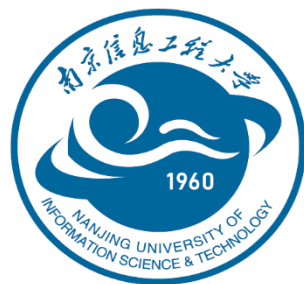


分类号: _____
密 级: _____

单位代码: 10300
学 号: 20181222025

南京信息工程大学 硕士学位论文



基于病理图像分析的癌症恶性程度自动评估

申请人姓名: 闫朝阳

指导教师: 徐军 教授

学科名称: 控制科学与工程

研究方向: 病理图像计算

培养学院: 自动化学院

提交时间: 2021.06.20

二〇二一年六月

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。本论文尽我所知，除了文中特别加以标注和致谢的内容外，不包含其他人或其他机构已经发表或撰写过的研究成果，也不包含为获得南京信息工程大学或其他教育机构的学位或证书而使用过的材料。与我一同工作同志对本研究所做的贡献均已在论文中作了声明并表示谢意。

本人学位论文及涉及相关资料若有不实，愿意承担一切相关的法律责任。

学位论文作者签名：_____ 签字日期：_____

关于论文使用授权的说明

本人授权南京信息工程大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档；允许论文被查阅和借阅；可以将学位论文的全部或部分内容编入有关数据库进行检索；可以采用影印、缩印或扫描等复制手段保存、汇编本学位论文。本文电子文档的内容和纸质论文的内容相一致。论文的公布（包括刊登）授权南京信息工程大学研究生院办理。

非国家秘密类涉密学位论文在解密后适用本授权书。

公开 保密（_____年_____月）（保密的学位论文在解密后应遵守此协议）

学位论文作者签名：_____ 签字日期：_____

指导教师签名：_____ 签字日期：_____

目 录

摘 要.....	I
Abstract	II
第一章 绪论.....	1
1.1 研究背景及意义	1
1.2 国内外研究现状	3
1.3 论文的主要内容及章节安排	6
第二章 相关理论基础.....	7
2.1 腺癌临床分级病理学诊断标准	7
2.2 同源性分布概述	8
2.3 多任务学习概述	9
第三章 基于同源性分布统计表示的前列腺癌自动格里森分级方法.....	13
3.1 研究动机	13
3.2 研究方法	14
3.2.1 病理图像预处理.....	14
3.2.2 同源性分布统计表示.....	15
3.2.3 分类器建模.....	18
3.3 实验设计	19
3.3.1 实验数据.....	19
3.3.2 对比方法设计.....	19
3.3.3 性能评估.....	20
3.4 实验结果及讨论	20
3.4.1 图像块层面的评估分析.....	20
3.4.2 切片层面的评估分析.....	21
3.4.3 统计特征分析.....	23
3.4.4 敏感性分析.....	25
3.4.5 全景病理图像预测.....	26
3.5 本章小结	27
第四章 基于多任务学习和先验知识的结直肠腺癌良恶自动评估方法.....	29

4.1	研究动机	29
4.2	研究方法	29
4.2.1	特征学习骨干网络	30
4.2.2	分割分类多任务分支	31
4.2.3	腺体结构空间注意力约束	32
4.2.4	多任务损失	32
4.3	实验设计	33
4.3.1	实验数据	33
4.3.2	实验细节	34
4.3.3	对比方法设计	34
4.3.4	性能评估	35
4.4	实验结果及讨论	35
4.4.1	相关方法自动分级对比	35
4.4.2	多任务学习和先验知识的有效性评估	36
4.4.3	先验知识有效性的可视化分析	38
4.4.4	分割分支评估	40
4.5	本章小结	41
第五章	总结与展望	43
5.1	本文工作内容及创新之处	43
5.2	今后工作的展望	43
致 谢	45
参考文献	46
作者简介	52

摘要

腺癌是一种常发生于上皮腺体组织的癌症类型，在结直肠癌、前列腺癌、乳腺癌、肺癌等多种恶性肿瘤中都十分常见。在病理形态上，腺癌的发生往往伴随着腺体结构分化变差甚至不分化，即腺体组织呈现出异常变形扭曲、内部空腔结构被细胞核侵占致其缩小甚至消失。这种不同程度的腺体分化直接与腺癌的恶性程度相关联。因此，在临床病理诊断中，腺体组织结构的分化程度是病理学家确定腺癌等级乃至决策治疗方案的决定性因素。

为了实现腺癌恶性程度的自动评估分级，本文从腺体结构分化程度这一病理诊断标准出发：一方面，本文设计了领域特定的手工特征，提出了基于同源性分布统计表示的前列腺癌自动格里森分级方法。该方法首先基于同源性分布算法计算病理图像的同源性分量，通过描述腺体周围细胞核间的拓扑排列对腺体结构分化程度进行定量化；然后用统计量对获得的同源性序列特征进行二次表征；最后采用加权 K 近邻分类器算法建模实现前列腺癌病理图像的自动分级。实验结果表明，相比无监督学习方法 SSAE、有监督方法 DLGg 和传统病理组学方法 MATF，本文方法分级表现更出色，特征也具有更好的鲁棒性。而且，本文提出的特征表示方法具有生物学基础，可解释性强。

另一方面，启发于对腺体结构自动分割结果进行精确描述以实现诊断的思路，本文提出基于多任务学习和先验知识的腺癌良恶自动评估方法。该方法首先利用特征学习骨干网络进行特征提取，提取到的特征分别被送入分割分支和分类分支进行腺体结构自动分割和腺癌图像的自动良恶分级；同时，分割分支的腺体结构预测作为先验知识被编码为空间注意力融合到分类分支中以约束分级的推理偏好。实验结果验证了本文基于先验知识约束的多任务网络模型的有效性，在腺癌测试集上达到最高 97.04% 的准确率和 0.9971 的 AUC 值。而且，本文的先验知识约束思想同样具有解释性。

本文研究的基于病理图像分析的癌症恶性程度自动评估分级方法，立足自腺体结构分化这一病理诊断标准，因而可以为病理医生提供可解释的辅助诊断支持。

关键词：腺癌，自动分级，同源性分布，先验知识，多任务学习

Abstract

Adenocarcinoma is a type of cancer that often occurs in epithelial gland tissues. It is common in malignant tumors such as colorectal cancer, prostate cancer, breast cancer, and lung cancer. In terms of pathological pattern, the occurrence of adenocarcinoma is often accompanied by poor or undifferentiated gland structure. The glands are abnormally stretched and distorted, and the cavity structures are invaded by the nuclei, causing it to shrink or even disappear. This situation is directly related to the degree of malignancy of adenocarcinoma. Therefore, the degree of differentiation of glandular structure is a decisive factor for pathologists to determine the grade of adenocarcinoma and decide the treatment plan in clinical pathological diagnosis.

In this paper, to realize the automated grading of the malignant degree of adenocarcinoma, the study evolves from the degree of gland structure differentiation which is a pathological diagnostic criteria. On the one hand, this paper designs domain-specific hand-crafted features and proposes an automated gleason grading method for prostate adenocarcinoma based on the statistical representation of homology profile. This method first calculates the homology components of pathological images based on the homology profile algorithm, and quantifies the degree of gland structure differentiation by describing the topological arrangement of nuclei around the gland. Then it employs statistical methods for secondary characterization on the obtained homology sequence characteristics. The method finally models the weighted K-nearest neighbor classifier algorithm for automated gleason grading of pathological images on adenocarcinoma. Experimental results show that this method performs better in grading and features robustness, compared with unsupervised learning method SSAE, supervised method DLGg and traditional pathomics-based method MATF. In addition, the proposed homology feature representation method has a strong biological basis and is highly interpretable.

On the other hand, inspired by the idea of accurately describing the automatic segmentation of glandular structures for diagnosis, this paper proposes an automated assessment method for colorectal adenocarcinoma based on multi-task learning and prior knowledge. Firstly, this method employs the backbone network for feature extraction, and the extracted features are respectively sent to the segmentation branch and classification branch for automated gland structure segmentation and grading of adenocarcinoma images. Meanwhile, the automated

gland structure prediction from the segmentation branch acts as the prior knowledge, and it is encoded as spatial attention and merged into the classification branch to constrain the reasoning. The experimental results verify the effectiveness of the prior knowledge-aware multi-task network, achieving the highest accuracy of 97.04% and AUC value of 0.9971 on the test set. Moreover, the idea of prior knowledge constraint in this paper is also interpretable.

The automated assessment grading methods based on pathological image analysis are proposed for cancer malignancy. Importantly, the glandular structure differentiation acting as the pathological diagnosis criteria is fundamental to them. The methods thus can provide pathologists with interpretable auxiliary diagnosis support.

Key words: Adenocarcinoma, Automated grading, Homology profile, Prior knowledge, Multi-task learning

第一章 绪论

1.1 研究背景及意义

癌症是世界范围内主要的公共卫生问题，关系着每一个人的健康。《2020 年全球癌症报告》^[1]显示，我国癌症发病率、死亡率均居全球首位。全球共有 57 个国家癌症为第一大死亡原因，如图 1.1 所示。据统计，全球发病率最高的 5 种癌症依次是乳腺癌、肺癌、结直肠癌、前列腺癌、胃癌。其中，腺癌作为一种发生于上皮涎腺组织的癌症类型，在结直肠癌、前列腺癌、乳腺癌、肺癌等多种恶性肿瘤中都十分常见^[2]。近年来，腺癌的发病率和死亡率都呈现逐年增高的趋势，严重危害到患者的身心健康。

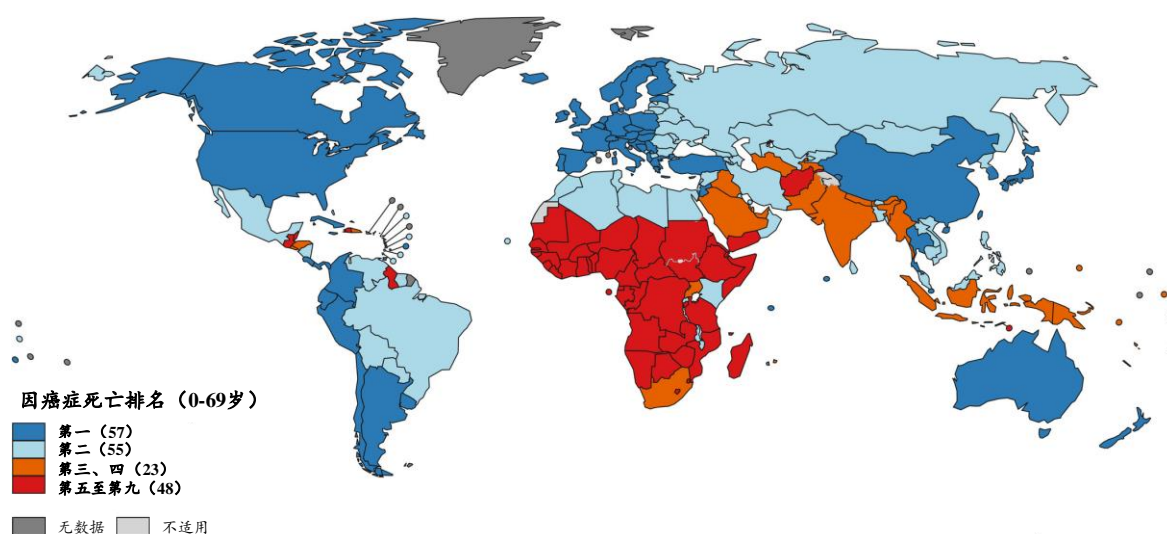


图 1.1 2019 年各国家年龄在 70 岁以下因癌症导致死亡的占比排名。其中，在中国、日本、美国等 57 个国家中癌症为第一大死亡原因。注 a：数据来源于世界卫生组织^[1]

腺癌常发生在腺体组织，而腺体是大部分器官系统中重要的组织结构。腺体的主要工作机制是分泌蛋白质和碳水化合物，保证细胞增殖、分化和凋亡之间的循环协调，以保护不同器官及组织免受不利的化学和物理环境的侵害^[3]。在病理形态上，上皮组织中分化良好的正常腺体结构通常是由内部的近圆空腔结构、中间的围绕在近圆空腔结构周围的柱状上皮细胞和外部的位于柱状上皮细胞基部的细胞核组成^[4]。而在细胞增殖、分化、凋亡的循环协调过程中，一旦腺上皮癌细胞增生使得这种平衡被破坏，腺上皮细胞再生的完整性丧失，便导致了腺癌^[5]。此时的腺体结构出现分化变差甚至不分化的情况，在病理形态上表现为近圆空腔结构异常变形扭曲、空腔结构被细胞核侵占导致其缩小甚至消失，如图 1.2 所示。这种不同程度的腺体分化直接与腺癌的恶性程度相关联。因此，在临床病理诊断中，腺体组织结构的分化程度是病理学家确定腺癌恶性程度、观察预后

状态乃至决策治疗方案的决定性因素，也是肿瘤学临床实践指南中的重要特征^[6-8]。如在前列腺癌中，格里森分级系统^[9]是进行前列腺癌诊断和风险评估的最广泛使用的分级方案^[10]，其根据腺体分化模式的差异确定不同等级的恶性程度（1至5级；其中，1级代表良好，5级代表最差，3级4级最为常见）。

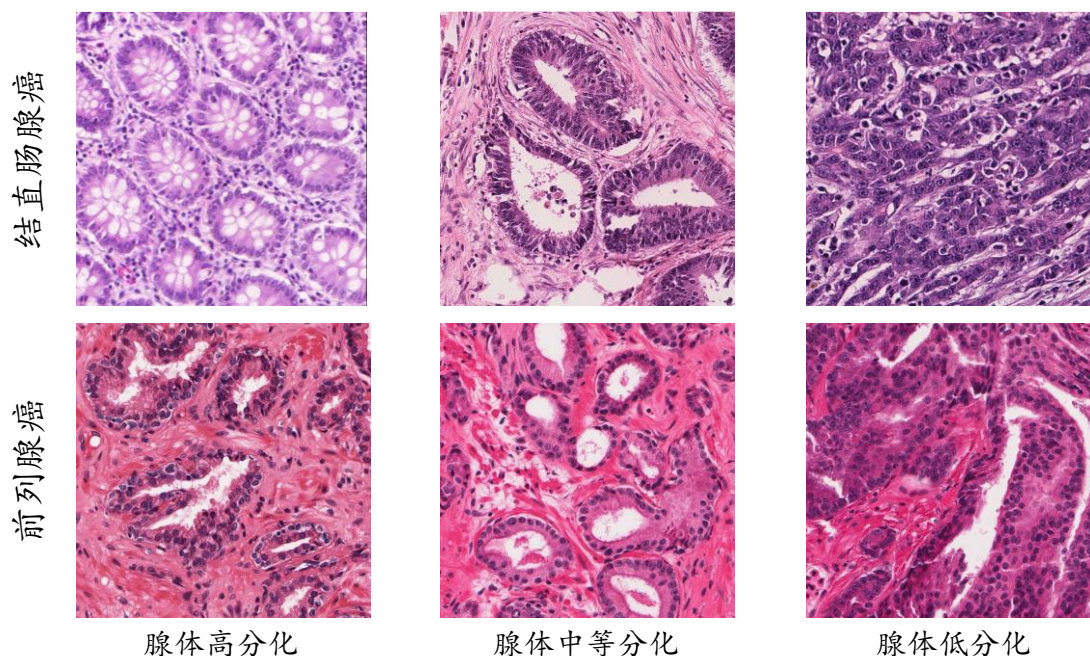


图 1.2 不同器官下不同恶性程度的腺癌伴随的腺体结构分化程度差异。其中，第一行为结直肠腺癌病理图像，第二行为前列腺癌病理图像。

在临床腺癌诊断过程中，病理医生通常依据肉眼对显微镜下组织病理切片进行观察，依据不同上皮组织区域的腺体结构分化程度来决策患者的恶性程度。整个过程非常繁琐而且耗时、费力，且医生的判断容易受到病理切片质量、染色等客观因素影响而产生偏差，这极大加重了医生的负担。甚至，同一病理医生在不同环境下的情况、不同医生的专业水平和经验等因素也影响着最终的诊断结果，这些问题使得人工诊断具有较大的主观性。因此，针对腺癌临床分级痛点，亟需一种快速、有效且标准化的计算机自动分析方法。

随着人工智能领域的快速发展，机器学习算法、特征描述工程、深度学习算法等各种新技术层出不穷，被应用到零售、安防、教育、医疗、金融、物流、自动驾驶等各个行业^[11]。尤其对于医疗行业，2020年，AlphaFold算法成功根据基因序列预测出蛋白质的三维结构^[12]，在结构生物学领域和医疗相关其他领域产生了巨大的影响。近年来，数字扫描技术的成熟，使得利用人工智能方法对组织病理切片图像进行定量化分析实现恶性程度自动评估成为可能。以数字病理^[13-14]为核心的计算病理领域^[15]受到越来越多研究

者的关注与重视，相关的计算病理方法^[15-18]被提出用于研究数字化组织病理图像并解决疾病诊断、预后、治疗反应预测等问题。

尽管基于病理图像分析进行自动诊断评估具有很多优势，但由于数字化的全景病理图像具有尺寸大、分辨率高等特点，使得病理图像分析十分具有挑战性，如图 1.3 中病理切片信息所示。而且，切片制片、染色和不同扫描设备成像质量差异也会带来不可忽略的影响。此外，就病理图像内容而言，如图 1.3 所示，包含有肿瘤上皮、间质、坏死、淋巴等多种组织，结构复杂、形态差异大。更为困难的是，腺癌发生过程中腺体组织呈现连续动态发展状态^[19]，病理表现为同一等级肿瘤差异性高，不同等级肿瘤相似性大，这给病理图像分析方法和自动诊断目标增加了更大难度。

