MiCo: 带有上下文感知聚类的多实例学习用于全幻 灯片图像分析

Junjian Li 1, Jin Liu $^{1,2\star},$ Hulin Kuang 1, Hailin Yue 1, Mengshen He 1, and Jianxin Wang 1,2

¹ Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, Changsha 410083, Hunan, China

² Xinjiang Engineering Research Center of Big Data and Intelligent Software, School of Software, Xinjiang University, Urumqi 830091, China

liujin06@csu.edu.cn

摘要 多实例学习(MIL)在组织病理学全切片图像(WSI)分析中的癌症诊断和预后方面显示出巨大的潜力。然而,WSIs内在的空间异质性提出了关键挑战,因为形态相似的组织类型通常分散在遥远的解剖区域。传统的MIL方法难以建模这些散布的组织分布并有效捕捉跨区域的空间交互作用。为了解决这些问题,我们提出了一种新的具有上下文感知聚类(MiCo)的多实例学习框架,旨在增强WSIs中的跨区域内组织相关性和强化组织间的语义关联。MiCo首先通过聚类实例来提炼区分性形态模式,其中聚类中心作为语义锚点。为了增强跨区域内的组织相关性,MiCo采用了一个集群路由模块,该模块通过特征相似性动态链接远距离区域的同一组织类型的实例。这些语义锚点充当上下文枢纽,传播语义关系以细化实例级别的表示。为消除语义碎片化并强化组织间的语义关联,MiCo整合了一个聚类缩减模块,该模块在合并冗余锚点的同时增强不同语义组之间的信息交换。在九个大规模公共癌症数据集上的两项具有挑战性的任务中进行的广泛实验展示了MiCo的有效性,并证明了其优于最先进的方法。代码可在https://github.com/junjianli106/MiCo获取。

Keywords: 多实例学习·整张幻灯片图像·上下文感知聚类。

1 介绍

全景图像(WSIs)提供了细胞形态和肿瘤微环境的全面详细表示,这两者对于准确的诊断评估和预后判断都至关重要[2,3,22,6]。然而,组织病理学

 $^{^{\}star}$ means the corresponding author.

WSIs 通常表现出极高的分辨率,可达到 100,000×100,000 像素。WSIs 的高分辨率以及缺乏像素级标注,在其分析和建模中带来了显著挑战 [12,14,9,8]。尽管多实例学习(MIL)已作为一种弱监督 WSI 分析范式出现 [1,20],现有方法仍然难以应对 WSIs 的空间复杂性。

这一挑战通过两个相互关联的现象表现出来:在宏观层面,同一类型的病理性显著组织结构往往在远端区域表现出不连续分布。例如,淋巴血管肿瘤栓子可能从瘤周区域延伸到远处的血管系统 [11],而沿神经束扩散的周围神经侵犯灶可以跨越多个解剖平面 [13]。在显微层面,这些分布式结构与异质性的细胞微环境相互作用,形成复杂的病理网络,需要同时建模局部判别特征和跨区域语义依赖。

尽管取得了显著的进步,现有的 MIL 方法仍然难以建模形态上距离遥远但病理相关的区域之间的长程依赖关系,这是由 WSIs 固有的空间异质性带来的挑战。例如,传统的基于注意力的 MIL 方法(如 AMIL[5])利用注意力机制动态识别和聚合关键实例,这些方法倾向于关注实例级别的特征学习,却无法捕捉到远距离病理区域之间的语义关系。基于聚类的方法,比如DeepAttnMISL[21],在集群内部学习补丁依赖性但未能捕捉组织区域间的上下文交互作用,这对预后至关重要。基于变换器的方法(如 TransMIL[15])使用变换器来学习 WSIs 的全局上下文信息,但将所有交互视为同质的,无法捕捉跨肿瘤微环境的动态异质性相互作用。基于图的方法(如 PatchGCN[1])依赖于固定的邻域定义,而分层方法如 HVTSurv[16] 将局部学习限制在预定义的子区域上,这两种方法都不足以建模跨越空间分布的组织级别的形态连续性。

