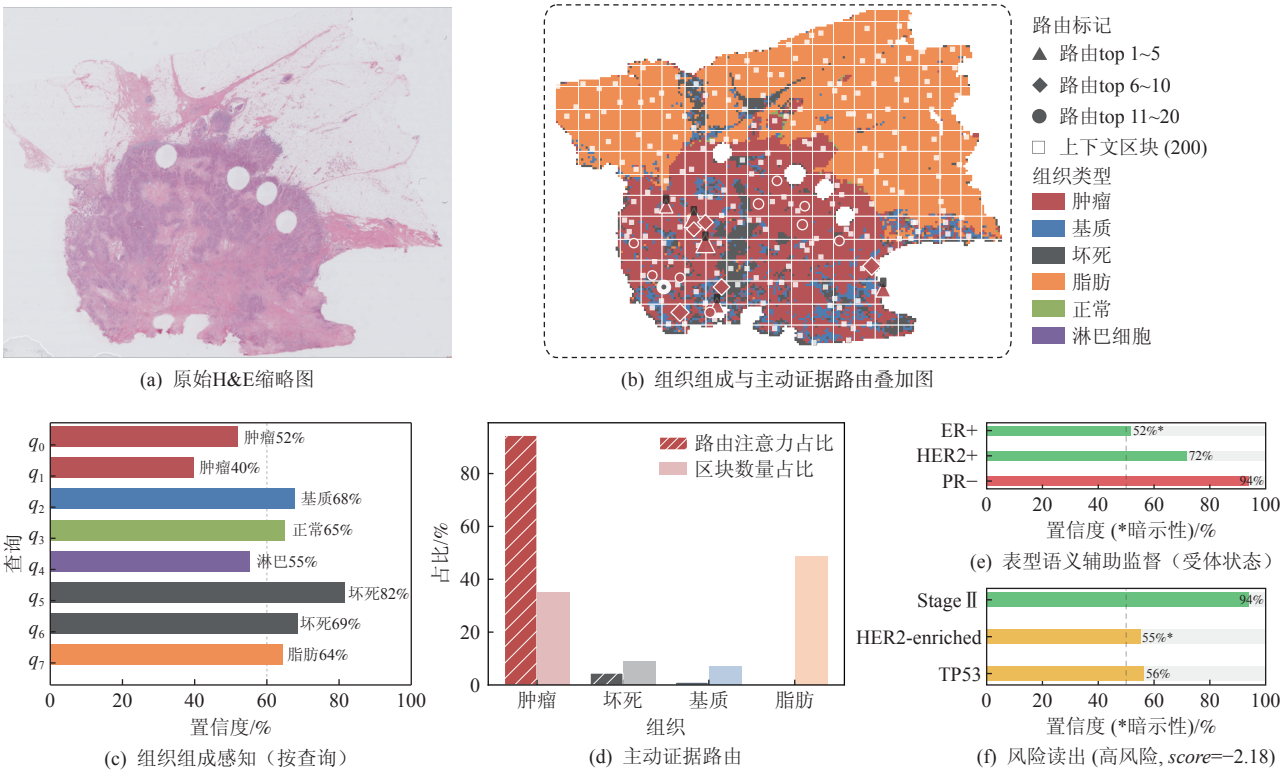
2)主动证据路由。给出全切片范围内每种组织的“路由注意力占比”与“区块数量占比”的对照：肿瘤组织以 35% 的区块数量贡献了 94% 的路由注意力，相应过表达倍数约 2.7；其余组织(基质、坏死、脂肪等)的注意力占比则普遍低于其区块占比。这与 3.6 节在 2 391 张切片上的统计结论一致，并具体地展示了路由对该切片诊断证据的聚焦行为。

Table 3 Tile-level Attention Concentration of Active Evidence Routing

表 3 主动证据路由的区块级注意力集中度

癌种（n切片）	肿瘤	间质	坏死	脂肪组织	正常组织	淋巴细胞
BRCA (n = 922)	**2.04**/0.95	0.02/0.09	0.05/0.14	0.02/0.03	0.02/0.08	0.00/0.02
STAD (n = 342)	**1.90**/0.89	0.27/0.12	0.36/0.16	0.46/0.24	0.88/0.42	0.48/0.23
COADREAD (n = 292)	**2.84**/0.96	0.00/0.01	0.15/0.08	0.00/0.02	0.46/0.29	0.00/0.03
HNSC (n = 412)	**2.56**/0.98	0.14/0.14	0.13/0.05	0.00/0.03	0.13/0.15	0.03/0.07
BLCA (n = 423)	**2.69**/0.90	0.03/0.05	0.38/0.25	0.01/0.10	0.05/0.06	0.03/0.11