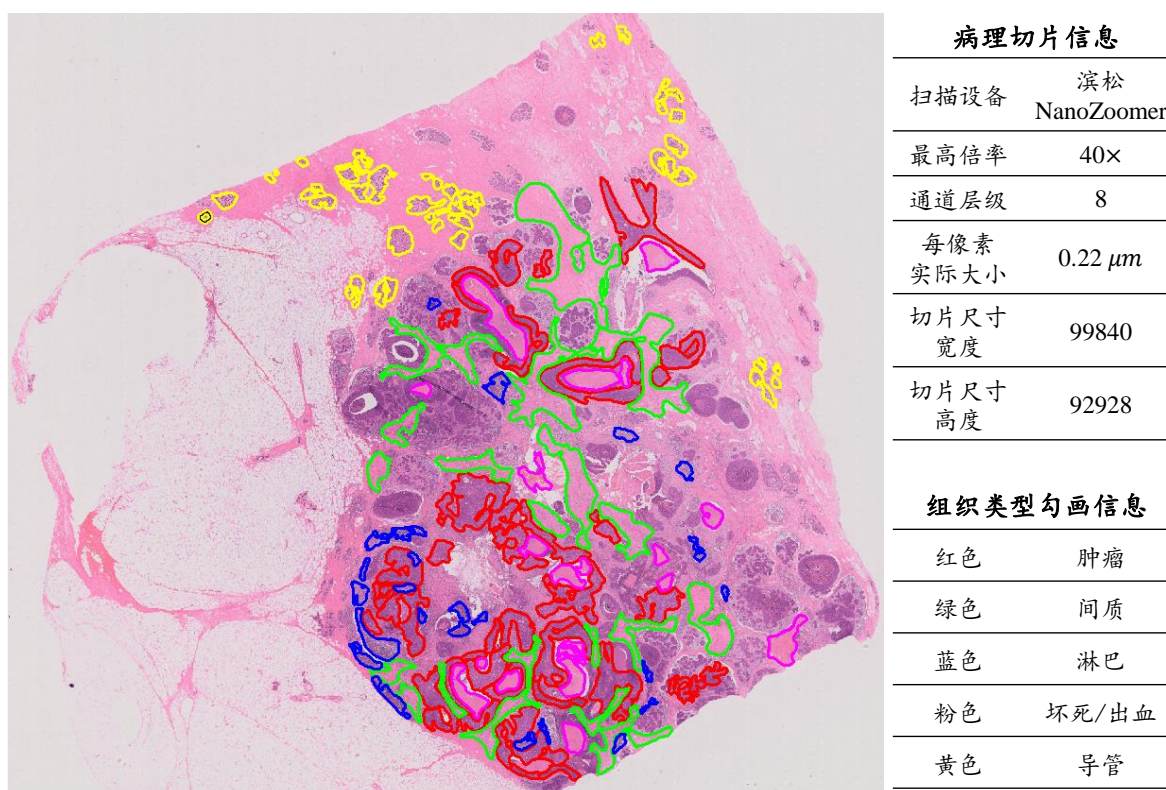


图 1.3 全景病理图像示例图。右侧展示了该切片的属性信息及对应不同组织类型的勾画信息。

因此，本文研究旨在对腺癌患者组织病理图像进行标准、快速、有效的定量化分析，以实现腺癌恶性程度的自动评估分级，最终辅助医生进行临床诊断乃至治疗方案决策。本文的研究重点在于从腺体结构分化这一病理诊断标准出发，基于病理图像分析技术实现腺癌这一癌症类型的恶性程度自动评估分级。

1.2 国内外研究现状

近年来，数字病理领域快速发展^[20]，许多用于研究数字化组织病理切片的计算病理

学算法被提出,以帮助解决与疾病诊断、预后和治疗反应预测相关的问题^[15-18]。Srinidhi 等人^[21]综述了病理图像分析领域先进的深度学习方法,从机器学习策略(包含监督学习、无监督学习、弱监督学习、迁移学习等)、病理图像分析目标(包含良恶分类、分级、有丝分裂检测、细胞核分类检测分割、肿瘤区域分割、病理组织结构分割等)和计算病理领域公开数据(包含乳腺癌、结直肠癌、前列腺癌、肺癌等)多个角度进行综合性概述,堪称计算病理领域的“资治通鉴”。对于腺癌而言,许多病理图像分析算法被提出以实现癌变区域识别和恶性程度自动评估分级^[22-29]。Doyle 等人^[22]提出增强贝叶斯多分辨率系统用于识别病理穿刺活检切片上的癌变区域。Monaco 等人^[23]使用类加权的马尔科夫随机场方法来识别全景病理图像上的癌变感兴趣区域。龚磊等人^[30]和何雪英等人^[31]使用深度学习方法对乳腺癌组织病理图像进行自动分级。Li 等人^[25]提出了基于区域的卷积神经网络框架用于多任务的上皮细胞检测和自动格里森分级。Poojitha 等人^[29]设计了深度学习统一混合结构以进行前列腺癌的自动格里森分级。此外, *The Lancet Oncology* 期刊中的两项研究^[32-33]都采用深度学习方法在穿刺活检病理图像中检测前列腺癌的感兴趣区域并进行自动格里森分级。

表 1.1 计算病理领域^[15]在癌症器官类型、染色类型、扫描数据类型、研究方法、研究区域、研究目标等多个角度的研究方向。

计算病理	癌症器官类型	乳腺、前列腺、结直肠、肺、肾脏、口腔、头颈等。
	染色类型	H&E、IHC、PAS、FIL、ER、PR、Ki-67 等。
	数据类型	术后、穿刺活检、微阵列等。
	研究方法	病理组学特征工程、机器学习、卷积神经网络等。
	研究内容	细胞、核、组织结构、肿瘤、淋巴、间质、坏死等。
	研究目标	染色评估、检测、识别、分割、分类、诊断、预测、预后等。

得益于强大的特征表示能力,深度学习方法在病理图像分析中获得了优越的性能表现,这也使得其在包含腺癌恶性程度评估分级的计算病理领域应用十分广泛^[34-35],但大部分方法的黑箱性质引发了业界关于可解释性和临床应用的讨论^[17]。Bera 等人^[17]综述了计算病理领域的各种图像分析算法,包括深度学习方法和手工特征。其在讨论两类特征方法利弊时指出:计算病理领域的深度学习方法获取的特征目前仍旧很难被解释,临床应用中只适用于诸如细胞结构检测分割等底层任务;而手工特征通常与疾病领域知识相关联,具有更好的可解释性以及生物学基础,因此更可能用于疾病预后、治疗预测等

高层级决策任务。如 Lu 等人^[36]依据肺癌中细胞核的局部空间排列这一领域知识提出局部细胞聚类图算法,用于预测肺癌患者的长短期生存状态和乳头瘤病毒状态两个高层级目标。James 等人^[37]从肿瘤微环境这一领域知识出发,提出了结合细胞和组织结构的手工特征以捕捉肿瘤特异性和生物学相关特征,实现对同源重组缺陷和免疫检查点蛋白表达的预测。而在腺癌领域中,病理医生进行恶性程度自动评估分级时依赖的关键领域知识是腺体的分化程度^[6-8]。Mosquera-Lopez 等人^[38]讨论了多种基于图像纹理特征的腺癌检测和分级方法。Ali 等人^[39]通过提取细胞核形状、结构和纹理特征集合以区分不同等级的癌变区域。Farooq 等人^[24]采用 Gabor 和局部二值模式纹理特征,结合机器学习分类器实现腺癌自动分级。Niazi 等人^[40]提取前列腺癌的腺体结构中的管腔和结构特征来区分低级别和高级别的前列腺癌。Zhou 等人^[41]针对腺体形态和细胞信息提出基于细胞核的图卷积神经网络,实现了结直肠腺癌自动评估分级。

此外,也有一些研究通过对腺体结构自动分割结果进行定量描述以实现腺癌的恶性程度评估分级。Xu 等人^[42]采用了测地线活动轮廓模型对前列腺癌中的腺体结构进行分割然后进行分级。Ali 等人^[39]提出基于形状先验的自适应主动轮廓算法对腺体结构的细胞核分割以进行后续的分级工作。Lee 等人^[43]提出了一种腺体倾斜度计算方法,对腺体结构紊乱程度进行定量化以区分良性和恶性腺体,这些特征也被证明是前列腺切除术后复发风险的预测因素。Awan 等人^[44]首先对腺体结构分割,然后提出最佳匹配指标进行腺体结构统计度量,实现了腺癌的恶性程度评估。Bulten 等人^[33]首先训练腺体分割模型,然后对整张穿刺病理切片进行腺体预测,最后依据整张切片中不同腺体结构的分化程度计算百分比来确定患者的格里森分级结果。

基于以上研究,对于腺癌的恶性程度自动分级评估,本文从腺体结构分化程度这一病理诊断标准出发。一方面,本文设计领域特定的手工特征,提出基于同源性分布统计表示的前列腺癌自动格里森分级方法,通过对腺体结构周围细胞核的拓扑排列进行描述以定量化腺体结构分化程度,并采用机器学习分类算法进行前列腺癌的自动格里森分级。这种特征表示方法可解释性强,具有更好的生物学基础。另一方面,启发于对腺体结构分割结果的精确描述以进行恶性程度评估分级,本文将腺体自动分割和恶性程度分级评估融合到一个端对端的网络模型中,因此提出基于多任务学习和先验知识的结直肠腺癌良恶自动评估方法,该方法同时将腺体结构这一重要的病理先验知识编码为空间注意力融合到良恶分级模型中,从而约束分级模型对腺体结构的推理偏好。这种先验知识约束的思想同样具有更好的解释性。

1.3 论文的主要内容及章节安排

本文共分为五个章节，具体章节内容安排如下：

第一章：绪论部分，该章节首先介绍了本文的研究背景及意义，包括研究背景、研究问题、研究意义、研究难点，然后是研究问题的国内外研究现状，最后简单介绍全文的结构安排。

第二章：相关理论基础部分，介绍了腺癌这一癌症类型的临床分级病理学诊断标准，还包括了对同源性分布算法和多任务学习的概述。

第三章：该章节介绍了基于同源性分布统计表示对前列腺癌进行自动格里森分级的方法，详细阐述了该方法中的病理图像预处理、同源性分布统计表示和分类器建模部分；并从多个层面进行实验以验证该方法的有效性。

第四章：该章节介绍了基于多任务学习和先验知识的结直肠腺癌良恶自动评估分级方法，详细描述了该方法中的特征学习骨干网络、分割分级多任务分支和腺体结构空间注意力约束三个部分；同时定量、定性实验评估了该方法的优越性。

第五章：总结与展望，主要包括对全文研究工作的概括性讨论和总结，以及对未来研究工作的展望。

第二章 相关理论基础

基于腺体结构分化程度这一病理诊断标准进行腺癌恶性程度自动评估分级的工作主要从两个角度开展。一个是设计手工特征进行定量描述，构建基于同源性分布特征表示的分级模型；另一个是提取腺体结构这一先验知识作为空间注意力约束分级模型，构建基于先验知识的分割分类多任务学习卷积网络模型。在详细介绍研究方法之前，先对腺癌这一组织学癌症类型的临床分级病理学诊断标准进行介绍，并对同源性分布和多任务学习相关理论基础进行概述。

2.1 腺癌临床分级病理学诊断标准

腺癌是前列腺癌、结直肠癌、乳腺癌、肺癌等恶性肿瘤中常见的组织学类型，其发生位置大多为上皮涎腺组织。对于大多数腺癌的病理学分级诊断，腺体组织结构的分化程度是病理学家确定癌症恶性程度的决定性因素，也因此被写入肿瘤病理学临床实践指南的诊断标准中^[6-8]。



图 2.1 前列腺癌格里森分级系统中三级四级分化模式示意图。

对于前列腺癌而言，格里森分级系统^[9]是进行癌症诊断和风险评估的最广泛使用的分级方案^[10]。格里森分级系统将组织区域内腺体结构不同的分化模式划分为 1 至 5 级，其中，1 级代表分化良好，5 级代表分化最差。临床中 3 级 4 级前列腺癌最为常见且最难区分，其对应的分化模式如图 2.1 所示，具体描述为：

- 3 级：腺体结构中分化；出现小腺体或腺泡以及呈浸润性生长，或出现小型筛状结构腺体；具有少数可识别的腺体结构，大小和形状各不相同；
- 4 级：腺体结构中低分化；多个腺体结构出现共壁甚至融合的情况，腺体之间难以区分，或出现大型筛状结构腺体；几乎没有可识别的完整腺体结构。

对于结直肠癌腺癌，尽管目前有许多分级系统，但缺乏一个统一使用的分级标准。大

多数系统依据腺体结构的分化程度和形态将肿瘤分为 3 个等级或 4 个等级，从等级 1 到等级 4 分别为分化良好、中等分化、分化差和不分化。其中，分化良好和中等分化的等级 1、2 通常又被归为低等级，分化差和不分化的等级 3、4 被归为高等级。然而，鉴于目前的预后价值，按照简单性和可重复性原则，Washington 等人^[6]建议对低等级和高等级结直肠腺癌使用 2 层分级系统。提出以下仅基于腺体结构分化的分级标准，目前已被广泛接受。

- 低等级：大于或等于 50%的腺体分化程度；
- 高等级：少于 50%的腺体分化程度。

2.2 同源性分布概述

同源性分布算法是拓扑数据分析领域一种用于测量对象几何拓扑形状的方法^[45]。在给定拓扑空间的情况下，同源性分布算法对数据的连接组件、回路、孔洞以及其他高维特征进行代数表示。同源性分布算法是拓扑分析理论中一个相对较新的概念，为研究结构复杂的数据提供了一个易于理解的理论框架，它可以通过线性代数的方式进行特征计算。此外，该方法对于输入数据中的小扰动具有鲁棒性，因此在多个领域下应用广泛，包含机器人、神经科学、脑动脉分析、蛋白质结构、突变谱和生存分析、形状建模、时间序列建模和自然语言处理等。

对于同源性分布算法而言，可供研究的数据集类型包含有限度量空间、数字图像、实值函数水平集和无向网络。因此，该算法中一个最为关键的部分在于如何将研究数据抽象建模为可研究对象类型。考虑到本文研究数据为图像类型，因此接下来以图像类型为示例引入同源性分布算法，如图 2.2 所示。

数字化图像具有立方体结构，由像素（对于二维图像而言）或者体素（对于三维图像而言）给出。因此，一种研究数字图像的同源性分布方法使用“立方复合物”的组合结构^[46]。大致上，立方复合物是由顶点、边、正方形、立方体和高维超立方体的并集构成的拓扑空间。从二维数字图像构建立方复合体的一种有效方法是为每个像素分配一个顶点，然后通过一条边将与相邻像素对应的顶点连接起来，并填充生成的正方形。对于三维图像，以类似的方式进行。然后，用与像素的灰度值相对应的整数标记每个顶点，并用相邻顶点的最大值标记边缘。进而可以构造一个立方复合体的嵌套序列 $S_0 \subseteq S_1 \subseteq \dots \subseteq S_{255}$ 。其中，对于每个 $i \in \{0, 1, \dots, 255\}$ ，立方复合体 S_i 包含了以小于或等于 i 的数字标记的所有顶点、边、正方形和立方体。这样的立方复合物序列也被称为“滤波立方复合

物”。

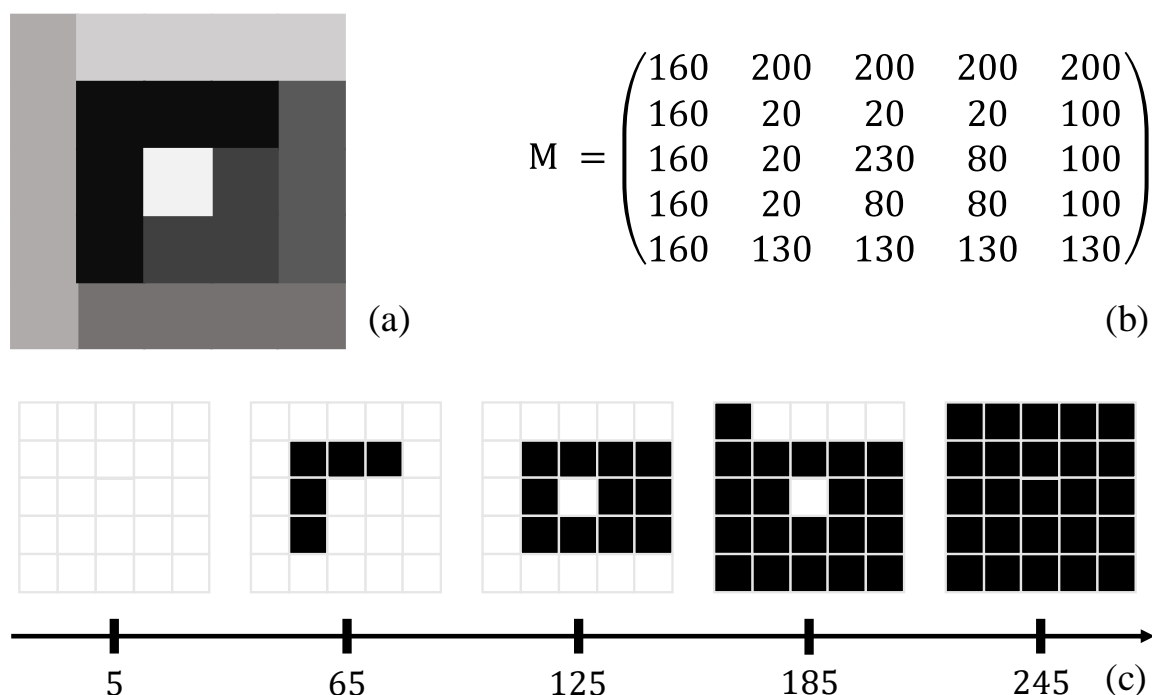


图 2.2 灰度数字化图像的同源性分布示例。(a)为灰度化数字图像示例；(b)为灰度化图像对应的像素灰度值矩阵；(c)为基于不同灰度值 $i \in \{5, 65, \dots, 245\}$ 获取的立方复合物拓扑结构。