为了克服这些限制,我们提出了 MiCo,一个具有上下文感知聚类的新型多实例学习框架。MiCo 采用了一个集群路由模块,通过聚合和传播同一组织类型分散补丁的信息来增强组织内部的语义关联,从而优化实例级别的表示。此外,MiCo 集成了一个集群缩减模块,以消除语义碎片化并强化组织之间的语义关联,在保持病理多样性的同时整合语义冗余的锚点。在九个大规模公开癌症数据集上的两个具有挑战性的任务中进行了广泛实验,验证了 MiCo 的有效性,并展示了其在生存预测和癌症亚型分类任务中的优越性能。

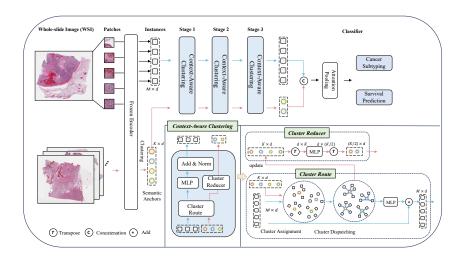


图 1: 概述 MiCo。MiCo 由多层上下文感知聚类模块组成。每个模块由一个 集群路由模块组织,该模块聚合并传播语义信息,以及一个集群缩减器模 块,该模块整合冗余锚点同时增强语义不同锚点之间的信息交换。

2 方法论

2.1 MiCo 概览

为了解决组织病理学 WSI 中的空间异质性挑战,我们提出了 MiCo, 一种具有语境感知聚类功能的新颖多示例学习框架。如图 1 所示,MiCo 由多层语境感知聚类模块组成,每个模块均由两个核心组件组织: Cluster Route (CluRoute) 模块和 Cluster Reducer (CluReducer) 模块。MiCo 开始时通过对实例进行聚类以提炼出具有辨别力的形态模式,簇中心作为语义锚点。CluRoute 基于特征相似性动态链接同一组织类型的实例,即使这些实例分布在遥远的区域,汇集并传播语义信息以优化实例级表示。这个过程通过建立独立于解剖边界组织关系来增强跨区域内的组织相关性。CluReducer 整合冗余锚点同时促进不同语义组之间的信息交换,从而消除语义碎片化并加强组织间的语义关联。这些组件协同工作,在区域内传播上下文信息,通过语义聚合优化实例级表示,并生成用于下游诊断和预后任务的稳健袋级别特征。

2.2 WSI 预处理

给定一个 WSI W_j ,我们首先将组织区域裁剪为不重叠的块 $\{p_{j,m}\}_{m=1}^M$ 。每个块使用多模态全幻灯片基础模型 TITAN[4] 编码成特征向量 $h_{j,m} \in \mathbb{R}^d$,该模型通过自监督和视觉语言目标进行了预训练。这个过程为每个 WSI 生成一个特征矩阵 $H_j \in \mathbb{R}^{M \times d}$ 。为了简化表示,我们省略下标 j。为了初始化语义锚点,我们对训练实例应用 K-均值聚类,获得聚类中心 $S = \{s_k\}_{k=1}^K$ 。这些中心作为代表不同组织模式的语义锚点。

2.3 上下文感知聚类

形态相似的组织类型通常分散在遥远的解剖区域,这使得捕捉它们的上下文关系和语义一致性变得具有挑战性。MiCo 通过其多层上下文感知聚类模块解决了这一问题,该模块由两个核心组件组成:聚类路径和聚类简化器。聚类路径通过连接远距离区域中的相同组织类型实例来聚合和传播语义信息,而聚类简化器则整合冗余锚点并促进不同语义组之间的信息交换。

聚类路由。给定一个包含 M 实例的 WSI,每个补丁 p_m 都由特征嵌入 $h_m \in \mathbb{R}^d$ 表示。为了建立跨区域的组织内相关性,我们计算每个实例特征 h_i 和 K 可 学习语义锚点 $\{s_k\}_{k=1}^K$ 之间的余弦相似度:

$$A_{l(m,k)} = \frac{h_m^{\top} s_k}{|h_m| \cdot |s_k|},\tag{1}$$

其中 $A_l \in \mathbb{R}^{M \times K}$ 衡量了实例与语义锚点之间的语义对齐程度。

argmax 操作确保每个实例都被分配到与其最相关的语义锚点,从而在 病理结构中保持语义一致性。然而, argmax 是非可微的,这会中断梯度流 向锚点。为了解决这一问题,我们使用一个直通估计器 [18] 来实现实例到锚 点的分配同时维持梯度传播:

$$\hat{A}_l = \text{one-hot} \left(\arg \max(A_l) \right) + A_l - \operatorname{sg}(A_l), \tag{2}$$