注：黑体数值表示集中度中位数大于 1，即有效聚焦。“/”左侧为癌种组织的集中度中位数，“/”右侧为集中度中位数大于 1 的组织比例。



(a) 原始H&E缩略图 (b) 组织组成与主动证据路由叠加图

(c) 组织组成感知（按查询） (d) 主动证据路由 (e) 表型语义辅助监督（受体状态） (f) 风险读出 (高风险, *score*=−2.18)

Fig. 4 Interpretability snapshot of four query types in a BRCA slide (TCGA-3C-AALI) of PathRoute

图 4 PathRoute 在一例 BRCA 切片（TCGA-3C-AALI）上的 4 类查询可解释性快照

3）表型语义辅助监督。给出 ER/HER2/PR 三个临床受体状态的预测：在该例上，模型给出 ER+（52%）、HER2+（72%）、PR−（94%）的预测，对应 PAM50 推断标签下的真实表型。表型查询的中间输出与 PathRoute 整体判断之间通过方向性掩码保持单向可见性，因此该输出可作为独立可读的临床先验接口供下游审计。

4）风险读出。综合前序信息给出最终的临床判断：病理分期、分子亚型、可能突变靶点（如 TP53）以及生存风险等级。在该例上，风险查询给出 HIGH 风险等级（*score* =−2.18），与该患者的临床随访结果一致。

总体来看，PathRoute 的 4 类查询在单次前向传播中各自产出独立可读、独立可监督的中间表示：组织组成感知确立全局组织学先验；主动证据路由聚焦诊断相关区块；表型语义提供受体状态等临床先验；风险查询完成最终预后判断。这一“设计时可解释”的中间表示链条，使模型的内部信息流不必依赖事后注意力可视化即可被审计。

4 结　论

本文提出 PathRoute，一个面向 WSI 生存预测的结构化证据组织框架。其核心由 2 部分构成：独立的路由 Transformer 通过软注意力主动从全量区块中聚合诊断证据，使下游推理的序列长度与切片大小解耦；推理 Transformer 通过方向性注意力掩码约束组织组成感知、表型监督与风险读出之间的前向可见性方向，使风险预测在架构上成为前序证据表示的下游读出。

在 5 个 TCGA 癌种的统一特征输入与相同患者级 5 折交叉验证划分下，PathRoute 仅以 WSI 作为输入即在五癌种 C-index 上的 Overall 优于代表性多模态融合方法与代表性纯 WSI 方法。进一步的语义对齐分析表明，主动证据路由在仅由生存损失驱动的条件下，仍在数千张切片上跨 5 种解剖部位完全不同的癌种稳定地将注意力质量自发集中于肿瘤组织区域，为结构化证据组织的有效性提供了直接的可解释性证据。

本文工作的局限主要体现在 2 个方面：1）所有实验均在 TCGA 数据上进行，跨机构、跨扫描仪的泛化能力仍有待外部多中心数据集评估。2）PathRoute 当前不消费基因组与转录组信号，在分子层信息占主导的癌种（如 BLCA）上其性能上限受限于纯影像信息的可恢复程度；未来可在保留结构化证据组织的前提下，将基因组与文本临床信息以可控方式接入，扩展框架的多模态可用性。在更广阔的方向上，结

构化证据路由可推广至分子标志物预测、治疗响应预测等其他病理任务，推动计算病理方法向辅助诊断工具的转化。

作者贡献声明：吕千一负责研究构思、方法学、软件实现、验证、实验实施、形式分析、可视化并撰写论文；罗汉文负责方法学，撰写并审阅论文；王峻、李真、余国先负责监督指导、撰写并审阅论文。

参 考 文 献

[1] Chen R J, Ding T, Lu M Y, et al. Towards a general-purpose foundation model for computational pathology[J]. Nature Medicine, 2024, 30(3): 850−862

[2] MahmoodLab. UNI2-h: An updated pathology foundation model [EB/OL]. 2025 [2026-07-10]. https://huggingface.co/MahmoodLab/UNI2-h