代数拓扑可对一些基本对象内容进行计数，将向量空间甚至更复杂的拓扑结构与代数数值相关联来捕获这些对象内容的基本特征。因此，对于以上构建的立方复合物拓扑结构嵌套序列 $\{S_i\}_{i=0}^{255}$ ，使用代数拓扑可以测量各个空间 S_i 的多个特征 $H_j(S_i)$ 。抽象来说，对于每个自然数 $j \in \{0, 1, 2, \dots\}$ 和 $i \in \{0, 1, \dots, 255\}$ ，代数拓扑能将向量空间 $H_j(S_i)$ 与空间 S_i 相关联。 $H_0(S_i)$ 的维度表示空间 S_i 中连通分量的数量， $H_1(S_i)$ 的维度表示空间 S_i 中孔洞的数量， $H_2(S_i)$ 的维度表示更高维对象的数量，这里的 $H_j(S_i)$ 即表示同源性分布算法获得的同源性特征。这些同源性特征 $H_j(S_i)$ 的一个重要特性是鲁棒性，因为当基础空间 S_i 经过弯曲、拉伸或其他变换而变形时，特征 $H_j(S_i)$ 不会改变。从技术上讲，它们是同态不变的^[46]。而在病理图像中，由于切片制片等主观因素导致细胞、腺体组织结构弯曲变形，使得即便是同一等级的腺体结构也呈现高度差异性，因此具有同态不变特性的同源性分布算法将能很好地解决这一问题。

2.3 多任务学习概述

近年来，深度神经网络模型在分类、目标检测、分割等很多任务上都展示出卓越的

性能^[21]。然而，大部分任务都是通过各自训练一个单独的神经网络来实现，这不符合现实中的很多问题。通常来说，许多现实问题都是多维度的。例如自动驾驶问题，为了使汽车能够安全地行驶，它需要捕捉视野内的所有对象并确定类别、距离以及预计路线。这就要求汽车系统能够根据一个视野高效地处理多个任务。这促使了研究者开发多任务学习模型，即根据一个输入能够预测多个目标，实现多个期望输出。

在深度学习领域，多任务学习定义为：给定 m 个学习任务 $\{\mathcal{T}_i\}_{i=1}^m$ ，能够同时从 m 个任务的监督信息中学习共享特征表示的一种网络设计模式^[47]。相较于各个任务各自解决的传统单任务学习，多任务学习卷积网络具有以下几个理论性优势。首先，多任务网络中共享网络层的存在大大减少了模型训练时的内存显存占用。其次，对多任务网络中的每个目标任务而言，共享网络层显式地减少了模型对特征图的重复计算，这提高了模型的推理预测效率。另外，如果多个任务之间具有信息互补或者正则作用，那么通过共同优化的方式能够提高共享特征的鲁棒性和整个模型的泛化性，甚至可以提高各个任务的性能结果。

具体地，多任务学习的设计分为两个层次。首先是网络结构层面的设计，即如何构建网络层以实现多个任务的信息互补和共享特征表示。常见的网络结构可以分为硬参数共享和软参数共享两种，如图 2.3 所示。硬参数共享机制指多个任务使用参数完全一致的共享骨干网络进行特征提取，然后将该特征用于特定于任务的多分支；软参数共享机制指的是多个任务的骨干网络被加以约束限制（如距离正则）以保证参数相似化，获取的特征同样用于特定于任务的多分支。Vandenhende 等人^[48]从任务交互发生的位置（即任务之间交换或共享信息的网络层位置）这一角度确定网络结构的设计：基于编码器的多任务学习网络结构和基于解码器的多任务学习网络结构，如图 2.4 所示。前一种情况指的是多个任务仅在编码阶段共享特征，此后特定于任务的多分支预测多个任务的各自输出。后一种情况更为复杂，指多个任务在解码阶段（即在特定于任务的多分支中）共享信息，以优化每个任务的初始预测。

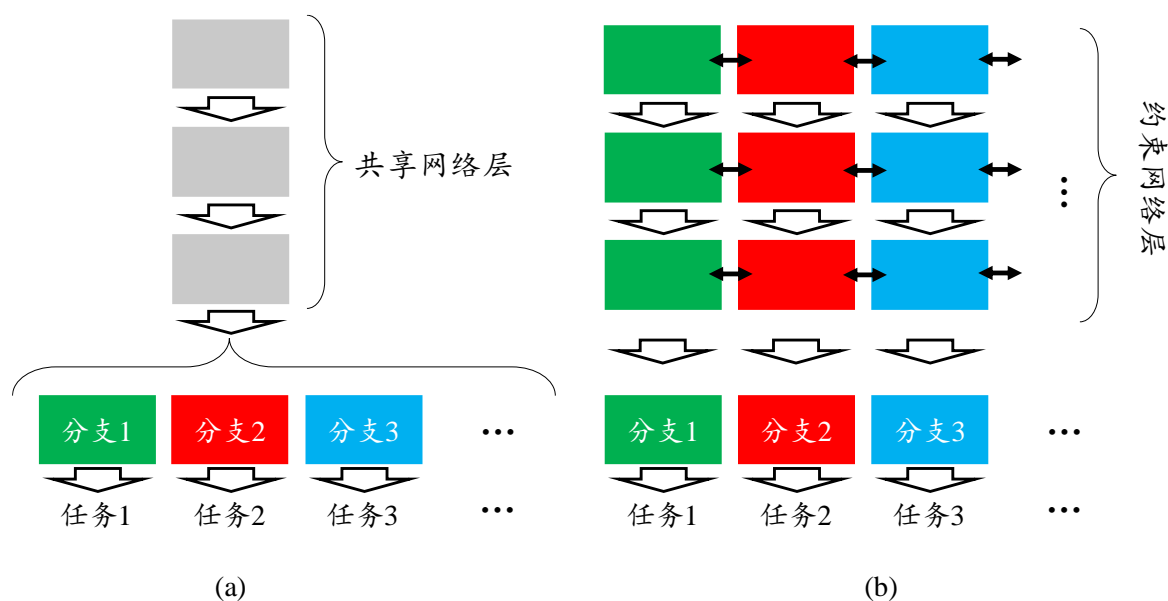


图 2.3 多任务学习中常见的网络结构层面的设计机制。(a)为硬参数共享机制的网络结构设计模式；(b)为软参数共享机制的网络结构设计模式。

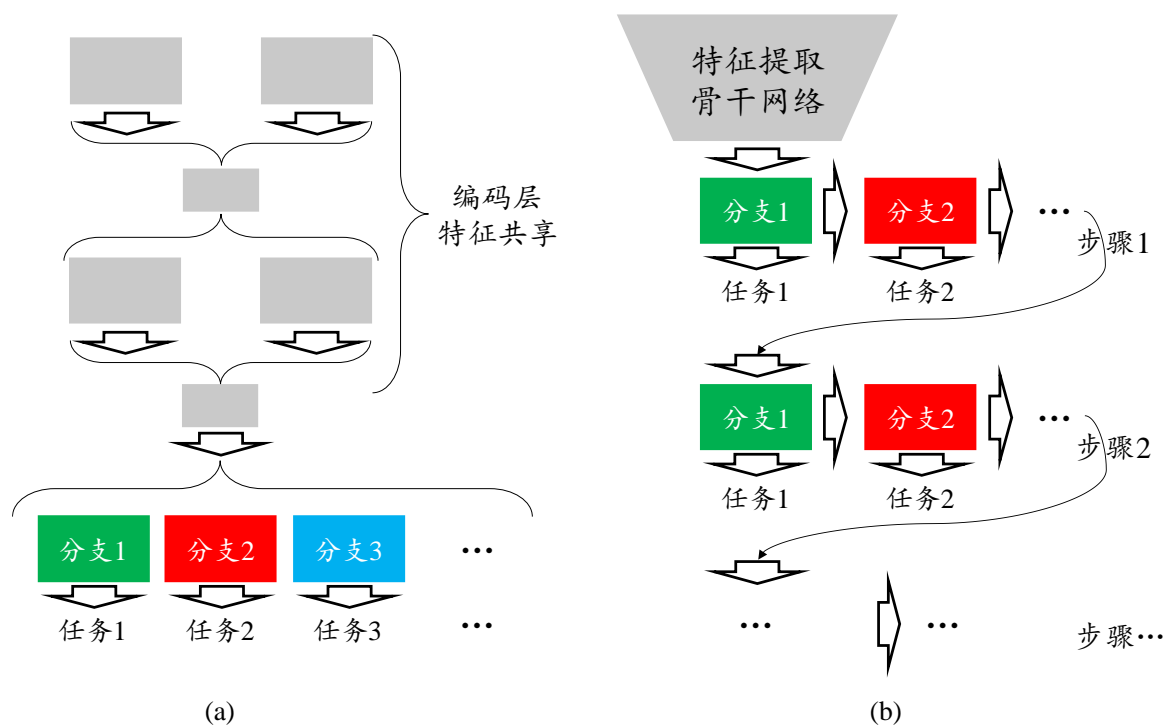


图 2.4 基于任务交互发生位置定义的多任务学习网络结构设计机制。(a)为基于编码器的多任务学习网络结构设计模式；(b)为基于解码器的多任务学习网络结构设计模式。

多任务学习另一层次的设计为损失优化策略，即如何设计损失函数以平衡所有任务间的联合学习，以避免某一个或某几个任务在联合学习过程中占主导地位。通常情况下多任务学习损失函数 \mathcal{L}_{MT} 可以归纳为：

$$\mathcal{L}_{MT} = \sum_i^m w_i \mathcal{L}_{\mathcal{T}_i} \quad (2.1)$$

其中, $\mathcal{L}_{\mathcal{T}_i}$ 表示为第*i*个任务 \mathcal{T}_i 的损失函数, w_i 表示该任务对应的调节权重。一般地, 当使用随机梯度下降优化算法最小化多任务损失 \mathcal{L}_{MT} 时, 共享网络层中的权重参数 W_{share} 按照如下规则更新:

$$W_{share} = W_{share} - \gamma \sum_i w_i \frac{\partial \mathcal{L}_{\mathcal{T}_i}}{\partial W_{share}} \quad (2.2)$$

其中, γ 表示学习率。可见, 一旦某个任务 \mathcal{T}_j 的梯度值远大于其他任务, 权重 W_{share} 的更新很容易被其主导。因此 w_i 被设置用于平衡所有任务 $\{\mathcal{T}_i\}_{i=1}^m$ 间的梯度幅度。目前最常见的做法是针对各个任务的 $\mathcal{L}_{\mathcal{T}_i}$ 在 \mathcal{L}_{MT} 中的幅度比例确定调节权重 w_i 的值。

第三章 基于同源性分布统计表示的前列腺癌自动格里森分级方法

3.1 研究动机

对于前列腺癌而言，格里森分级系统^[9]是评估恶性程度的重要依据，也是进行癌症诊断和风险评估的最广泛使用的分级方案^[10]。格里森分级系统根据组织区域内腺体结构的不同分化模式确定 1 至 5 级，其中，1 级代表分化良好，5 级代表分化最差。前列腺癌临床中 3 级 4 级最为常见且最难区分。图 3.1 展示了一张包含病理医生手工等级标注的数字病理切片示例。

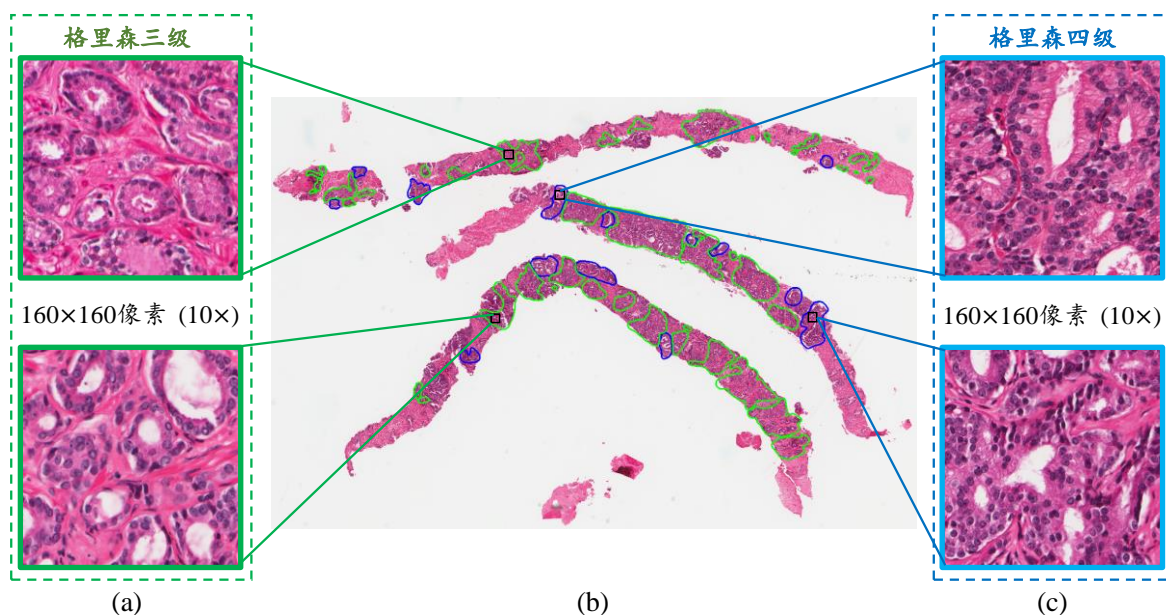


图 3.1 包含病理医生手工等级标注的数字全景病理图像切片示例。其中，(a)、(c)分别展示了 160×160 像素大小的格里森三级、四级病理图像块。

随着数字病理领域的快速发展^[20]，许多计算病理学算法被提出用于研究数字化组织病理切片，帮助解决疾病诊断等相关问题^[22-33]。但目前应用广泛的深度学习方法的黑箱性质引发了业界关于可解释性和临床应用的讨论。Bera 等人^[17]综述了计算病理领域包括深度学习方法和手工特征等病理图像分析算法，其讨论指出：计算病理领域的深度学习方法获取的特征目前仍旧很难被解释，在临床应用中只适用于诸如细胞结构检测分割等底层任务；而手工特征通常与疾病领域的知识相关联，具有更好的可解释性以及生物学基础。

因此，本文从腺体结构分化程度这一病理诊断标准出发，设计领域特定的手工特征，提出同源性分布统计表示特征对腺体结构分化程度进行量化，以实现前列腺癌的自动格里森分级。

3.2 研究方法

本文所提出的基于同源性分布统计表示的前列腺癌自动格里森分级方法，旨在通过对病理图像中腺体周围细胞核的拓扑排列进行描述，以定量化腺体结构分化程度，进而实现对前列腺癌的自动格里森分级。该方法包含了三个主要部分：a).病理图像预处理；b).病理图像的同源性分布统计表示特征提取；c).分类器建模。图 3.2 展示了该方法的整体流程图。

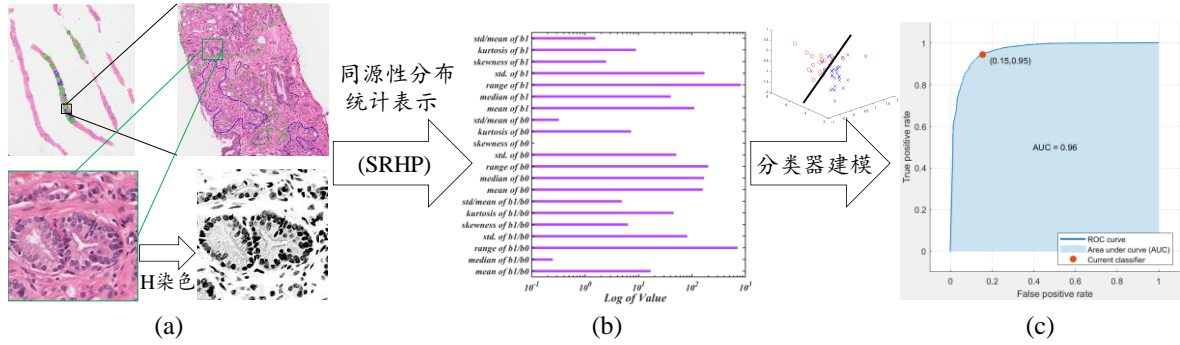


图 3.2 本文提出的基于同源性分布特征表示的前列腺癌自动分级方法流程图。(a)为病理图像预处理部分；(b)为经过同源性分布统计表示提取到的特征；(c)为经过分类器建模后获取到的模型。