其中 $sg(\cdot)$ 表示停止梯度运算符。在前向传播过程中, $\hat{A}_l \in \mathbb{R}^{M \times K}$ 是一个二进制矩阵,每一行都是一热编码,将每个实例分配给与其最相似的簇。在反向传播过程中,梯度绕过非可微的 $arg \max$ 操作并通过 A_l 传递,从而实现端到端训练。

对于每个语义锚点 s_k ,我们聚合其分配实例的特征以生成一个上下文感知表示:

$$\tilde{s}_k = \frac{1}{N_k} \sum_{m=1}^{M} \hat{A}_{l(m,k)} \cdot h_m, \quad N_k = \sum_{m=1}^{M} \hat{A}_{l(m,k)}.$$
 (3)

其中 $\tilde{s}_k \in \mathbb{R}^d$ 捕获锚组内共享的形态语义(例如肿瘤间质或免疫浸润),而 N_k 表示分配给第 k 个锚点的实例数量。此聚合步骤统一了空间上分散但 具有相似语义的实例,直接减轻了空间异质性。我们将所有 \tilde{s}_k 收集到矩阵 $\tilde{S} \in \mathbb{R}^{K \times d}$ 中,该矩阵作为聚类缩减模块的输入。

为了传播跨区域的上下文,我们将每个实例特征 h_m 与其分配的语义锚点语义对齐:

$$h'_{m} = h_{m} + MLP(h_{m} + \hat{A}_{l(m,k)} \cdot \tilde{s}_{k}).$$
 (4)

更新后的特征 h_m' 结合了实例级特征与跨区域的语义上下文,形成了空间分散组织结构的统一表示。该增强的特征随后传递到下一个上下文感知聚类模块进行进一步优化,使模型能够逐步捕捉复杂的解剖关系。

聚类简化器。为了减少语义冗余,Cluster Reducer 通过建模锚点之间的关系来合并具有相似特征的锚点。给定初始语义锚点 $\tilde{S} \in \mathbb{R}^{K \times d}$,我们把 \tilde{S} 转置为 $\tilde{S}^{\mathsf{T}} \in \mathbb{R}^{d \times K}$ 并应用多层感知器(MLP)来建模非线性交互:

$$\tilde{S}' = MLP(\tilde{S}^{\top}), \tag{5}$$

其中, $\tilde{S}' \in \mathbb{R}^{d \times K/2}$ 表示冗余减少后的优化锚点。多层感知器捕捉语义依赖关系,例如通过整合语义相似的锚点来合并分散的肿瘤浸润淋巴细胞簇。将 \tilde{S}' 转置回 $\tilde{S} \in \mathbb{R}^{K/2 \times d}$ 可以得到紧凑且有意义的锚点,这些锚点保留了组织的基本特征。

3 实验与结果

3.1 数据集

生存预测。我们在实验中使用了来自 TCGA 的七个公开可用的癌症数据集 (¹ (**膀胱癌腺癌**, **BRCA**, **GBMLGG**, **HNSC**, **肾细胞癌**, **肾乳头细胞癌**和 **肺癌腺癌**)。这些数据集共同包含了来自 3,523 名患者和 4,091 张 H&E 诊断 全幻灯片图像的数据。

癌症亚型分类。我们在两个具有挑战性的公共数据集上进行了比较实验: TCGA-BRCA 和 TCGA-非小细胞肺癌。TCGA-BRCA 数据集包含 1,034 张 H&E 切片,涉及两种侵袭性癌症亚型:浸润性导管癌 (IDC) 和浸润性小叶癌 (ILC)。TCGA-NSCLC 数据集包括来自两种亚型的 1,030 张 H&E 切片:肺鳞状细胞癌 (TCGA-LUSC) 和肺腺癌 (TCGA-LUAD)。