[3] Zimmermann E, Vorontsov E, Viret J, et al. Virchow2: Scaling self-supervised mixed magnification models in pathology [EB/OL]. 2024 [2026-07-10]. https://arxiv.org/abs/2408.00738

[4] Lu M Y, Chen B, Williamson D F K, et al. A visual-language foundation model for computational pathology[J]. Nature Medicine, 2024, 30(3): 863−874

[5] Ma J, Guo Z, Zhou F, et al. A generalizable pathology foundation model using a unified knowledge distillation pretraining framework [J]. Nature Biomedical Engineering, 2026, 10: 545−564

[6] Ilse M, Tomczak J M, Welling M. Attention-based deep multiple instance learning [C] //Proc of the 35th Int Conf on Machine Learning. Cambridge, MA: PMLR, 2018: 2127−2136

[7] Shao Z, Bian H, Chen Y, et al. TransMIL: Transformer based correlated multiple instance learning for whole slide image classification [C] //Advances in Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2021, 34: 2136−2147

[8] Zhang H, Meng Y, Zhao Y, et al. DTFD-MIL: Double-tier feature distillation multiple instance learning for histopathology whole slide image classification [C] //Proc of the IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2022: 18802−18812

[9] Chen R J, Lu M Y, Weng W H, et al. Multimodal co-attention transformer for survival prediction in gigapixel whole slide images [C] //Proc of the IEEE/CVF Int Conf on Computer Vision. Piscataway, NJ: IEEE, 2021: 4015−4025

[10] Jaume G, Vaidya A, Chen R J, et al. Modeling dense multimodal interactions between biological pathways and histology for survival prediction [C] //Proc of the IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2024: 11579−11590

[11] Bu Y, Niu Q, Li Z, et al. Cancer survival prediction by cyclic generation and multi-grained alignment [C] //Proc of the AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2026, 40(24): 19781−19789

[12] Hou J, Zhang R, Xie Y, et al. Multimodal deep learning for cancer prognosis prediction with clinical information prompts integration[J]. npj Digital Medicine, 2026, 9: 76

[13] Oquab M, Darcet T, Moutakanni T, et al. DINOv2: Learning robust visual features without supervision[J/OL]. Transactions on Machine Learning Research, 2024 [2026-07-16]. https://openreview.net/forum?id=a68SUt6zFt

[14] Mobadersany P, Yousefi S, Amgad M, et al. Predicting cancer outcomes from histology and genomics using convolutional networks [J]. Proceedings of the National Academy of Sciences, 2018, 115(13): E2970−E2979

[15] Cox D R. Regression models and life-tables[J]. Journal of the Royal Statistical Society: Series B (Methodological), 1972, 34(2): 187−202

[16] Kather J N, Heij L R, Grabsch H I, et al. Pan-cancer image-based detection of clinically actionable genetic alterations[J]. Nature Cancer, 2020, 1(8): 789−799

[17] Wang Meixia, Wang Huiying, Huang Miaomiao, et al. Survey on multi-instance multi-label learning[J]. Journal of Computer Research and Development, 2026, 63(1): 66−89 (in Chinese)
(王梅霞, 王慧颖, 黄苗苗, 等. 多示例多标记学习综述[J]. 计算机研究与发展, 2026, 63(1): 66−89)

[18] Campanella G, Hanna M G, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images[J]. Nature Medicine, 2019, 25(8): 1301−1309

[19] Ding Shifei, Sun Yuting, Liang Zhizhen, et al. Survey on support vector machine algorithms in weakly supervised scenarios[J]. Chinese Journal of Computers, 2024, 47(5): 987−1009 (in Chinese)
(丁世飞, 孙玉婷, 梁志贞, 等. 弱监督场景下的支持向量机算法综述[J]. 计算机学报, 2024, 47(5): 987−1009)

[20] Lu M Y, Williamson D F K, Chen T Y, et al. Data-efficient and weakly supervised computational pathology on whole-slide images [J]. Nature Biomedical Engineering, 2021, 5(6): 555−570

[21] Li B, Li Y, Eliceiri K W. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning [C] //Proc of the IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 14318−14328

[22] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers [G] //LNCS 12346: Proc of the European Conf on Computer Vision. Cham: Springer, 2020: 213−229