3.2.1 病理图像预处理

本文提出的同源性分布统计表示的意义在于基于细胞核之间的拓扑排列来定量描述腺体结构的分化程度。因此，获取清晰明朗的细胞核群以捕获细胞核之间的拓扑关系起基础作用。对于苏木精-伊红（Hematoxylin & Eosin, H&E）染色的病理石蜡切片，细胞核内的染色质被苏木精染色剂作用呈现紫蓝色，细胞质成分和细胞外基质被伊红染色剂染为红色。根据这一特性，本文利用染色反卷积算法^[49]对 H&E 图像进行染色分离，以获取具有清晰细胞核结构的 H 染色通道。如图 3.3 (c)所示。然而，由于切片、制片、染色、数字化扫描等主观客观因素的影响，不同的 H&E 病理图像染色呈现高度差异性，而染色反卷积算法中的 Ruifrok 固定染色矩阵 M_r 无法代表所有 H&E 图像的染色矩阵，这导致直接利用该固定染色矩阵参数进行染色分离并不合适。因此，本文采用 Macenko 等人^[50]提出的染色估计方法为每张病理图像计算染色矩阵 M_e 以观察不同染色之间的差异性，同时对图像应用本队列平均染色矩阵参数 \bar{M}_e 进行染色分离以消除染色差异性、提高后续模型的有效性。两种方式对应的染色矩阵如下所示：

$$M_r = \begin{bmatrix} 0.6442 & 0.7166 & 0.2668 \\ 0.0928 & 0.9541 & 0.2831 \end{bmatrix} \quad (3.1)$$

$$\bar{M}_e = \begin{bmatrix} 0.5155 \pm 0.0307 & 0.7234 \pm 0.0169 & 0.4576 \pm 0.0205 \\ 0.1501 \pm 0.0200 & 0.7723 \pm 0.0308 & 0.6149 \pm 0.0391 \end{bmatrix} \quad (3.2)$$

其中, M_r 染色矩阵为染色反卷积算法中的 Ruifrok 固定矩阵参数, \bar{M}_e 为采用染色估计方法^[50]计算得到的本队列所有病理图像块的平均染色矩阵; 其中 M_r 和 \bar{M}_e 染色矩阵的第一行和第二行分别代表了苏木精染色通道和伊红染色通道。图 3.3 展示了本队列所有病理图像块采用染色估计方法得到的苏木精通道染色矩阵的三维散点图以及 H&E 图像应用 M_r 和 \bar{M}_e 不同染色矩阵得到的 H 染色通道图像。

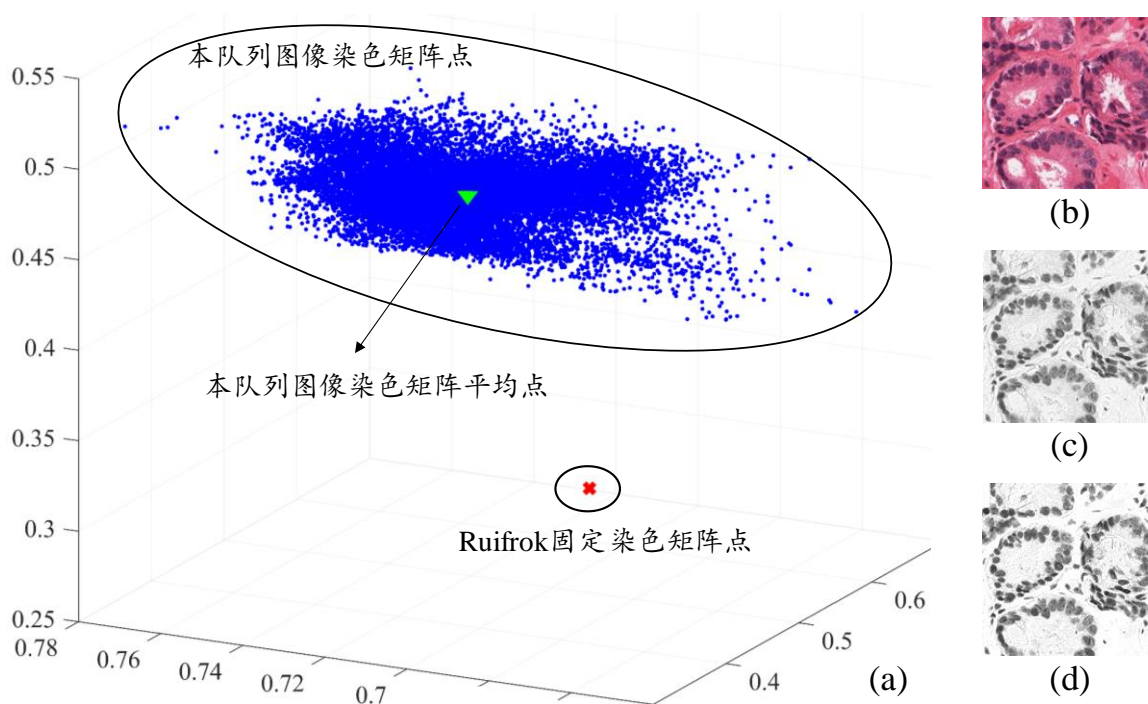


图 3.3 病理图像 H 染色通道提取示意图。(a)为染色矩阵苏木精通道的三维散点图; 红色叉号、蓝色点和绿色三角形分别表示了 Ruifrok 固定染色矩阵 M_r 、采用染色估计方法得到的本队列所有病理图像块的染色矩阵 $\{M_e\}$ 以及队列平均染色矩阵 \bar{M}_e ; (b)为本队列某一 H&E 病理图像块; (c)为应用 Ruifrok 固定染色矩阵 M_r 得到的 H 染色通道图像; (d)为应用队列平均染色矩阵 \bar{M}_e 得到的 H 染色通道图像。

3.2.2 同源性分布统计表示

同源性分布算法是用于量化拓扑空间中对象的几何拓扑特征的一种拓扑数据分析工具。对于二维图像, 同源性分布算法通过计算多组阈值化后的二值图像的连通分量和孔洞的数量 (分别表示为 b_0 、 b_1) 来进行区域描述, 以获取该图像中结构拓扑特征的变化规律。本文利用同源性分布算法计算得到的 b_0 、 b_1 集合来描述腺体周围细胞核之间的拓

扑排列，从而量化腺体结构的分化程度。

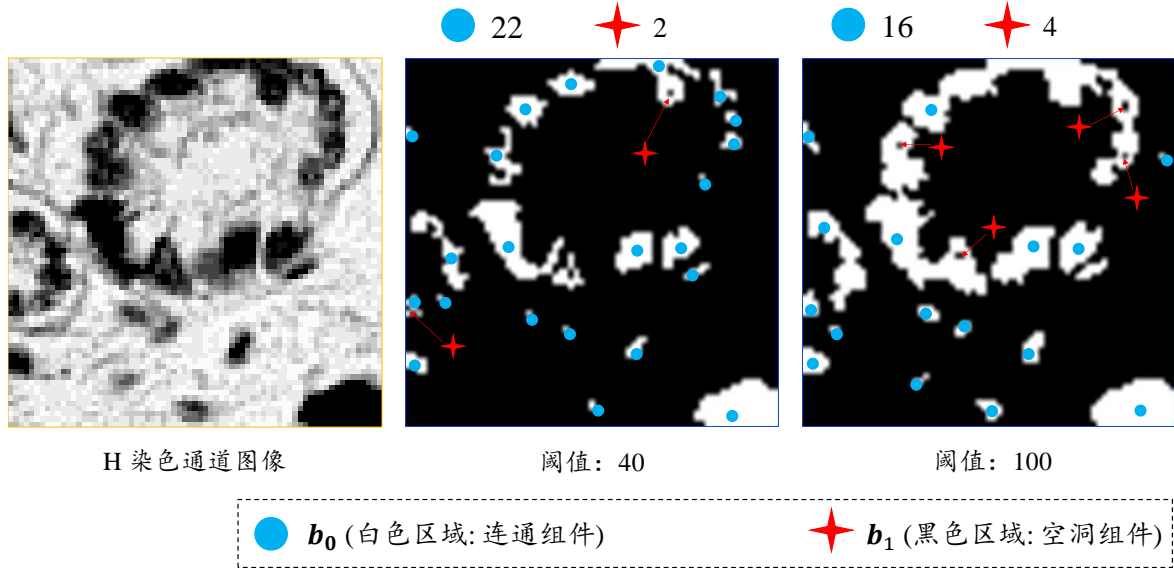


图 3.4 H 染色通道图像在不同阈值 t (40 和 100) 下二值化为 \widetilde{X}_t 后对 b_0 、 b_1 分量的计算。蓝色点和红色星号分别表示 b_0 、 b_1 分量。

依据同源性分布算法，本文首先应用阈值 t 将 H 染色通道图像 X 二值化为 X_t ，为保证细胞核及腺体结构为感兴趣对象，对 X_t 取反以获得新的二值图像，表示为 \widetilde{X}_t ，即，将图像 X 中小于等于 t 的像素所在位置确定为属于感兴趣对象（包含细胞核的腺体结构），用像素值 1 表示，其他位置用像素值 0 表示。得到的二值图像 \widetilde{X}_t 可以被认为是 H 染色通道图像 X 关于像素值 t 的滤波表示。然后，分别基于 \widetilde{X}_t 和 $\widetilde{X_{255}} \setminus \widetilde{X}_t$ 计算特定对象的连通分量得到 b_0^t 、 b_1^t ，其中 $\widetilde{X_{255}}$ 是以阈值 255 确定的取反二值图像， $\widetilde{X_{255}} \setminus \widetilde{X}_t$ 是 \widetilde{X}_t 在 $\widetilde{X_{255}}$ 中的补。此时 $\widetilde{X_{255}} \setminus \widetilde{X}_t$ 中特定对象的连通分量即为 \widetilde{X}_t 的空洞组件。当阈值 t 从 0 到 k 逐渐变化时，H 染色通道图像 X 关于像素值 t 的滤波表示 \widetilde{X}_t 满足如下特性：

$$\widetilde{X}_1 \subseteq \widetilde{X}_2 \subseteq \dots \subseteq \widetilde{X}_k \subseteq \dots \subseteq \widetilde{X_{254}} \subseteq \widetilde{X_{255}} = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \quad (3.3)$$

因此，阈值 t 从 0 到 k 逐渐变化时，得到的各个二值图像 \widetilde{X}_t 中腺体结构呈现动态、有序的变化过程，这使得从二值化图像 \widetilde{X}_t 下获取的 b_0^t 、 b_1^t 值同样呈现增大或减少的动态状态。图 3.4 展示了不同阈值 t 下计算感兴趣对象的连通分量得到 b_0 、 b_1 的过程。图 3.5 展示了 H 染色通道图像 X 经连续阈值化后腺体结构沿阈值 t 方向的有序演化过程及其对应该得到的 $\{b_0^t, b_1^t\}_{t=1}^{255}$ 集合的动态变化曲线。这里，本文将阈值 t 范围设置为 1~255，增量为 1。通过应用该同源性分布算法，每一病理图像块被 b_0 、 b_1 的同源性序列特征 B 所表示，

其中：

$$B = \{b_0^t, b_1^t\}_{t=1}^{255}, \quad B \in \mathbb{N}^{2 \times 255} \quad (3.4)$$

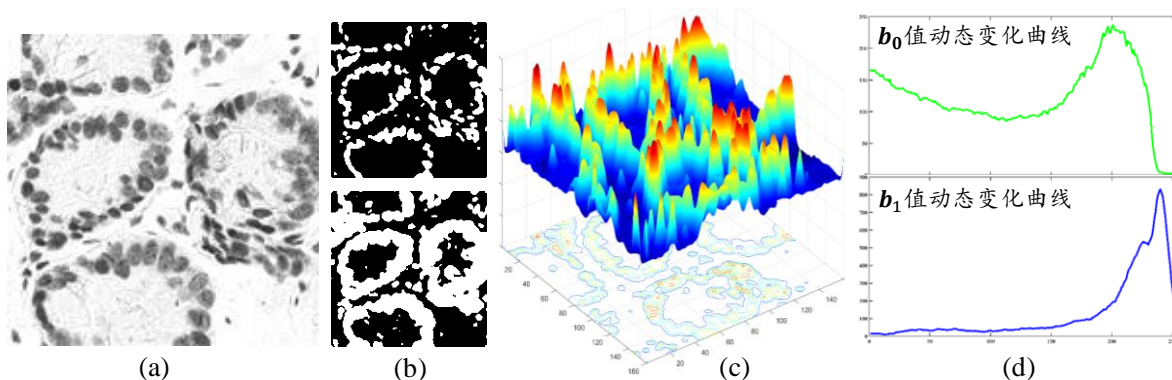


图 3.5 病理图像的同源性分布量计算示意图。(a)为 H 染色通道图像 X ；(b)为对图像 X 应用不同阈值 t 获得的二值化图像 \widetilde{X}_t 的示例展示；(c)展示了 H 染色通道图像 X 经连续阈值化后二维腺体结构沿阈值 t 方向（ z 轴）的有序演化过程，其中底部平面为所有阈值 t 下二维腺体结构累加形成的等高线图；(d)为该 H 染色通道图像 X 随阈值 t 变化的 b_0 、 b_1 变化曲线。

对于从每一病理图像块 X 获得的同源性序列特征 B ，一种做法是使用对称 Kullback-Leibler 散度度量标杆图像块和待测试图像块的序列特征之间的相似性^[51]。然而，这种方法往往严重依赖于标杆图像的选择。而对于前列腺癌而言，即便是相同的格里森等级区域也会存在显著的形态异质性^[52]，因此选取具有代表性的格里森三级和四级的标杆样本非常有挑战性。本文因此不局限于局部标杆样本图像而从队列数据的全局出发，利用统计方法^[53]对同源性序列特征 B 进行二次表征，以描述本队列前列腺癌数据不同类别下的规律性。

通过同源性分布算法计算得到的 b_0^t 、 b_1^t 值随阈值 t 变化而变化，对应的特征 B 因此具有序列性，可以看作是离散分布 Y_0 、 Y_1 。考虑到同源性分量 b_0^t 、 b_1^t 在二值化图像 \widetilde{X}_t 的依存关系，两者之间的比例 b_1^t/b_0^t 被纳入，得到的连续变化量被视为离散分布 Y_r 。此时离散分布 Y_0 、 Y_1 、 Y_r 可以从多个维度进行统计描述：（1）分布的中心趋势，（2）分布的分散程度，（3）分布的形状。这里，均值和中位数统计量被选取用来描述分布的中心趋势，标准差、范围和变异系数统计量用来表示分布的分散程度，分布的形状则用偏度和峰度统计量进行描述。分布 Y_0 、 Y_1 、 Y_r 被以上 7 个统计量统计描述，即，每一病理图像块 X 的同源性序列特征 B 作为三个离散分布 Y_0 、 Y_1 、 Y_r 被 7 个统计量进行描述，图像因此被二次表征获得 21 维同源性分布统计表示特征 $V \in \mathbb{R}^{1 \times 21}$ 。表 3.1 展示了同源性分布统计表示

特征计算的算法流程。

表 3.1 本文提出的同源性分布统计表示特征算法流程。

算法 1: 同源性分布统计特征计算

输入: H&E 病理图像数据集 $\{X^i\}_{i=1}^n = \{X^1, X^2, \dots, X^n\}$

输出: 对应的特征集 $\{V^i\}_{i=1}^n = \{V^1, V^2, \dots, V^n\}$

- 1: 平均染色矩阵 $\bar{M}_e = \text{染色估计方法}(\{X^i\}_{i=1}^n)$
- 2: **for** $i = 1; i \leq n$ **do**
- 3: H 染色通道 $X^{H^i} = \text{染色分离方法}(X^i; \bar{M}_e)$
- 4: **for** $t = 1; t \leq 255$ **do**
- 5: 滤波表示 $\widetilde{X}_t^i = \text{二值化并取反}(X^{H^i}; t)$
- 6: 同源性值 $b_0^t, b_1^t = \text{对象连通分量计算}(\widetilde{X}_t^i)$
- 7: 同源性序列特征 $B^i = \{b_0^t, b_1^t, b_1^t/b_0^t\}_{t=1}^{255}$
- 8: 离散分布 $Y_0^i, Y_1^i, Y_r^i \equiv B^i$
- 9: 同源性分布统计表示 $V^i = \text{统计量描述}(Y_0^i, Y_1^i, Y_r^i)$
- 10: **return** 同源性分布统计表示特征集 $\{V^i\}_{i=1}^n = \{V^1, V^2, \dots, V^n\}$

3.2.3 分类器建模

为了实现前列腺癌的自动分级，基于距离度量的加权 K 近邻分类器（Weighted K Nearest Neighbor, wKNN）算法用于对病理图像块的同源性分布统计表示特征进行建模。wKNN 算法依据欧几里德距离确定 K 个最近邻居并为其分配权重向量 $W \in \mathbb{R}^{1 \times K}$ ：

$$W_{ij} = \frac{1}{\sum_{n=1}^N (x_i^n - x_j^n)^2}, \quad j \in \{1, \dots, K\} \quad (3.5)$$

其中 W_{ij} 被定义为测试样本 x_i 关于第 j 最近邻居样本 x_j 的权重值， x_i^n 表示样本 x_i 的第 n 维特征， x_j^n 表示样本 x_i 在训练样本集中选取的第 j 最近邻居样本 x_j 的第 n 维特征值。N 表示特征总维数，本文为 21。考虑到 wKNN 算法为基于距离度量的分类器算法，其对特征维度之间的尺度差异十分敏感，因此在应用 wKNN 分类算法之前，本文依据训练样本集对本队列病理图像块的同源性分布统计表示特征 V 进行归一化，以消除不同维度的特征尺度的影响，提高分类器的区分能力。

3.3 实验设计

本文提出了一种基于同源性分布统计表示 (Statistical Representation of Homology Profile, SRHP) 的前列腺癌自动分级方法。为表明该方法在前列腺癌病理图像上自动分级的可行性和有效性, 将该方法与另外三种常用的分级方法进行对比分析。