¹ https://portal.gdc.cancer.gov

BLCA BRCA GBMLGG HNSC KIRC Models KIRP LUAD MEAN AMIL [5] $0.551_{0.10} \ \ 0.576_{0.11} \ \ 0.736_{0.15} \ \ 0.590_{0.02} \ \ 0.681_{0.05} \ \ 0.778_{0.04} \ \ 0.584_{0.04}$ 0.642TransMIL [15] $0.610_{0.07} \ \ 0.555_{0.04} \ \ \ 0.760_{0.14} \ \ \ 0.565_{0.02} \ \ 0.694_{0.05} \ \ 0.773_{0.07} \ \ 0.585_{0.09}$ 0.649 DeepGraphConv [10] $0.563_{0.08} \ \ 0.581_{0.10} \ \ \ 0.739_{0.14} \ \ \ 0.591_{0.02} \ \ 0.681_{0.02} \ \ 0.738_{0.04} \ \ 0.581_{0.02}$ 0.639 PatchGCN [1] $0.607_{0.05}$ $0.600_{0.12}$ $0.736_{0.16}$ $0.584_{0.01}$ $0.694_{0.05}$ $0.783_{0.05}$ $0.582_{0.04}$ 0.655ILRA [20] $0.612_{0.04} \ \ 0.551_{0.07} \ \ 0.722_{0.16} \ \ 0.600_{0.01} \ \ 0.663_{0.04} \ \ 0.688_{0.05} \ \ 0.590_{0.04}$ 0.632 HVTSurv [16] $0.606_{0.05} \ \ 0.579_{0.09} \ \ 0.792_{0.01} \ \ 0.575_{0.03} \ \ 0.692_{0.04} \ \ 0.778_{0.06} \ \ 0.591_{0.03}$ RRTMIL [17] $0.588_{0.05} \ \ 0.542_{0.08} \ \ \ 0.738_{0.16} \ \ \ 0.597_{0.02} \ \ 0.699_{0.04} \ \ 0.770_{0.06} \ \ 0.579_{0.04}$ 0.644 $0.607_{0.07} \ \ 0.582_{0.09} \ \ \ 0.752_{0.16} \ \ \ 0.579_{0.02} \ \ 0.704_{0.06} \ \ 0.763_{0.09} \ \ 0.595_{0.06}$ WiKG [7] 0.654微协 $\mathbf{0.619}_{0.03} \ \mathbf{0.608}_{0.12} \ \mathbf{0.813}_{0.02} \ \mathbf{0.605}_{0.01} \ \mathbf{0.715}_{0.05} \ \mathbf{0.800}_{0.07} \ \mathbf{0.604}_{0.04}$

表 1: 七种癌症类型中的生存预测性能比较。

表 2: 两种癌症类型中的癌症亚型分类性能比较。

方法	TCGA-BRCA			TCGA-肺癌特殊细胞类型癌症			平均值		
	ACC	F1	AUC	ACC	F1	AUC	ACC	F1	AUC
AMIL [5]	0.9090.01	$0.858_{0.02}$	$0.948_{0.01}$	$0.923_{0.02}$	$0.923_{0.02}$	$0.979_{0.01}$	0.916	0.890	0.964
TransMIL [15]	$0.906_{0.01}$	$0.857_{0.02}$	$0.945_{0.02}$	$0.919_{0.02}$	$0.919_{0.02}$	$0.977_{0.02}$	0.913	0.889	0.961
${\rm DeepGraphConv}\ [{\color{blue}10}]$	$0.913_{0.01}$	$0.859_{0.01}$	$0.947_{0.01}$	$0.913_{0.02}$	$0.913_{0.02}$	$0.969_{0.02}$	0.913	0.886	0.958
PatchGCN [1]	$0.917_{0.01}$	$0.871_{0.02}$	$0.944_{0.02}$	$0.919_{0.02}$	$0.919_{0.02}$	$0.979_{0.01}$	0.918	0.895	0.962
ILRA [20]	$0.899_{0.01}$	$0.836_{0.03}$	$0.931_{0.02}$	$0.918_{0.02}$	$0.918_{0.02}$	$0.973_{0.02}$	0.909	0.885	0.952
RRTMIL [17]	$0.910_{0.01}$	$0.855_{0.01}$	$0.945_{0.02}$	$0.915_{0.02}$	$0.915_{0.02}$	$0.976_{0.01}$	0.913	0.890	0.961
WiKG [7]	$0.915_{0.02}$	$0.859_{0.03}$	$0.948_{0.01}$	$0.920_{0.03}$	$0.920_{0.03}$	$0.980_{0.01}$	0.917	0.889	0.964
微协	$0.922_{0.01}$	$0.875_{0.02}$	$0.952_{0.01}$	$0.931_{0.02}$	$0.931_{0.02}$	$0.981_{0.01}$	0.927	0.903	0.967