[23] Li Wensheng, Zhang Jing, Zhuo Li, et al. Overview of Transformer-based visual segmentation[J]. Chinese Journal of Computers, 2024, 47(12): 2760−2782 (in Chinese)
(李文生, 张菁, 卓力, 等. 基于 Transformer 的视觉分割技术进展[J]. 计算机学报, 2024, 47(12): 2760−2782)

[24] Jaegle A, Gimeno F, Brock A, et al. Perceiver: General perception with iterative attention [C] //Proc of the Int Conf on Machine Learning. Cambridge, MA: PMLR, 2021: 4651−4664

[25] Jaegle A, Borgeaud S, Alayrac J B, et al. Perceiver IO: A general architecture for structured inputs and outputs[C/OL]// Proc of the 10th Int Conf on Learning Representations. Amherst, MA: OpenReview. net, 2022 [2026-07-16]. https://openreview.net/forum?id=fILj7WpI-g

[26] Li J, Li D, Savarese S, et al. BLIP-2: Bootstrapping language-image pre-training with frozen image encoders and large language models

[C] //Proc of the Int Conf on Machine Learning. Cambridge, MA: PMLR, 2023: 19730−19742

[27] Zeng Cheng, Ge Yunjie, Zhao Lingchen, et al. Survey of multimodal vision-language representation learning models and their adversarial examples attack and defense techniques[J]. Journal of Computer Research and Development, 2025, 62(9): 2208−2232 (in Chinese)
(曾诚, 葛云洁, 赵令辰, 等. 多模态视觉语言表征学习模型及其对抗样本攻防技术综述[J]. 计算机研究与发展, 2025, 62(9): 2208−2232)

[28] Zhang J, Kapse S, Ma K, et al. Prompt-MIL: Boosting multi-instance learning schemes via task-specific prompt tuning [C] //Proc of the Int Conf on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2023: 624−634

[29] Selvaraju R R, Cogswell M, Das A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization [C] //Proc of the IEEE Int Conf on Computer Vision. Piscataway, NJ: IEEE, 2017: 618−626

[30] Hashimoto N, Fukushima D, Koga R, et al. Multi-scale domain-adversarial multiple-instance CNN for cancer subtype classification with unannotated histopathological images [C] //Proc of the IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 3852−3861

[31] Koh P W, Nguyen T, Tang Y S, et al. Concept bottleneck models [C] //Proc of the Int Conf on Machine Learning. Cambridge, MA: PMLR, 2020: 5338−5348

[32] Kuhn H W. The Hungarian method for the assignment problem[J]. Naval Research Logistics Quarterly, 1955, 2(1-2): 83−97

[33] Zadeh S G, Schmid M. Bias in cross-entropy-based training of deep survival networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(9): 3126−3137

[34] Hutter C, Zenklusen J C. The cancer genome atlas: Creating lasting value beyond its data[J]. Cell, 2018, 173(2): 283−285

[35] Harrell F E, Califf R M, Pryor D B, et al. Evaluating the yield of medical tests[J]. JAMA, 1982, 247(18): 2543−2546

[36] Klambauer G, Unterthiner T, Mayr A, et al. Self-normalizing neural networks [C] //Advances in Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2017, 30: 971−980

[37] Yao J, Zhu X, Jonnagaddala J, et al. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks[J]. Medical Image Analysis, 2020, 65: 101789



Lü Qianyi, born in 2002. Master candidate. His main research interests include computational pathology and machine learning.
吕千一, 2002 年生。硕士研究生。主要研究方向为计算病理学、机器学习。



Luo Hanwen, born in 2002. PhD candidate. His main research interests include causal discovery and bioinformatics.
罗汉文, 2002 年生。博士研究生。主要研究领域为因果发现、生物信息学。



Wang Jun, born in 1983. PhD, professor, PhD supervisor. Senior member of CCF. Her main research interests include machine learning and bioinformatics.
王　峻, 1983 年生。博士, 教授, 博士生导师。CCF 高级会员。主要研究方向为机器学习、生物信息学。



Li Zhen, born in 1984. PhD, professor, chief physician. Her main research interest includes new technologies and methods for the study of digestive system diseases.
李　真, 1984 年生。博士, 教授, 主任医师。主要研究方向为消化系统疾病研究新技术与新方法。



Yu Guoxian, born in 1985. PhD, professor, PhD supervisor. Distinguished member of CCF. His main research interests include data mining and bioinformatics.
余国先, 1985 年生。博士, 教授, 博士生导师。CCF 杰出会员。主要研究方向为数据挖掘、生物信息学。