3.3.1 实验数据

本章实验使用的数据是来自宾夕法尼亚大学附属医院的前列腺癌穿刺活检数字化切片队列。该队列共包含 43 张由 Aperio 光学扫描仪获取的 H&E 染色全景病理图像和对应标记文件。其中, 图像的最高放大倍数为 40 倍, 切片实际大小 (微米) 与图像像素比为 0.252, 队列平均像素大小约在 62400×47000 ; 标记文件由宾夕法尼亚大学附属医院的专业病理学家根据临床格里森分级标准进行手工勾画, 得到格里森三级和格里森四级感兴趣区域, 本文将它们作为自动分级方法的标签。

考虑到本文方法是基于细胞核之间的拓扑排列来定量描述腺体结构的分化程度, 而不是针对细胞核个体定量化, 因此本文根据病理学家标注的感兴趣区域直接在 10 倍放大倍数下进行滑动窗采样提取 160×160 像素大小的图像块, 这些图像块的标签等级根据病理学家对该图像块所在区域的勾画类别来确定。其中滑动窗采样的步长为 80 像素。最终, 本文从 43 张全景病理图像队列中采样获得 9597 张格里森三级图像块和 8679 张格里森四级图像块。

3.3.2 对比方法设计

为了表明本文提出方法的有效性和优越性, 将该方法与其他现有的应用于病理图像分类分级的三种主流方法进行对比。这三种方法简要说明如下:

- 传统的病理组学手工特征方法^[39]——表示为 MATF (Morphological, Architectural and Textural Features)。提取细胞核形态、结构和纹理等大量病理组学特征以实现前列腺癌格里森三级四级的区分。
- 端到端的无监督学习方法^[54]——Xu 等人提出的堆叠式稀疏自动编码器算法 (Stacked Sparse Auto Encoder, SSAE), 通过滑动窗策略将每个图像块分类为细胞核或非细胞核, 实现了病理图像的细胞核检测。
- 端到端的监督学习方法^[26]——将深度卷积神经网络方法直接应用到前列腺癌组织微阵列病理图像进行多类格里森分级 (良性, 三级, 四级和五级)。本文将这种方法

表示为 DLGg (Deep Learning for Gleason Grading)。由于本文实验数据相对较少，因此在进行该对比方法实验时，采用该深度卷积神经网络方法的预训练模型在本文实验数据上进行微调。

此外，为了定性化评估本文的方法，本文在测试阶段实现了全景病理图像的预测。完整的全景病理图像推理预测实验步骤如下：首先，对一张全景病理图像进行无重叠滑动窗采样，同时基于大津阈值算法进行背景区域过滤；然后，本文基于卷积神经网络方法构建了良恶性分类模型，对非背景图像块进行良恶性预测，预测为恶性的图像块继而送入到 SRHP 模型中用于区分格里森三级或四级；最终，本文将每一图像块的预测结果（背景、良性、格里森三级、格里森四级）按原图像块位置拼接在一起，形成完整的全景图像预测。

3.3.3 性能评估

为了公平准确地比较不同方法的性能，本文采用准确率 (Accuracy, ACC)、召回率 (Recall, REC)、精度 (Precision, PRC)、特异性 (Specificity, SPC)、F1 得分 (F1 Score)、受试者操作特征曲线下面积 (Area Under the Curve, AUC) 等不同指标进行量化评估。本章节所有实验采用 10 折交叉验证策略，最终实验结果为各指标在多折模型上的平均值。

3.4 实验结果及讨论

3.4.1 图像块层面的评估分析

为了展现本文提出的同源性分布统计表示方法区分前列腺癌格里森等级的有效性，本实验从图像块层面对比了 SRHP 和 MATF、SSAE 及 DLGg 方法。表 3.2 展示了图像块层面应用几种方法在 AUC、准确率、召回率、精度、特异性和 F1 得分指标上的定量结果，其中，实验结果为 10 折交叉验证模型上的平均值。

从表 3.2 中可以看出，在图像块层面的定量评估中，本文提出的 SRHP 方法要优于其他三种对比方法，达到了最高的 AUC 值、准确率、召回率和 F1 得分，分别为 0.96、89.02%、0.94 和 0.89。SSAE 表现最差，仅有 0.79 的 AUC 值和 72.07% 的准确率。而 MATF 和 DLGg 方法的性能处于两者中间。可见，本文提出的 SRHP 方法具有更好的优越性。

与传统的针对单一拓扑空间下个体细胞核的病理组学手工特征方法 MATF 相比，SRHP 方法旨在关注连续演化过程中细胞核之间的拓扑排列变化来序列化、量化描述

腺体结构的分化程度，因而在区分格里森三级和四级癌变区域上更有针对性。另外，实验结果表明，SRHP 方法 AUC、召回率和 F1 得分指标上远好于端到端的 DLGg 和 SSAE 方法，这说明基于数据驱动的 DLGg 和 SSAE 方法无法很好地区分队列图像样本的等级；尤其对于 SSAE 方法，基于自编码器对输入图像进行优化学习，只是提取到图像的通用特征表示而无法聚焦在腺体区域针对腺体结构进行建模。因此，这再次表明了本文根据腺体分化程度这一病理诊断标准提出的 SRHP 方法在区分前列腺癌格里森三级和四级的针对性和有效性。对于精度 PRC 这一指标，本文 SRHP 方法只达到 0.84，相较于最高精度值的 DLGg 方法稍有差距，这说明本文方法预测有更多的假阳性样本，尽管如此，本文方法在 AUC 这一重要指标上达到了最好的结果。

表 3.2 本文提出的 SRHP 方法与其他对比方法在病理图像块层面的定量评估结果。实验结果为 10 次交叉验证模型在 AUC、准确率、召回率、精度、特异性和 F1 得分指标上的平均值。（最佳结果以粗体显示）

图像块层面	AUC	ACC	REC	PRC	SPC	F1 Score
DLGg	0.91	85.04%	0.72	0.90	0.94	0.80
SSAE	0.79	72.07%	0.65	0.73	0.79	0.69
MATF	0.94	86.41%	0.84	0.87	0.89	0.85
SRHP	0.96	89.02%	0.94	0.84	0.84	0.89

3.4.2 切片层面的评估分析

为了展现本文提出的 SRHP 方法区分前列腺癌格里森等级的稳定性，本实验从病理穿刺切片层面对比了 SRHP 方法和 MATF、SSAE 及 DLGg 方法。队列中同时包含格里森三级和四级感兴趣区域的 20 张切片被选取进行逐切片地定量评估，表 3.3 中的定量结果为 20 张切片上各自 AUC、准确率、召回率、精度、特异性和 F1 得分指标值的平均结果。

从表 3.3 可以看出，在切片层面的定量评估中，不同方法整体上保持了和图像块层面大致一致的实验结果：SRHP 方法最好、MATF 次之、DLGg 再次之、SSAE 最差。相较于 MATF、SSAE 及 DLGg 方法，SRHP 方法达到了 0.99 平均 AUC、98.75% 平均准确率和 0.97 平均召回率的最好性能，这再次验证了 SRHP 方法在单一切片上进行格里森自动分级的有效性和可靠性。表 3.3 同样报告了所有评估指标在不同切片间的标准差，SRHP 和 MATF 方法在 AUC、准确率、召回率、F1 得分指标上都有着较低的标准差，

这表明了 SRHP 和 MATF 方法在多切片验证时具有很好的稳定性,这也印证了利用统计方法对同源性序列特征进行二次表征的优越性。

表 3.3 本文提出的 SRHP 方法与其他对比方法在病理切片层面的定量评估结果。实验结果为 20 张穿刺切片上各自 AUC、准确率、召回率、精度、特异性和 F1 得分指标值的平均值和标准差(最佳结果以粗体显示)。其中,20 张穿刺切片同时包含有格里森三级和四级感兴趣区域。

切片层面	AUC	ACC	REC	PRC	SPC	F1 Score
DLGg	0.93±0.13	89.35%±0.14	0.80±0.25	0.82±0.20	0.91±0.13	0.78±0.21
SSAE	0.62±0.18	71.31%±0.12	0.42±0.30	0.48±0.36	0.73±0.21	0.53±0.25
MATF	0.94±0.01	86.45%±0.02	0.84±0.03	0.87±0.05	0.89±0.03	0.85±0.03
SRHP	0.99±0.01	98.75%±0.01	0.97±0.03	0.93±0.12	0.98±0.02	0.95±0.09

进一步地,图 3.6 展示了以每张病理切片为单位的不同对比方法的评估结果散点图。其中,每一个点都表示一张病理切片,横轴、纵轴分别代表了 AUC 和准确率指标,不同的颜色表示不同的方法。从图中可以发现,表示 SRHP 方法的红色点几乎都聚集在散点图的右上角,而表示 DLGg 方法的蓝色点、表示 MATF 方法的绿色点和表示 SSAE 方法的黑色点都分散分布在红色点的下方、左下方、图左侧甚至部分超出坐标范围域,这说明本文提出的 SRHP 方法相比于其他方法而言达到了更好的性能。此外,表示 SRHP 方法的红色点和表示 MATF 方法的绿色点都各自呈现聚集的状态,这直观表明了 SRHP 特征方法和 MATF 方法在不同切片间的鲁棒性。观察散点图发现本文 SRHP 方法有一例样本的性能游离在整体之外,但该样本的性能仍旧好于其他方法下的大部分样本。

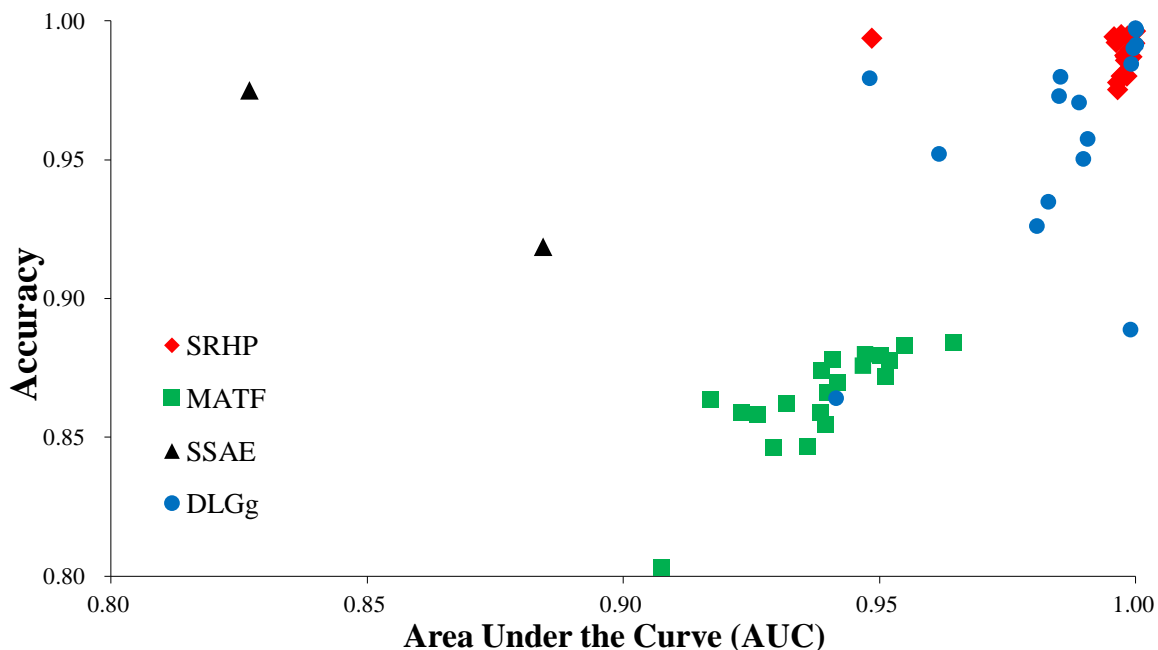


图 3.6 不同对比方法下每张病理切片评估性能的散点图。散点图分别以 AUC 和准确率指标作为横轴、纵轴；其中每个红色、绿色、黑色、蓝色点分别代表了 SRHP、MATF、SSAE 和 DLGg 方法下每张切片的评估结果。注 a: 部分超出坐标范围域的样本点未被绘出

3.4.3 统计特征分析

本实验的目的在于分析本文方法涉及的所有 21 维同源性分布统计表示特征区分格里森三级和四级的有效性。图 3.7 展示了用于格里森自动分级的所有 21 维同源性分布统计表示特征的箱形图。

图 3.7 的箱形图表明了本文提出的 SRHP 特征在进行格里森三级、四级分级上具有可区分性。观察发现同源性分量 b_1 的第 15~21 维统计量特征（第 20 维除外）的上下四分位点距离宽，说明同一类样本在孔洞数量这一特性上具有较大类内差异性。即便如此，比较这些统计量特征的格里森三级、四级的箱形图看出，其仍能很好地将不同类别区分开来，这得益于同源性分量 b_1 统计表示特征对不同等级腺体结构的差异化定量描述。另外观察发现，相较于 b_1 的第 15~21 维特征，同源性分量 b_0 的第 8~14 维特征的上下四分位点距离更窄，这表明了同源性分量 b_0 、 b_1 之间的互补性。此外，从所有特征的箱形图发现格里森三级的特征值普遍高于格里森四级，说明同源性分量 b_0 、 b_1 对应的细胞核连通分量和腺体结构孔洞数量三级普遍高于四级，这符合病理诊断标准中不同等级的状态：三级分化良好，腺体结构形成各自独立完整腔体；四级分化差，产生细胞核粘连，腺体结构不具有清晰完整的腔体。

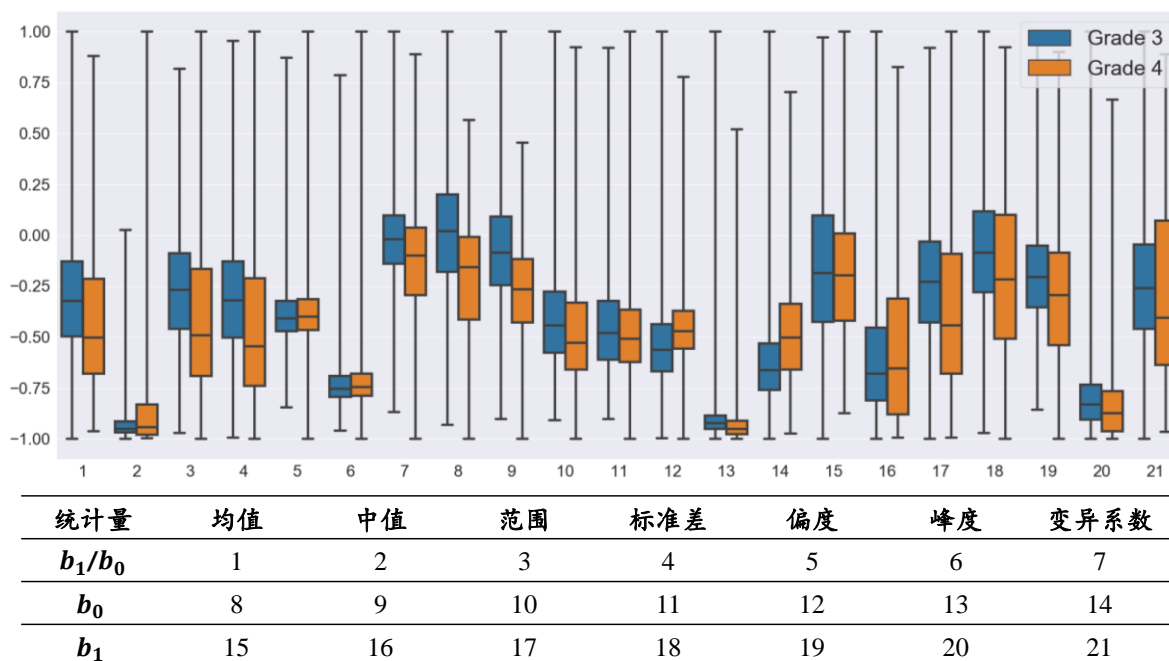


图 3.7 用于格里森自动分级的 21 维同源性分布统计表示特征的箱形图。深蓝色、橙色箱体分别代表格里森三级、四级。图下方的表格列出了每一维特征对应于某一同源性分量的某一统计量。

本实验中同样分析了三种独立的同源性分量 b_0 、 b_1 、 b_1/b_0 的统计表示特征各自进行自动分级的差别。图 3.8 为 b_0 、 b_1 、 b_1/b_0 以及 SRHP 的对比评估结果。

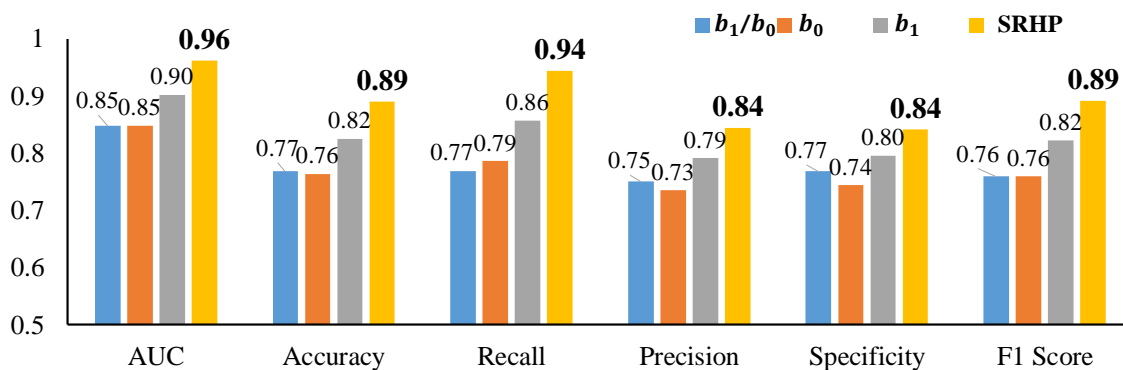


图 3.8 同源性分量不同组合的对比评估结果。柱状图表示 SRHP 方法和三种独立的同源性分量 b_0 、 b_1 、 b_1/b_0 的统计表示特征各自进行自动分级的性能。