3.2 实现细节

我们将MiCo与八种最先进的方法进行了比较: AMIL[5], TransMIL[15], DeepGraphConv[10], Patch-GCN[1], ILRA[20], HVTSurv[16], RRTMIL[17] 和 WiKG[7]。WSIs 使用 TITAN[4] 进行预处理,以提取 448×448 像素在 20×倍放大下的特征。所有实验均采用 4 折交叉验证,数据集按 60:15:25 的比例划分为训练集、验证集和测试集。语义锚点数量设置为 64。训练配置在各种方法中统一:200 个周期,批次大小 1,学习率 2e-4(Ranger 优化器 [19]),提前停止 8,每 2 步进行一次梯度累积。生存预测性能使用Concordance Index (C-Index) 及其标准差 (std) 进行评估,而癌症亚型分类则通过 Accuracy (ACC)、F1-score (F1) 和 Area Under the Curve (AUC) 指标进行评价。

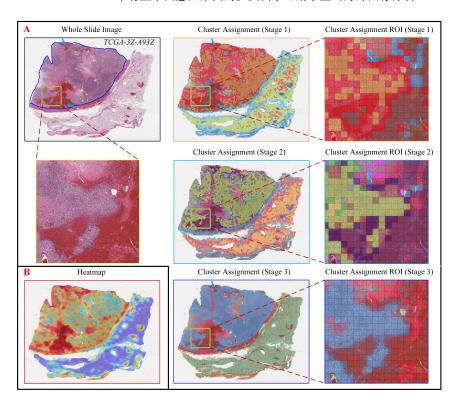


图 2: **A.** 聚类分配的定性结果,展示了 WSI 缩略图、聚类分配结果以及不同 阶段对应的感兴趣区域 (ROI)。蓝色轮廓表示真实的肿瘤区域,而被分配到 同一聚类的区域则用相同的颜色显示。**B.**MiCo 的可解释性分析,热图中的 红色区域指示高注意力区域。

3.3 结果与讨论

生存预测。实验结果表明,MiCo 在七种癌症类型的生存预测任务中表现出色。如表 1 所示,MiCo 达到了 0.680 的平均 C 指数,超越了所有其他方法。这一表现突显了 MiCo 上下文感知聚类在应对组织病理学全玻片图像空间异质性挑战方面的有效性。MiCo 能够建模跨区域依赖关系并增强分散组织区域之间的语义关联,使其能够捕捉复杂的形态模式并提高生存预测的准确性。

癌症亚型分类。表 2 中所示的癌症亚型实验结果表明,MiCo 在两个主要数据集 TCGA-BRCA 和 TCGA-NLCSC 上表现出色。在准确率、F1 分数和 AUC 方面,MiCo 达到了最先进的结果,超越了所有其他方法。凭借 0.927

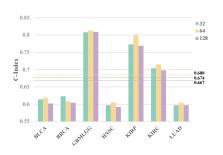


图 3: 语义锚点对生存预测的影响分析。

表 3: MiCo 的消融研究平均结果。

方法	Survival Prediction	癌症亚型分类			
77 144	C-Index	ACC	F1	AUC	
w/o Semantic Anchors Init	0.668	0.922	0.896	0.964	
w/o CluReducer	0.659	0.921	0.897	0.965	
w/o CluRoute	0.656	0.910	0.890	0.955	
微协	0.680	0.927	0.903	0.967	

的平均 ACC, MiCo 显著超过了最佳基线 PatchGCN, 后者达到了 0.918 的 平均 ACC。这些结果突显了 MiCo 凭借其上下文感知聚类机制和动态语义 锚点准确分类癌症亚型的能力,有效地缓解了组织病理学 WSI 中的空间异质性和语义碎片化挑战。