附录 A. 实验设置

A.1 超参数设置

表 A1 汇总了 PathRoute 的全部架构与训练超参数; 这些超参数在所有癌种、所有消融与敏感性实验中保持统一固定, 不作为调优变量。

A.2 模型参数量分布

表 A2 列出了 PathRoute 各组件的可训练参数量。UNI2 骨干(约 1.5×10^9 参数量)完全冻结, 不计入可训练参数。

附录 B. 表型语义监督

表型语义辅助监督是 PathRoute 向共享推理表征注入临床先验的接口, 其角色定位为辅助监督而非风险预测的强制中介。本节给出该机制的具体实现配置(任务、标签、头类型)以及表型查询的预测准确率, 作为方法实现细节与中间输出可读性的参考。

表 B1 给出 5 个 TCGA 癌种的样本量及所选用的表型监督任务。任务选取遵循“形态学高度相关、临床意义公认”原则: BRCA 选择激素受体(ER/PR/

Table A1 Architecture and Training Hyperparameters

表 A1 架构与训练超参数

分类	超参数	取值
架构	骨干特征维度 d	1 536
	隐藏维度 D	1 024
	查询数 $Q_1/Q_2/Q_3/Q_4$	8/16/16/8
	上下文采样数 C	200
	路由 Transformer 层数 $L_{routing}$	4
	推理 Transformer 层数 L/L_1	12/4
	注意力头数 H	16
	FFN 维度	4 096
	组织类型数 T	6
	生存风险区间数 K	4
训练	优化器	AdamW
	学习率	5E−5
	Weight decay	0.05
	学习率调度	Cosine
	Warmup 步数	500
	梯度累积/等效 batch	4/4
正则化	Tile dropout 率	0.3
	Drop path rate	0.1
	Label smoothing epsilon	0.1
硬件	GPU	NVIDIA L40, 47.6 GB

Table A2 Distribution of Model Parameters

表 A2 模型参数量分布

组件	参数量
Layer 1 内容编码器	约 3×10^6
Layer 2 路由 Transformer	约 25×10^6
Layer 2 推理 Transformer	约 160×10^6
Layer 3 任务输出头	约 20×10^6
总计	约 209×10^6

Table B1 Dataset Statistics and Sources of Phenotype Supervision Labels

表 B1 实验数据集统计与表型监督标签来源

癌种	样本量	表型监督	标签来源
BRCA	1 111	ER/PR/HER2	PAM50 推断
STAD	398	分级 G1~G3	病理报告
COADREAD	565	MSI 状态	CMS 推断
HNSC	467	分级 G1~G4	病理报告
BLCA	455	基底/腔面	mRNA

HER2), STAD 与 HNSC 选择病理报告中的组织学分级, COADREAD 选择微卫星不稳定性(MSI)状态, BLCA 选择基底/腔面亚型; 真值标签部分来自病理报告(STAD/HNSC 分级), 部分来自分子层面的代理标签(BRCA 的 PAM50 推断、COADREAD 的 CMS 推断、BLCA 的 mRNA 子型)。分子代理标签并非完美真值, 而是与对应癌种形态学之间存在文献支持的关联信号, 本框架将其作为软监督使用, 不对预测准确率本身追求最优。

表 B2 列出当前框架内 5 个癌种的具体注册条目。注册表以“癌种 ID -> 任务列表”形式组织, 每条任务声明其名称、标签字段与头类型。通用调度器在前向传播时根据当前 batch 的癌种 ID 检索对应注册项, 并依据预测 logit 维度自动推断损失类型: 输出维度为 1 的任务采用二元交叉熵(BCE), 整数序数任务采用累积 BCE, 其余采用多类交叉熵。该设计使新癌种的接入仅需追加 1 行配置而无需修改任何模型代码; 当前 batch 中标签缺失或癌种未注册任务的样本对表型损失贡献为零, 从而实现跨癌种联合训练与癌种异质表型语义的解耦。

Table B2 Configuration of the Cancer-aware Supervision Registry

表 B2 癌种感知监督注册表配置

癌种	监督任务	头类型
BRCA	ER/PR/HER2 状态	Binary
STAD	分级 G1/G2/G3	Ordinal
COADREAD	MSI/dMMR 状态	Binary
HNSC	分级 G1~G4	Ordinal
BLCA	基底/腔面亚型	Binary