从图 3.8 可以看出，相比于独立的同源性分量统计表示， b_0 、 b_1 、 b_1/b_0 共同组合的 SRHP 方式达到了最佳的评估结果，这表明了本文提出的 SRHP 特征集的有效性。除此以外，实验结果表明，同源性分量 b_1 的统计量特征的性能好于同源性分量 b_0 、 b_1/b_0 ，这种趋势与临床分级标准相符合。即，相比于细胞核连通分量数量的差异，腺体分化程度高低导致的腔体是否完整进而导致的孔洞数量差异在格里森分级中起着更为关键的作用，这表明本文提出的 SRHP 特征可以捕获腺癌中腺体结构中的拓扑排列，能够实现对

腺体分化程度的定量描述。

3.4.4 敏感性分析

本实验的目的是对 wKNN 分类器的超参数 K 值进行敏感性分析。此外，本实验还评估了 SRHP 方法关于是否采用 H 染色通道这一策略的敏感性。图 3.9 展示了超参数 K 在区间[3, 20]内取不同值的分析结果，表 3.4 展示了基于 H 染色通道和基于原始 H&E 染色图像的 SRHP 特征的性能敏感性分析结果。

从图 3.9 实验结果发现，即便超参数 K 处在不同的取值，本文提出的 SRHP 方法的分级 AUC 指标值稳定在 0.96 左右，其他指标也达到了几乎一致的稳定性能。这表明本文提出的方法对分类器的超参数取值不敏感，进一步表明 SRHP 特征的鲁棒性。在本章所有实验中，将 K 值都设置为 10。

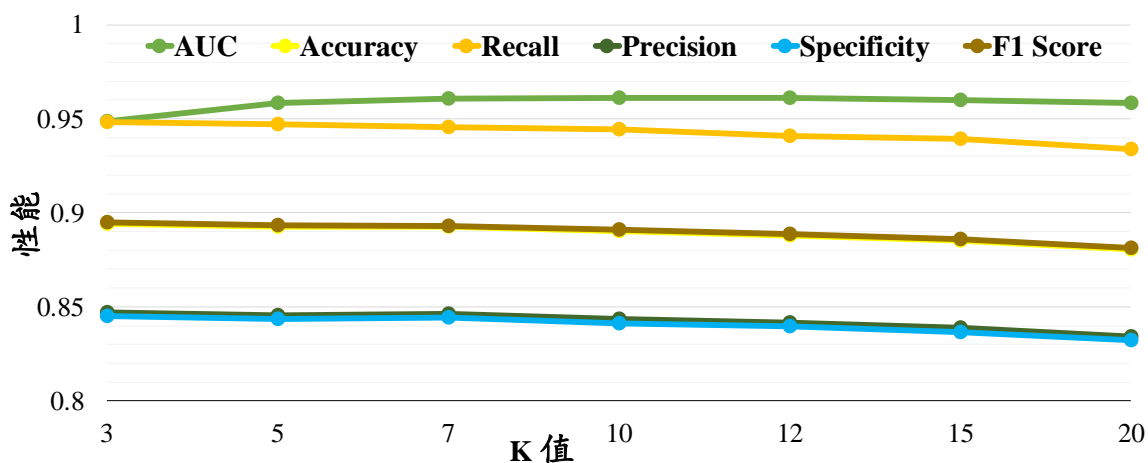


图 3.9 本文 wKNN 分类器模型的超参数 K 取不同值时的性能敏感性分析结果。K 值取值为{3, 5, 7, 10, 12, 15, 20}。

表 3.4 本文提出的 SRHP 方法关于是否采用 H 染色通道的敏感性分析评估结果。原始图像、H 染色图像所在行定量结果分别为基于原始 H&E 染色图像和基于 H 染色通道图像获取同源性序列特征实现自动分级的性能。（最佳结果以粗体显示）

	AUC	ACC	REC	PRC	SPC	F1 Score
原始图像	0.9479	87.33%	0.9011	0.8443	0.8478	0.8718
H 染色图像	0.9612	89.02%	0.9442	0.8435	0.8413	0.8910

表 3.4 的实验结果中，采取 H 染色通道获得的同源性分布统计表示特征在格里森分级上达到了 AUC 值 0.9612、F1 得分 0.8910 的结果，相比直接采用 H&E 染色图像的 0.9479 和 0.8718 具有更好的表现，说明本文采用 H&E 染色反卷积算法提取 H 染色通道

获得了更为清晰的细胞核和腺体结构，这对源性分量计算及格里森自动分级起到了积极作用。

3.4.5 全景病理图像预测

本实验的目的在于定性评估 SRHP 方法在全景病理切片上进行自动分级的可行性。图 3.10 展示了原始 H&E 染色全景病理图像和基于卷积神经网络方法构建的良恶性分类模型在整张全景病理图像上预测的恶性类别的概率图。图 3.11 展示了在图 3.10 恶性类别概率图的基础上对恶性图像块应用 SRHP 模型进行格里森三级四级区分得到的全景预测结果。

从图 3.10 的恶性类别概率图看出，尽管存在部分边缘腺体区域被检测为恶性，但良恶性分类模型的预测结果和病理学家勾画的格里森三级、四级感兴趣区域仍具有较高一致性，甚至，从图 3.10(b)中右下角放大显示的白色方框中观察到，该模型能识别到病理学家遗漏勾画的恶性区域（已和病理学家沟通证实），这表明基于卷积神经网络方法构建的良恶性分类模型达到了很好的区分效果，为接下来的基于 SRHP 模型进行格里森分级奠定坚实的基础。图 3.11 的定性结果表明，SRHP 模型在全景病理图像上达到了较好的预测。观察发现，SRHP 方法能很好地识别出格里森四级区域而格里森三级区域预测较差，导致了更多的假阳性区域，这与图像块层面定量结果中精度较低这一结论一致。

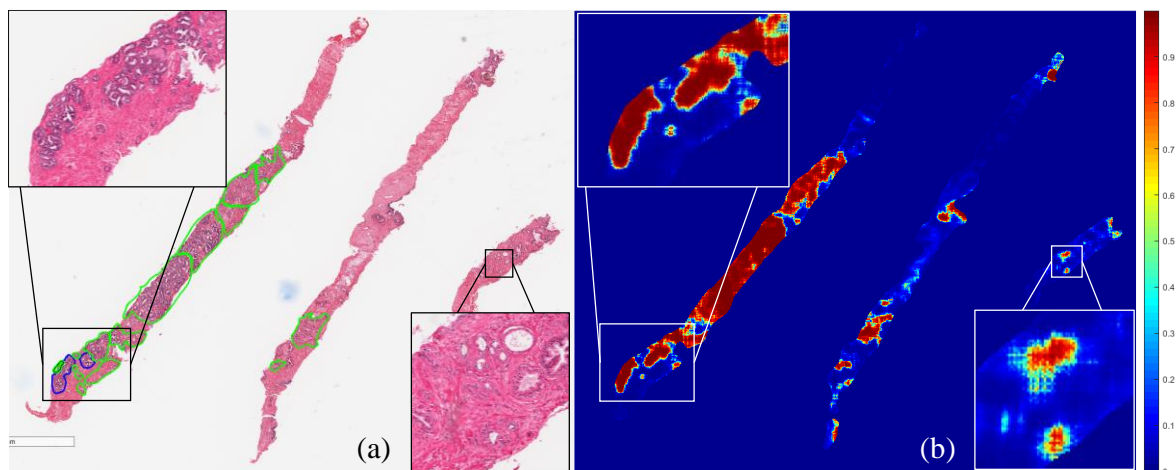


图 3.10 全景病理图像的良恶性预测结果。(a)为原始 H&E 染色全景病理图像；其中绿色和蓝色轮廓分别表示病理学家勾画的格里森三级、四级感兴趣区域。(b)为全景病理图像上对应的预测为恶性类别的概率图。为便于观察，选定两个区域用白色方框进行放大显示。

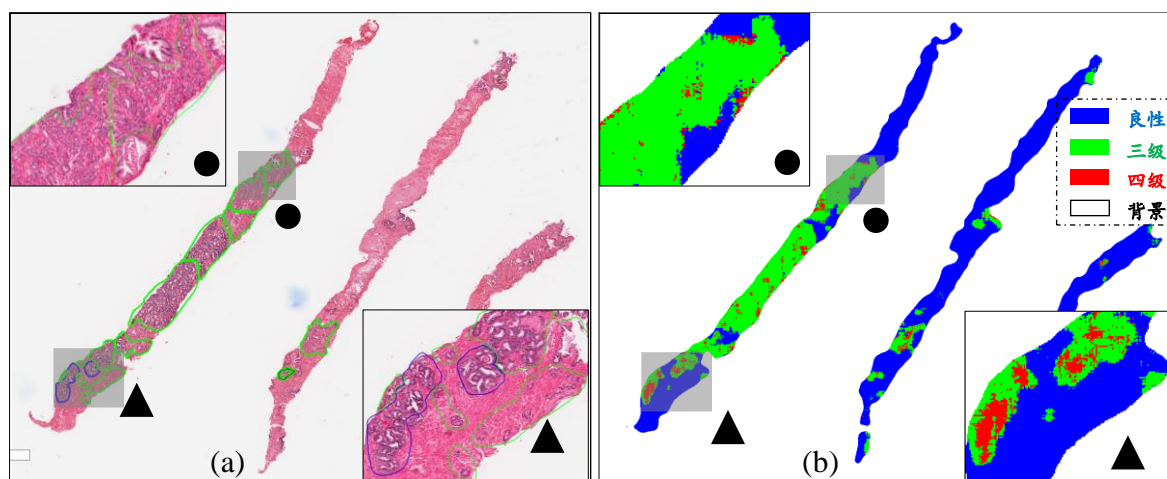


图 3.11 基于良恶性预测的全景病理图像格里森三级四级预测结果。(a)为原始 H&E 染色全景病理图像；(b)为基于本文提出的 SRHP 方法进行全景病理图像上的格里森三级四级预测结果。其中不同类别用 RGB 颜色编码表示，白色、蓝色、绿色和红色分别表示背景、良性、格里森三级和四级。为便于观察，选定两个区域用方框进行放大显示。

3.5 本章小结

本文提出了一种基于同源性分布统计表示(SRHP)的前列腺癌自动格里森分级方法。该方法首先利用同源性分布算法计算病理 H 染色图像的同源性分量 b_0 、 b_1 ，用以描述腺体周围细胞核间的拓扑排列以定量化腺体结构的分化程度；然后针对获得的同源性序列特征用 7 维分布统计量进行二次表征以捕捉格里森三级四级的不同规律性；最后采用加权 K 近邻分类器算法对病理图像块的同源性分布统计表示特征进行建模以实现腺癌病理图像的自动分级。

本文使用来自宾夕法尼亚大学附属医院的实验数据对 SRHP 分级方法进行评估。图像块层面和全景病理切片层面的定量评估结果都表明，相比于有监督学习方法 DLGg、无监督学习算法 SSAE 和传统的病理组学特征方法 MATF，本文提出的 SRHP 分级方法性能表现更出色，特征具有更好的全局性和泛化性。此外，特征分析和敏感性分析评估结果表明，本文提出的统计表示特征不仅与病理诊断标准相符合，而且对分类器的超参数取值更鲁棒。定性的全景病理图像预测实验结果表明，SRHP 方法应用到全景病理切片上进行等级预测具有一定的可行性，可以被用作前列腺癌病理图像格里森三级、四级分级的临床决策支持工具。

尽管如此，本文的方法仍有一些局限性。本文的 SRHP 模型在精度这一指标上结果较差，存在较高的假阳性预测。在最新的分级指南中，腺体结构的筛状模式起着重要作

用^[55]，但本文没有特定地关注 SRHP 方法处理筛状腺体结构的评估表现。另外，由于同源性分布算法需要对 H 染色图像遍历所有阈值进行二值化，导致同源性分量 b_0 、 b_1 的计算耗费了大量的时间；这可以通过调整阈值遍历的步长来缓解。此外，本文的方法仅限于在单一数据单位评估，无法验证其在不同数据中心的泛化性，未来可以采集更多医院的数据进行外部验证及算法的更新优化。

总之，本文提出了一种新的特征表示方法（SRHP：同源性分布统计表示），用于区分前列腺癌穿刺活检全景病理图像上的格里森三级、四级区域。作为一种新的计算病理学特征，其能够捕获腺体周围细胞核的拓扑排列以量化腺体结构的分化程度，不仅具有可区分性，还具有直观的临床可解释性。

第四章 基于多任务学习和先验知识的结直肠腺癌良恶自动评估方法

4.1 研究动机

对于结直肠腺癌而言,尽管目前有许多分级系统,但缺乏一个统一使用的分级标准。大多数系统依据腺体结构的分化程度和形态将肿瘤分为3个等级或4个等级,从等级1到等级4分别为分化良好、中等分化、分化差和不分化。其中,分化良好的等级1可以被归为良性,中等分化、分化差和不分化的等级2、3、4被归为恶性。

近年来,深度学习方法在病理图像分析中获得了优越的性能表现^[56],这使其在计算病理领域应用广泛^[34-35],但如前述章节中讨论,大部分深度学习方法的黑箱性质引发了业界关于可解释性和临床应用的讨论^[17]。因此,一些研究尝试对自动分割的腺体结构进行定量描述再进行恶性程度评估分级。Xu 等人^[42]采用了测地线活动轮廓模型对前列腺癌中的腺体结构进行分割并进行分级。Ali 等人^[39]提出基于形状先验的自适应主动轮廓算法对腺体结构的细胞核分割以进行后续的分级工作。Awan 等人^[44]首先对腺体结构分割然后提出最佳匹配指标进行腺体结构统计度量以实现腺癌的自动分级。Bulten 等人^[33]首先训练腺体分割模型,然后对整张穿刺病理切片进行腺体预测,依据整张切片中不同腺体结构的分化程度计算百分比以确定患者的格里森分级结果。

启发于对腺体结构自动分割结果进行精确描述以进行恶性程度自动评估分级的研究思路,本文构建了一个端对端的自动分割分级多任务模型,并将腺体结构自动分割结果这一先验知识编码为空间注意力融合到自动分级分支中以约束模型推理偏好。

4.2 研究方法

本文所提出的基于多任务学习和先验知识的结直肠腺癌良恶自动评估方法,旨在通过对腺癌病理图像的腺体结构进行自动分割进而作为空间注意力以约束结直肠腺癌的良恶自动分级。该方法包含了三个主要部分:a).特征学习骨干网络;b).分割分类多任务分支;c).腺体结构空间注意力约束。图4.1展示了该方法的整体流程图。

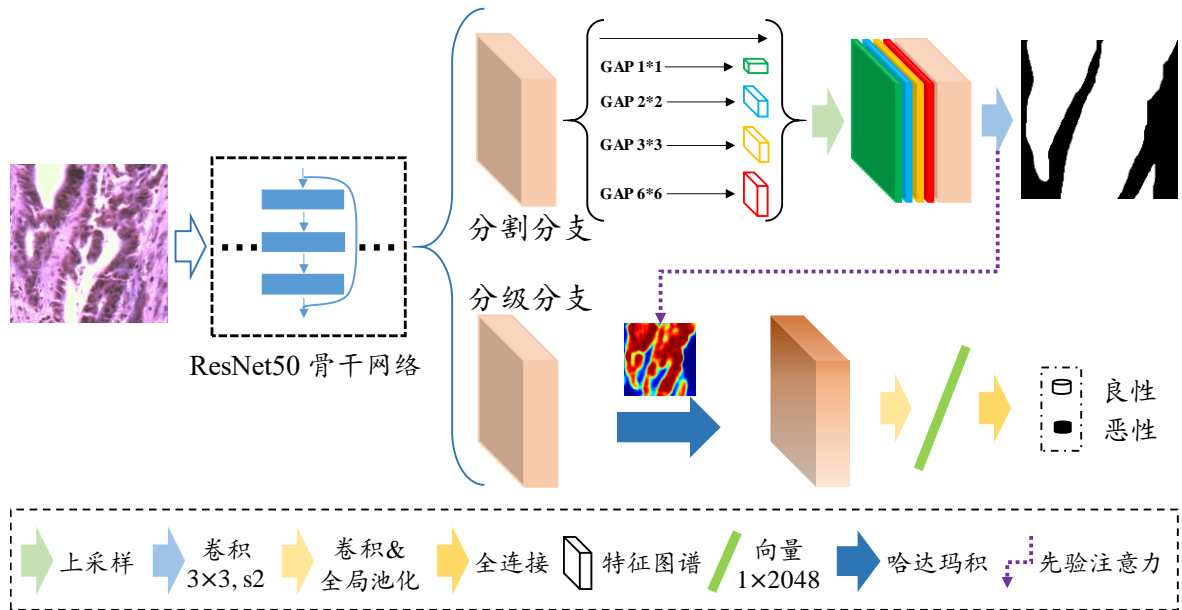


图 4.1 本文提出的基于多任务学习和先验知识的结直肠腺癌良恶自动分级模型。该模型由特征提取骨干网络和两个并行任务分支组成。骨干网络提取到的特征图同时被用于分割分支和分级分支以实现腺体结构分割和病理图像良恶分类。其中，先验注意力信息传递部分负责将腺体结构预测概率图作为空间注意力约束良恶分级分支。