消融研究。为了进一步确定 MiCo 中每个关键组件的有效性,我们在生存预测和癌症亚型分类任务上进行了一系列消融研究。如表 3 所总结的平均结果强调了每个模块在提升 MiCo 整体性能方面的重要性。这些消融实验评估了移除关键组件的影响,包括语义锚点初始化、CluReducer 和 CluRoute。移除语义锚点初始化将 C 指数降低到 0.668,突显了其在捕捉判别性形态模式中的重要性。省略 CluReducer 进一步降低了 C 指数至 0.659,强调了它在整合冗余锚点和增强锚间信息交换方面的作用。同样地,移除 CluRoute 导致 C 指数为 0.656,证明了其建立跨区域语义关系并细化实例级表示的有效性。这些结果验证了每个模块对 MiCo 整体性能的关键贡献。

可解释性分析。如图 2 所示,我们通过聚类分配和热图分析了 MiCo 的可解释性。在图 2.A 中,语义相关的区域逐渐合并,展示了 MiCo 改进表示能力和增强语义连贯性的能力。热图结果进一步揭示了与图 2.B 中的真实肿瘤注释的强对齐性,证实了 MiCo 在肿瘤定位方面的高精度。

语义锚点分析。如图 3 所示,我们使用不同数量的语义锚点(32、64 和 128)评估 MiCo 的生存预测性能。平均 C-index 值分别为 0.674(32 个锚点)、0.680(64 个锚点)和 0.667(128 个锚点)。64 个锚点配置实现了最高的平均 C-index,平衡了形态多样性和冗余性。

4 结论

本文提出了一种新的具有上下文感知聚类的多实例学习框架 MiCo,旨在解决组织病理学 WSI 中的空间异质性问题。MiCo 包含一个 Cluster Route模块以加强跨区域的组织内相关性,并且包含一个 Cluster Reducer模块以整合冗余锚点并促进不同语义组之间的信息交换。在九个癌症数据集上的广泛实验验证了 MiCo 在生存预测和癌症亚型分类任务中优于最先进的方法。致谢。2 本工作部分得到了新疆维吾尔自治区重点研发计划(编号 2024B03039-3)、国家自然科学基金(编号 U24A20256)、湖南省教育厅科学研究基金(编号 23A0020)、长沙市科技重大专项(编号 kh2502004)、中央引导地方科技发展资金(编号 ZYYD2025QY25)、中南大学创新驱动研究计划(编号 2023CXQD018)以及中南大学中央高校基本科研业务费专项资金(编号 2024ZZTS0108)的支持。本工作部分使用了中南大学高性能计算中心的计算资源。

参考文献

- Chen, R.J., Lu, M.Y., Shaban, M., Chen, C., Chen, T.Y., Williamson, D.F., Mahmood, F.: Whole slide images are 2d point clouds: Context-aware survival prediction using patch-based graph convolutional networks. In: Medical Image Computing and Computer Assisted Intervention—MICCAI 2021: 24th International Conference, Strasbourg, France, September 27—October 1, 2021, Proceedings, Part VIII 24. pp. 339–349. Springer (2021)
- Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., Madabhushi, A.: Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: Medical imaging 2014: Digital pathology. vol. 9041, p. 904103. SPIE (2014)
- Das, K., Conjeti, S., Roy, A.G., Chatterjee, J., Sheet, D.: Multiple instance learning of deep convolutional neural networks for breast histopathology whole slide classification. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). pp. 578–581. IEEE (2018)
- Ding, T., Wagner, S.J., Song, A.H., Chen, R.J., Lu, M.Y., Zhang, A., Vaidya, A.J.,
 Jaume, G., Shaban, M., Kim, A., Williamson, D.F.K., Chen, B., Almagro-Perez,
 C., Doucet, P., Sahai, S., Chen, C., Komura, D., Kawabe, A., Ishikawa, S., Gerber,