表 B3 报告 5 折交叉验证下表型语义查询的预测准确率。该指标的合理范围由各任务的固有难度与代理标签噪声共同决定: 具有鲜明“分子-形态”对应关系的任务准确率较高(BRCA 的 ER/HER2 与 COADREAD 的 MSI 均超过 87%), 而组织学分级(STAD 为 59.6%、HNSC 为 52.4%)则受病理观察者间一致性的固有限制, 处于文献报道的人类病理医师一致性水平。该结果说明表型查询确实承载了与临床先验对齐的可读语义, 从侧面验证了表型语义辅助监督接口的有效性。

附录 C. 超参数敏感性分析

为评估 PathRoute 对架构与训练超参数的敏感性, 本节在 STAD、COADREAD、HNSC 三个癌种上对 3 个核心超参数进行扫描: 表型语义查询数 $Q_3 \in \{4,$

Table B3 Prediction Accuracy of Phenotype-semantic Queries

表 B3 表型语义查询预测准确率

癌种	任务	标签来源	准确率/%
BRCA	ER	PAM50	87.0±3.1
BRCA	PR	PAM50	76.5±7.8
BRCA	HER2	PAM50	91.7±2.7
STAD	分级	病理报告	59.6±10.2
COADREAD	MSI	CMS	88.3±1.0
HNSC	分级	病理报告	52.4±7.3
BLCA	亚型	mRNA	74.7±4.8
BLCA	分级	病理报告	95.7±5.3

注: “±”前后分别为 5 折均值±标准差。

8,16,32}、组织组成感知层数 $L_1 \in \{2,4,6,8\}$(推理 Transformer 共 12 层)、表型损失权重 $\lambda_{\text{phen}} \in \{0.25, 0.5,1.0,2.0\}$。每个扫描点采用与主实验相同的 5 折交叉验证。表 C1、表 C2 与表 C3 分别展示了 3 组扫描的敏感性, 其中标记为“*”的整行数值为默认配置。

Table C1 Sensitivity to the Number of Phenotype-semantic Queries Q_3

表 C1 表型语义查询数 Q_3 敏感性

Q_3	STAD	COADREAD	HNSC	均值
4	0.642±0.073	0.718±0.142	0.631±0.036	0.664
8	0.651±0.084	0.720±0.123	0.609±0.049	0.660
16*	0.665±0.075	0.717±0.143	0.646±0.075	0.676
32	0.642±0.084	0.728±0.153	0.618±0.052	0.663

注: “±”前后分别为 5 折 C-index 均值和标准差。标“*”的整行数据为默认配置。

Table C2 Sensitivity to the Number of Tissue-composition Sensing Layers L_1

表 C2 组织组成感知层数 L_1 敏感性

L_1	STAD	COADREAD	HNSC	均值
2	0.668±0.094	0.691±0.151	0.615±0.064	0.658
4*	0.665±0.075	0.717±0.143	0.646±0.075	0.676
6	0.645±0.078	0.730±0.124	0.605±0.037	0.660
8	0.654±0.100	0.720±0.159	0.590±0.045	0.655

注: “±”前后分别为 5 折 C-index 均值和标准差。标“*”的整行数据为默认配置。

整体上, 3 个超参数在合理取值范围内的最大波动不超过 0.02 C-index, 说明 PathRoute 对超参数不敏感。默认配置(Q_3=16, L_1=4, λ_{phen}=1.0)在 3 组扫描中均位于或接近峰值, 其他取值带来的差异均处于跨折标准差范围内。

Table C3 Sensitivity to the Phenotype Loss Weight λ_{phen}

表 C3 表型损失权重λ_{phen}敏感性

λ_{phen}	STAD	COADREAD	HNSC	均值
0.25	0.658±0.104	0.704±0.135	0.611±0.062	0.658
0.5	0.659±0.102	0.695±0.125	0.614±0.066	0.656
1.0*	0.665±0.075	0.717±0.143	0.646±0.075	0.676
2.0	0.656±0.091	0.713±0.146	0.601±0.038	0.657