4.2.1 特征学习骨干网络

特征学习骨干网络被设计用于提取腺癌病理图像的高维语义特征。本文采用基于残差模块^[57]的改进 ResNet50 网络的骨干层作为特征提取部分。残差块是 ResNet50 网络中的基本单元。每个残差块都由 1×1 、 3×3 、 1×1 卷积层堆叠组合而成，同时包含一个从输入到输出的恒等连接将输入特征与输出特征逐像素相加。其中， 1×1 卷积层负责对特征进行降维和升维以减少参数计算。

整个特征提取部分如图 4.1 中 ResNet50 骨干网络所示，对于输入图像 $X \in \mathbb{R}^{3 \times h \times w}$ ，首先由 1 个 3×3 卷积核大小步长为 2 的卷积层和 2 个 3×3 卷积核大小、步长为 1 的卷积层构成输入处理层，然后 3×3 卷积核大小、步长为 2 填充为 1 的最大池化层执行下采样，此时获得的特征图 F_1 尺寸减小至输入图像尺寸的四分之一。然后，分别包含了 3、4、6、3 个残差块的四组卷积组部分继续对上述特征图 F_1 进行卷积运算，其中，第二组卷积组部分利用步长为 2 的卷积实现下采样将特征图尺寸减少至特征图 F_1 的二分之一。为了获得更大的感受野以关注腺体结构的尺度信息，在第三、第四组卷积组采用膨胀卷积策略，膨胀系数分别为 2、4。因此，输入图像 X 经过以上操作的运算后得到其高维语义特征表示 F_5 。整个特征学习骨干网络结构具体设置如表 4.1 所示。

表 4.1 特征学习骨干网络结构设置。

网络层名称	网络层具体设置结构 (s 表示步长, p 表示填充, d 表示膨胀系数)	特征通道数	输出尺寸大小
输入		3	$h \times w$
输入处理层	$\begin{bmatrix} 3 \times 3, & s = 2 \\ 3 \times 3, & s = 1 \\ 3 \times 3, & s = 1 \end{bmatrix}, \begin{cases} p = 1 \\ d = 1 \end{cases}$	128	$\frac{h}{2} \times \frac{w}{2}$
池化层	3×3 maxpool, $\begin{cases} s = 2 \\ p = 1 \\ d = 1 \end{cases}$	128	$\frac{h}{4} \times \frac{w}{4}$
残差块 卷积组一	$\begin{bmatrix} 1 \times 1, & p = 0 \\ 3 \times 3, & p = 1 \\ 1 \times 1, & p = 0 \end{bmatrix} \times 3, \begin{cases} s = 1 \\ d = 1 \end{cases}$	256	$\frac{h}{4} \times \frac{w}{4}$
残差块 卷积组二	$\begin{bmatrix} 1 \times 1, & s = 1, & p = 0 \\ 3 \times 3, & s = 2, & p = 1 \\ 1 \times 1, & s = 1, & p = 0 \end{bmatrix} \times 1, \{d = 1\}$ $\begin{bmatrix} 1 \times 1, & p = 0 \\ 3 \times 3, & p = 1 \\ 1 \times 1, & p = 0 \end{bmatrix} \times 3, \begin{cases} s = 1 \\ d = 1 \end{cases}$	512	$\frac{h}{8} \times \frac{w}{8}$
残差块 卷积组三	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 6, \begin{cases} s = 1 \\ p = 2 \\ d = 2 \end{cases}$	1024	$\frac{h}{8} \times \frac{w}{8}$
残差块 卷积组四	$\begin{bmatrix} 1 \times 1 \\ 3 \times 3 \\ 1 \times 1 \end{bmatrix} \times 3, \begin{cases} s = 1 \\ p = 4 \\ d = 4 \end{cases}$	2048	$\frac{h}{8} \times \frac{w}{8}$
输出			

4.2.2 分割分类多任务分支

腺体结构分割分支

分割分支用于病理图像中腺体结构的自动分割。为了获取图像的上下文信息和全局信息,本文采用了金字塔池化模块^[58]。金字塔池化模块由四个并行的全局平均池化操作组成,随后是 1×1 卷积用以特征降维,并行输出特征图组 F_{set} :

$$F_{set} = \{F_{pi} \in \mathbb{R}^{512 \times i \times i}\}_{i=1}^U, \quad U := \{1,2,3,6\} \quad (4.1)$$

接下来,采用双线性插值操作对输出特征图组进行上采样以恢复至输入特征图 F_s 尺寸大小,因此获得新的特征图组 $\{F_{bi} \in \mathbb{R}^{512 \times \frac{h}{8} \times \frac{w}{8}}\}$ 。为了获得数据的多尺度信息,本文将 $\{F_{bi}\}$ 与输入特征图 F_s 进行特征融合后用作分割预测层。分割预测层由一个 3×3 卷积核大

小、填充为 1 的卷积层和一个 1×1 卷积核大小的卷积层组成，用于将融合特征变换为特征层预测 $P_{map} \in \mathbb{R}^{2 \times \frac{h}{8} \times \frac{w}{8}}$ 。最后，采用 softmax 层和双线性插值操作对特征层预测 P_{map} 进行 8 倍上采样，最终获得和输入图像大小一致的分割预测结果。

先验约束分级分支

先验约束分级分支实现了病理图像的不良恶性分类。这里，高维特征 F_s 首先被来自分割分支的特征层预测 P_{map} 进行了约束，具体细节如本章节 4.2.3 中所述。然后，3 个 3×3 卷积核大小的堆叠卷积层对约束后的语义特征进行特征变换，全局平均池化层和全连接层随后将特征拉成向量 V_{fc} 以进行特征降维。最后向量 V_{fc} 经过线性映射和 softmax 归一化得到分级预测概率 V_g ：

$$V_g \in \mathbb{Y}^{2 \times 1}, \quad \mathbb{Y} := [0,1] \quad (4.2)$$

4.2.3 腺体结构空间注意力约束

在腺体结构分割分支自动预测的过程中，本文提取其特征层预测 P_{map} 用于对分级分支的输入特征 F_s 进行特征重标定，以基于空间注意力约束来指导良恶性分级分支的模型推理。这也是本文网络模型在良恶性分级上具备可解释性预测价值的重要原因。这里，对特征 F_s 的空间注意力约束形式化为：

$$F_p = \left[F_s^k * \frac{e^{x_1}}{e^{x_0} + e^{x_1}} \right] \quad k \in \{1, 2, \dots, n\}, x \in P_{map} \quad (4.3)$$

其中， n 表示特征 F_s 的通道数量。 $\frac{e^{x_1}}{e^{x_0} + e^{x_1}}$ 表示将分割分支的特征层预测 P_{map} 经 softmax 函数映射为 $P_s \in \mathbb{Y}$ 并提取其归属于腺体结构的通道 P_s^1 。此时， P_s^1 所有像素值被映射到区间 $[0,1]$ 内，且空间尺寸大小与特征 F_s 一致，因此作为特征 F_s 的空间注意力与每一通道 F_s^k 进行哈达玛乘积运算实现特征重标定。由于 P_s^1 是归属于 P_s 的腺体结构通道，该特征重标定过程也是 F_s 被腺体结构这一先验知识显式约束的过程，获取的特征 F_p 作为分级分支的输入直接指导了后续分级分支的良恶评估预测。

4.2.4 多任务损失

优化算法的总体目标为：使任务的损失函数最小化。因此，损失函数在网络模型参数的收敛优化中起到了重要的作用。在本文的多任务学习卷积网络模型中，为了使分割

分支和分级分支联合学习，对于输入数据对 $\{X, c, s\}$ ，定义每张输入图像样本 X 的多任务损失函数 L_{mt} 为：

$$L_{mt} = L_{ce}(c, c^*) + \lambda \sum_{j=1}^w \sum_{i=1}^h L_{ce}(s_{ij}, s_{ij}^*) \quad (4.4)$$

这里，分割分支和分级分支都采用交叉熵损失函数 L_{ce} 进行优化。其中， c^* 表示分级分支的类别预测概率， s^* 表示分割分支的全像素预测概率图；而 c 和 s 分别代表了输入图像对应的类别标签和腺体结构掩膜。 (i, j) 表示腺体结构掩膜中的某一坐标位置；参数 λ 为平衡因子，用于调节多任务学习中的损失偏置。

4.3 实验设计

本文提出了基于多任务学习和先验知识的结直肠腺癌良恶自动评估方法。为表明该方法在腺癌病理图像上自动分级的可行性和优越性，将该方法在公开数据上进行实验对比分析。

4.3.1 实验数据

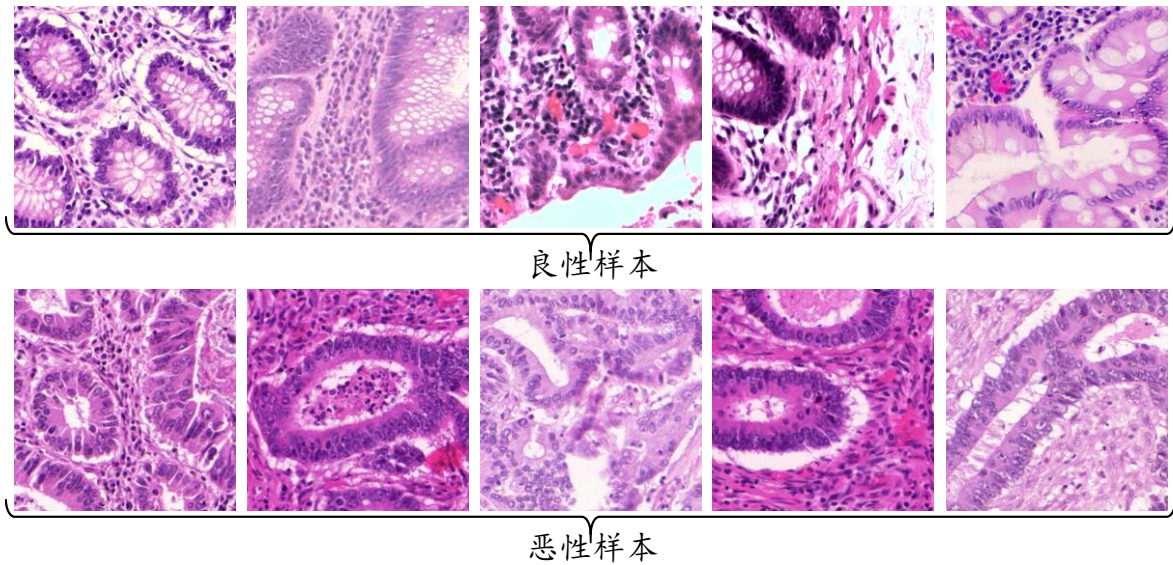


图 4.2 本文采用的 Warwick-QU 结直肠腺癌数据集中良恶性样本示例。

本章实验使用的是 Warwick-QU 结直肠腺癌公开数据集^[59]。该队列共包含 165 张由 Zeiss MIRAX MIDI 扫描仪获取的 H&E 染色图像、对应的腺体区域标记以及诊断信息（如良性、恶性）。其中，图像的最高放大倍数为 20 倍，切片实际大小（微米）与图像像素比为 0.62，队列平均像素大小约在 750×500 。整个数据集被分为训练集和测试集，训练集包含 37 张良性和 48 张恶性图像，测试集包含 37 张良性和 43 张恶性图像。部分

良恶性样本示例如图 4.2 所示。

这里,考虑到数据量与网络模型的匹配,对于每张病理图像和对应的腺体掩膜图像,本文从左上、左下、右上、右下、中心五个位置裁剪 320×320 大小的小图像对,并进行了颜色标准化^[60]、旋转增强、镜像增强以及空白块过滤等预处理操作,如图 4.3 所示。最终,本文基于 Warwick-QU 单任务分割数据集构建了多任务分级分割数据集,包括有 984 张良性和 1386 张恶性训练图像以及 1050 张良性和 1182 张恶性测试图像,每张图像都对应有各自的腺体结构掩膜。

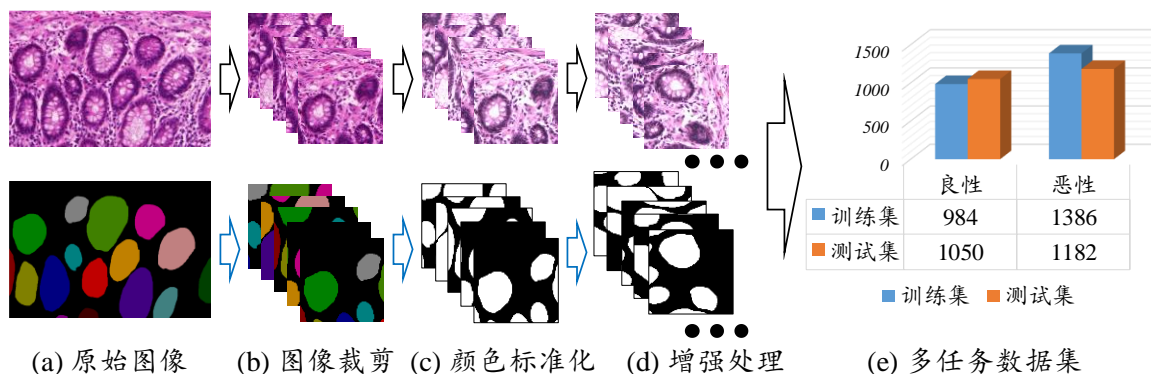


图 4.3 本文从 Warwick-QU 结直肠腺癌单任务分割数据集构建为多任务数据集的处理策略。

4.3.2 实验细节

本文搭建了如图 4.1 所示的多任务学习卷积神经网络模型。为了保证多任务网络模型优化的稳定性,按如下方式进行实验。首先,采用 kaiming 均匀策略对分割分支和分级分支进行参数初始化;然后,冻结分级分支权重参数,采用 Adam 优化器,学习率 0.001 仅对分割分支训练 10 个迭代周期。最后,解除分级分支参数冻结,同时将腺体结构特征层预测 P_{map} 传递至分级分支与 F_5 进行特征运算,此时,对整个多任务网络结构继续训练 100 个迭代周期。初始学习率仍为 0.001,每 10 个迭代周期对学习率进行一次减半衰减,其中,多任务损失平衡因子 λ 设置为 1。

4.3.3 对比方法设计

本文采用公开数据集进行实验,但公开方法仅评估了腺体结构的实例分割结果。而本文更多从分级层面进行验证。因此,为了表明本文所提方法的有效性和优越性,将该方法与其他现有的应用于病理图像分类分级的几种主流方法在以上多任务数据上进行分级实验对比。这三种方法简要说明如下:

- 传统的病理组学手工特征方法^[61]——表示为 Pathomics。包含细胞核形态、细胞图、

纹理等共 871 维病理组学特征。

- 端到端的无监督学习方法^[54]——Xu 等人提出的堆叠式稀疏自动编码器算法(SSAE), 通过滑动窗策略将每个图像块分类为细胞核或非细胞核, 实现病理图像的细胞核检测。
- 端到端的监督学习方法——将深度卷积神经网络方法直接应用到 Warwick-QU 数据进行腺癌良恶性分级。本文将这种方法表示为 DL-Baseline(Deep Learning Baseline)。

4.3.4 性能评估

为了公平准确地比较不同分级方法的性能, 本文采用准确率(ACC)、召回率(REC)、精度(PRC)、特异性(SPC)、F1 得分(F1 Score)、受试者操作特征曲线下面积(AUC)等不同指标进行量化评估。此外, 对于腺体结构自动分割分支, 本文采用交并比(Intersection over Union, IoU) 指标进行评估。

4.4 实验结果及讨论

4.4.1 相关方法自动分级对比

为了展现本文提出的基于腺体结构这一先验知识进行空间注意力约束的良恶分级方法的有效性, 本实验对比了 Pathomics、SSAE、DL-Baseline 和本文方法。表 4.2 展示了几种方法在各指标上的定量评估结果。

表 4.2 本文提出的基于腺体结构空间注意力约束的良恶分级方法与其他对比方法的定量评估结果。

(最佳结果以粗体显示)

分类	AUC	ACC	REC	PRC	SPC	F1 Score
Pathomics	0.8725	83.68%	0.8925	0.8259	0.7682	0.8579
SSAE	0.5951	59.95%	0.6091	0.6250	0.5885	0.6169
DL-Baseline	0.9780	93.82%	0.9036	0.9780	0.9771	0.9393
Ours	0.9971	97.04%	0.9898	0.9559	0.9486	0.9726

从表 4.2 可以看出, 相较于其他方法, 本文提出的方法在 AUC、ACC 指标上达到了最佳性能, 分别为 0.9971 和 97.04%, 这表明了本文提出的先验知识约束的多任务模型具有更好的优势。和 DL-Baseline 方法相比, 本文方法基于腺体结构空间约束, 具有更好的模型解释性。Pathomics 手工特征方法虽然也极具临床可解释性, 但在本数据上 AUC

指标仅为 0.8725，预测性能差于本文方法，而且本文方法实现了端对端的预测，更有潜力成为临床辅助诊断工具。此外，SSAE 方法预测性能最差，这说明使用无监督自编码器方法在本数据中很难学到更有目标针对性的区分性特征。

为了更好展示不同方法的区分能力，本文对以上四种方法的特征使用 tSNE 方法^[62]进行降维可视化。其中，对于 SSAE 方法，使用最后一个自编码器编码部分的输出特征；对于 DL-Baseline 和本文方法，使用全局池化层输出的向量特征。图 4.4 展示了四种方法的特征降维可视化结果，其中蓝色、红色点分别表示了标签为良性、恶性的样本。从图中可以直观地发现，DL-Baseline 方法和本文方法都能更好地将数据区分开来，但本文方法更胜一筹，而 SSAE 方法和 Pathomics 方法对数据的区分能力较差，这一结论和表 4.2 中的定量评估结果一致。