² 作者声明与本文内容无关的竞争利益。

- G., Peng, T., Le, L.P., Mahmood, F.: Multimodal whole slide foundation model for pathology (2024), https://arxiv.org/abs/2411.19666
- 5. Ilse, M., Tomczak, J., Welling, M.: Attention-based deep multiple instance learning. In: International conference on machine learning. pp. 2127–2136. PMLR (2018)
- 6. Lee, M.: Recent advancements in deep learning using whole slide imaging for cancer prognosis. Bioengineering **10**(8), 897 (2023)
- Li, J., Chen, Y., Chu, H., Sun, Q., Guan, T., Han, A., He, Y.: Dynamic graph representation with knowledge-aware attention for histopathology whole slide image analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11323–11332 (2024)
- 8. Li, J., Kuang, H., Liu, J., Yue, H., Wang, J.: Ca2cl: Cluster-aware adversarial contrastive learning for pathological image analysis. IEEE Journal of Biomedical and Health Informatics (2025)
- Li, J., Liu, J., Yue, H., Cheng, J., Kuang, H., Bai, H., Wang, Y., Wang, J.: Darc: Deep adaptive regularized clustering for histopathological image classification. Medical image analysis 80, 102521 (2022)
- Li, R., Yao, J., Zhu, X., Li, Y., Huang, J.: Graph cnn for survival analysis on whole slide pathological images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 174–182. Springer (2018)
- 11. Lia, M., Horn, L.C., Sodeikat, P., Höckel, M., Aktas, B., Wolf, B.: The diagnostic value of core needle biopsy in cervical cancer: A retrospective analysis. Plos one 17(1), e0262257 (2022)
- Lipkova, J., Chen, T.Y., Lu, M.Y., Chen, R.J., Shady, M., Williams, M., Wang, J., Noor, Z., Mitchell, R.N., Turan, M., et al.: Deep learning-enabled assessment of cardiac allograft rejection from endomyocardial biopsies. Nature medicine 28(3), 575–582 (2022)
- Medvedev, O., Hedesiu, M., Ciurea, A., Lenghel, M., Rotar, H., Dinu, C., Roman, R., Termure, D., Csutak, C.: Perineural spread in head-and-neck malignancies: Imaging findings—an updated literature review. Bosnian Journal of Basic Medical Sciences 22(1), 22 (2021)
- 14. Qu, L., Liu, S., Liu, X., Wang, M., Song, Z.: Towards label-efficient automatic diagnosis and analysis: a comprehensive survey of advanced deep learning-based weakly-supervised, semi-supervised and self-supervised techniques in histopathological image analysis. Physics in Medicine & Biology 67(20), 20TR01 (2022)
- 15. Shao, Z., Bian, H., Chen, Y., Wang, Y., Zhang, J., Ji, X., et al.: Transmil: Transformer based correlated multiple instance learning for whole slide image classification. Advances in neural information processing systems **34**, 2136–2147 (2021)

- Shao, Z., Chen, Y., Bian, H., Zhang, J., Liu, G., Zhang, Y.: Hvtsurv: hierarchical vision transformer for patient-level survival prediction from whole slide image.
 In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 2209–2217 (2023)
- Tang, W., Zhou, F., Huang, S., Zhu, X., Zhang, Y., Liu, B.: Feature re-embedding: Towards foundation model-level performance in computational pathology. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11343–11352 (2024)
- Van Den Oord, A., Vinyals, O., et al.: Neural discrete representation learning.
 Advances in neural information processing systems 30 (2017)
- 19. Wright, L.: Ranger-a synergistic optimizer. GitHub Repos. Available online at: https://github.com/lessw2020/Ranger-Deep-Learning-Optimizer (2019)
- 20. Xiang, J., Zhang, J.: Exploring low-rank property in multiple instance learning for whole slide image classification. In: The Eleventh International Conference on Learning Representations (2023)
- Yao, J., Zhu, X., Jonnagaddala, J., Hawkins, N., Huang, J.: Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. Medical image analysis 65, 101789 (2020)
- 22. Zarella, M.D., Bowman, D., Aeffner, F., Farahani, N., Xthona, A., Absar, S.F., Parwani, A., Bui, M., Hartman, D.J.: A practical guide to whole slide imaging: a white paper from the digital pathology association. Archives of pathology & laboratory medicine 143(2), 222–234 (2019)