注: “±”前后分别为 5 折 C-index 均值和标准差。标“*”的整行数据为默认配置。

附录 D. 主动证据路由与注意力池化的对比

我们在相同的 2 391 张切片、相同 UNI2 特征与 CONCH 伪标签下, 将 PathRoute 的路由注意力与代表性注意力池化方法 ABMIL 的区块级注意力进行逐切片对照, 并计算肿瘤集中度、归一化注意力熵、有效区块率和 top-1% 区块注意力质量 4 项指标, 如表 D1 所示。结果表明, 在 UNI2 强表征上, ABMIL 的单一门控注意力容易退化为近似均匀池化: 池化统计中其归一化注意力熵中位数为 0.96, 有效区块率约 0.70, 肿瘤集中度仅为 1.16, top-1% 区块仅承载 0.06 的注意力质量。相比之下, PathRoute 通过角色分化查询主动聚合证据, 池化统计中肿瘤集中度达到 2.30, top-1% 区块承载 0.68 的注意力质量, 说明其能够在海量区块中更主动、更精准地筛选关键诊断区域。PathRoute 较低的注意力熵与有效区块利

Table D1 per-slide Medians of Tile-level Attention: ABMIL vs Active Evidence Routing

表 D1 ABMIL 与主动证据路由在 per-slide 中位数的区块级注意力对照

癌种	方法	肿瘤集中度	注意力熵	有效区块率	top-1% 质量
BRCA	ABMIL	1.23	0.76	0.117	0.36
BRCA	PathRoute	2.04	0.49	0.011	0.82
STAD	ABMIL	1.06	0.99	0.931	0.03
STAD	PathRoute	1.90	0.75	0.122	0.36
COADREAD	ABMIL	1.47	0.96	0.733	0.05
COADREAD	PathRoute	2.84	0.63	0.040	0.58
HNSC	ABMIL	1.25	0.98	0.854	0.03
HNSC	PathRoute	2.56	0.60	0.031	0.64
BLCA	ABMIL	1.00	0.99	0.956	0.03
BLCA	PathRoute	2.69	0.57	0.019	0.73
池化 (2 391)	ABMIL	1.16	0.96	0.70	0.06
池化 (2 391)	PathRoute	2.30	0.59	0.025	0.68

用率并非“信息利用不足”，而是主动证据路由作为选择机制的预期且正确的行为：全切片中诊断证据本身高度稀疏，因此一个有效的证据选择器理应将注意力质量集中于少数诊断相关区块，而非在海量非诊断性背景上均匀分布；ABMIL 的高熵则源于其退化的近似均匀池化和丧失区分力（集中度仅为 1.16）。

这从“信息利用”和“关键区域筛选”两个维度定量支撑了主动证据路由的设计优势。更重要的是，这种更精准的证据选择并非以牺牲预测性能为代价：本文主结果中 PathRoute 在 5 个癌种的平均 C-index 为 0.666，显著高于 ABMIL 的 0.590。

附录 E. CONCH 组织伪标签的人工标注验证

我们在公开组织分类基准 CRC-VAL-HE-7K 上进行了人工标注验证。该数据集包含 7 180 个 224×224、0.5 μm/px（20×）H&E 组织块，具有病理专家标注，并包含肿瘤上皮类别。我们使用与正文 2.3 节完全相同的 CONCH 零样本分类管线，将预测与真值均二值化为“肿瘤/非肿瘤”，考察 CONCH-tumor 伪标签与专家标注的一致性。

结果如表 E1 所示，在排除背景的组织区块上，CONCH 判为肿瘤的区块中 99.1% 为专家标注的真实肿瘤（精确率为 0.991），肿瘤判别 AUROC 为 0.906。这说明本文用于路由语义对齐分析的 CONCH-tumor 伪标签是一个高纯度肿瘤标注。

Table E1 Agreement Between CONCH Tumor Pseudo-labels and Expert Annotations on CRC-VAL-HE-7K

表 E1 CONCH 肿瘤伪标签在 CRC-VAL-HE-7K 上与专家标注的一致性

评估子集	AUROC	精确率	准确率	召回率	*F*1
全部九类	0.864	0.786	0.882	0.430	0.556
组织区块（排除背景）	0.906	0.991	0.888	0.430	0.600

据此形成一条完整验证链：1）正文 3.6 节已表明 PathRoute 的路由注意力稳定对齐 CONCH-tumor 区块（集中度中位数为 2.30，$p<10^{-48}$）；2）附录 E 表明 CONCH-tumor 区块以约 99% 的纯度对应专家标注的真实肿瘤。因此，模型路由聚焦的区域是可传递地对应病理专家判定的真实肿瘤区域。