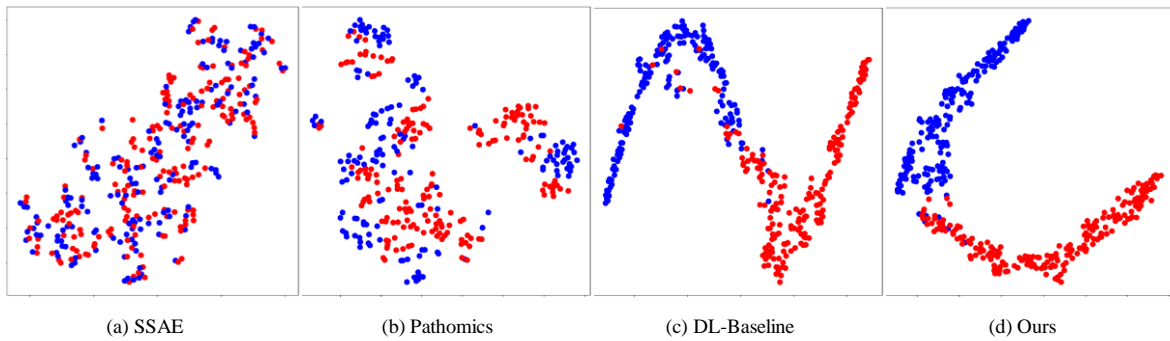


图 4.4 四种方法模型的特征 tSNE^[62]可视化结果。可视化图中蓝色点代表实际为良性的样本，红色点表示恶性样本。

4.4.2 多任务学习和先验知识的有效性评估

为了分别展现本文提出的多任务学习卷积网络框架和腺体结构这一先验信息在病理图像良恶自动分级上的有效性，本实验构建了其他三个网络模型与本文的基于先验信息约束的多任务模型（表示为 MT-P (Multi Tasks on Prior)）进行比较。首先，本实验构建了没有任何空间注意力约束的分类模型，即基于改进 ResNet50 骨干网络的单任务分类 DL-Baseline 模型，这里表示为 ST-no-P (Single Task without Prior)。然后，本文构建了 ST-P (Single Task on Prior) 模型，即，将数据集中的腺体结构掩膜直接转换为空间注意力约束特征的单任务分类模型。第三个模型则为多任务学习模型，该模型中仅包含腺体结构分割分支和分级分支，分割分支的特征层预测不作为空间注意力去约束分级分支，该模型称为 MT-no-P (Multi Tasks without Prior)。表 4.3 报告了以上四种网络结构模型的性能。

从表 4.3 可以看出, 本文提出的基于注意力约束的多任务模型 MT-P 好于其他三个模型, 达到了最高 0.9971 的 AUC、97.04% 的 ACC 和 0.9726 的 F1 得分, 这说明了本文方法采用先验约束和多任务学习架构的整体有效性。同样, 从实验结果中看出, 不仅 MT-P 模型优于 MT-no-P 模型, 而且 ST-P 模型要优于 ST-no-P 模型, 这表明不论对于单任务还是多任务网络而言, 加入了腺体结构这一先验信息的约束的确对自动良恶分级具有积极的指导作用, 这单独验证了本文提出的先验空间注意力约束的有效性。此外, 分析实验结果发现, 在不考虑是否加入先验信息这一条件时, MT-P 模型更优于 ST-P 模型, 而且 MT-no-P 模型的分级性能也要超过 ST-no-P 模型, 这表明本文的多任务模型的共享网络层可以学习到鲁棒的特征表示, 进而提升模型性能。这与 Vandenhende 等人^[48]的结论相一致: 当多个任务之间具有信息互补或者正则作用, 那么通过共同优化的方式能够提高共享特征的鲁棒性和整个模型的泛化性。

表 4.3 本文提出的多任务学习卷积网络框架和腺体结构这一先验信息在病理图像良恶自动分级上的有效性验证。(最佳结果以粗体显示)

分类	AUC	ACC	REC	PRC	SPC	F1 Score
ST-no-P	0.9780	93.82%	0.9036	0.9780	0.9771	0.9393
MT-no-P	0.9828	95.97%	0.9289	0.9946	0.9943	0.9606
ST-P	0.9939	96.77%	0.9746	0.9648	0.9600	0.9697
MT-P (Ours)	0.9971	97.04%	0.9898	0.9559	0.9486	0.9726

同样地, 本文对以上四种网络模型提取全局池化层输出的向量特征进行 tSNE 降维可视化以观察各自模型下数据的分布情况。图 4.5 展示了四种网络模型的特征可视化结果, 可视化图表明了本文 MT-P 方法在该数据上具有最佳的区分性。

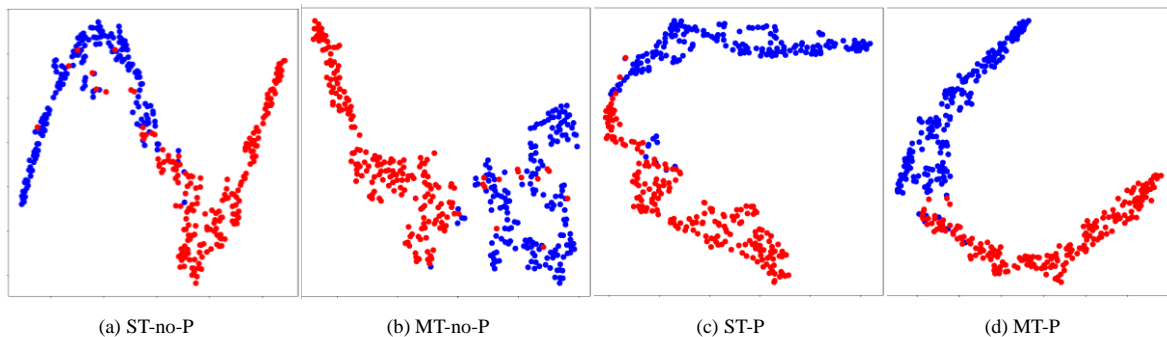


图 4.5 ST-no-P、MT-no-P、ST-P 和 MT-P 模型中的特征 tSNE 可视化结果。这里, 特征为网络模型全局池化层的输出。可视化图中蓝色点代表实际为良性的样本, 红色点表示恶性样本。

4.4.3 先验知识有效性的可视化分析

进一步地，为了更好地验证本文提出的先验知识约束思想在病理图像良恶分级上的有效性，本文采用类激活映射方法^[63]来可视化分析分级分支中的类激活热图。图 4.6 展示了不同类别图像样本分别被 MT-no-P 模型和 MT-P 模型预测的腺体结构分割结果、分级预测结果和分类器类激活热图。

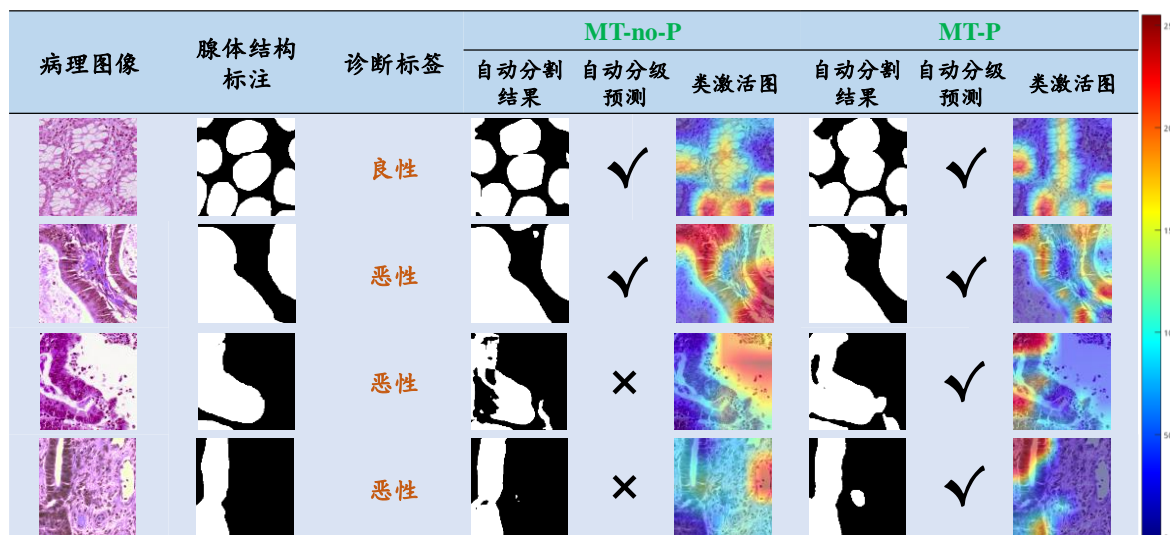


图 4.6 本文提出的先验知识约束思想的有效性可视化分析结果。这里，针对不同类别图像样本展示了其分别被 MT-no-P 模型和 MT-P 模型预测的腺体结构分割结果、分级预测结果和分类器的类激活热图。

从图中可以看出，对于前两个图像样本来说，MT-no-P 模型和 MT-P 模型均能正确预测类别并较好地分割腺体结构，而且，对应的类激活热图显示，不论分级分支是否被腺体结构约束，两个模型都能关注腺体区域（类激活热图中的红色部分）进而做出正确预测，但基于腺体先验信息约束的 MT-P 模型关注的区域更加精细化。而对于后两个图像样本，其腺体结构也都能被很好地分割，但却被 MT-no-P 模型错误分类而被 MT-P 模型正确预测。观察对应的类激活热图发现，MT-no-P 模型错误地关注了非组织区域（例如空腔和背景）导致识别错误，而 MT-P 模型则能从腺体结构约束中受益，从而关注腺体组织区域，做出正确预测。这说明在腺体结构先验信息的约束下 MT-P 模型的分级分支更具有针对性。

此外，为了探究先验知识约束在特征通道层面的有效性，本文选取良恶类别下各一个样本送入本文的先验知识约束多任务模型进行预测，同时提取腺体结构空间注意力图以及分级分支中的约束前和约束后的特征图谱进行可视化。

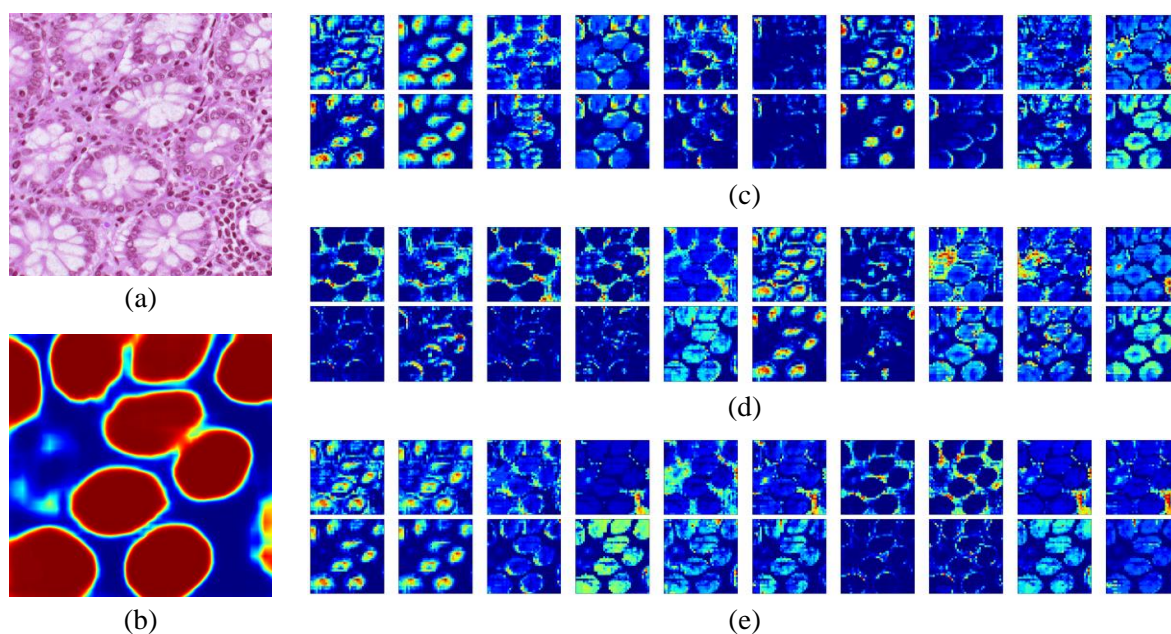


图 4.7 本文先验知识约束的多任务网络模型的分级分支特征图可视化结果。(a)为良性病理图像样本；(b)为分割分支预测的腺体结构，其被作为空间注意力约束分级分支；(c)、(d)、(e)分别为选取的通道 0~10、70~80、130~140 的特征图谱；其中，上部分为约束前的特征图，下部分为约束后的特征图。

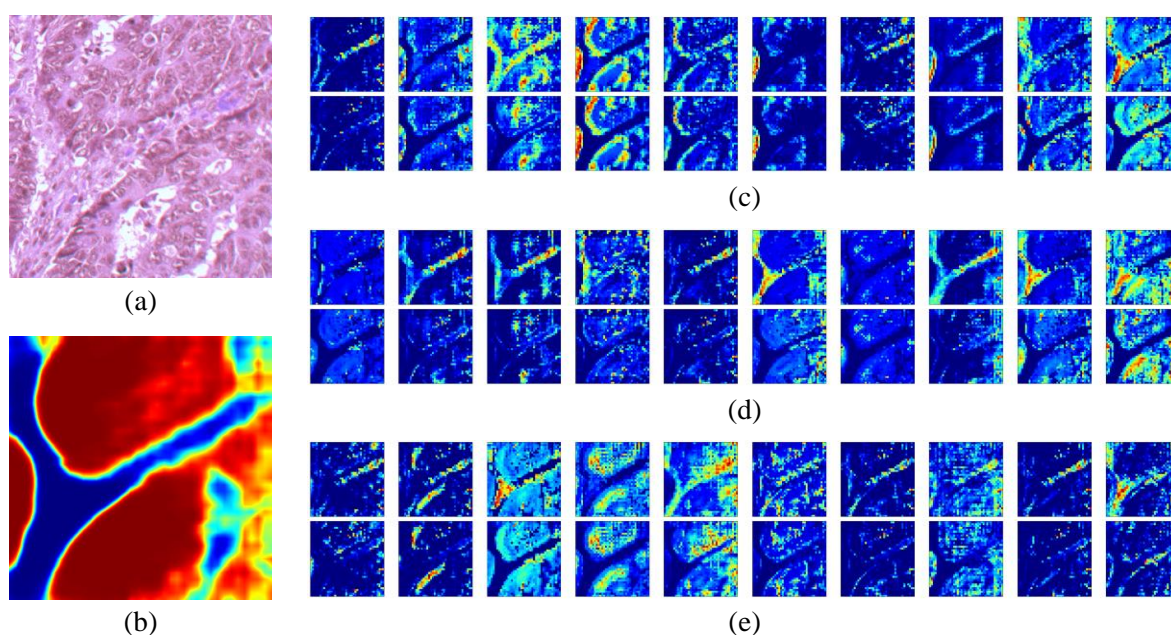


图 4.8 本文先验知识约束的多任务网络模型的分级分支特征图可视化结果。(a)为恶性病理图像样本；(b)为分割分支预测的腺体结构，其被作为空间注意力约束分级分支；(c)、(d)、(e)分别为选取的通道 0~10、100~110、180~190 的特征图谱；其中，上部分为约束前的特征图，下部分为约束后的特征图。

图 4.7 和图 4.8 分别展示了良、恶性图像样本在分级分支中的特征可视化结果以及

腺体结构空间注意力图。这里，对 2048 维特征图谱按连续 10 维通道为一组，随机选取了 3 组进行可视化；如图 4.7 和图 4.8(c~e)所示，其中每一组上半部分为先验知识约束前的特征图，下半部分为先验知识约束后的特征图。从图中观察腺体结构部分可以发现，本文提出的先验知识约束在特征图层面更好地抑制了原有特征图谱中的目标无关噪声并强化了组织区域的重要性。

4.4.4 分割分支评估

分割分支的作用在于自动分割腺体结构并提供结构约束以指导分级分支预测。为了验证其有效性，本文对该分支的自动分割结果进行定量化评估。同时，本文构建了单任务分割网络（表示为 ST-Seg (Single Tasks for Segmentation)）与其进行比较。表 4.4 报告了不同模型对比的自动分割评估结果。图 4.9 展示了不同模型下的腺体结构自动分割结果。

表 4.4 本文提出的多任务学习模型中分割分支的性能对比评估结果。（腺体缩写为 G）

分割	mIoU	G IoU	Non-G IoU	mACC
ST-Seg	0.8415	0.8528	0.8303	0.9146
MT-no-P	0.8086	0.8196	0.7976	0.8958
MT-P (Ours)	0.8134	0.8319	0.7948	0.8954

从表中可以看出，就 IoU 和 mACC 指标而言，多任务模型（包括 MT-no-P 和 MT-P）性能稍差于单任务分割模型。一个可能的原因是，分级分支是图像层面的预测，而分割分支则是像素层面的预测，在联合损失优化时分割分支的参数更新相较而言被削弱，进而导致多任务模型中的分割分支性能下降。从图 4.9 的定性分割结果同样发现多任务模型稍差于单任务分割模型。即便如此，本文的目标是基于腺体结构约束实现腺癌自动分级，而章节 4.4.1 至章节 4.4.3 的分级性能评估都表明分割分支输出的腺体结构预测已经具备较好的先验约束能力。

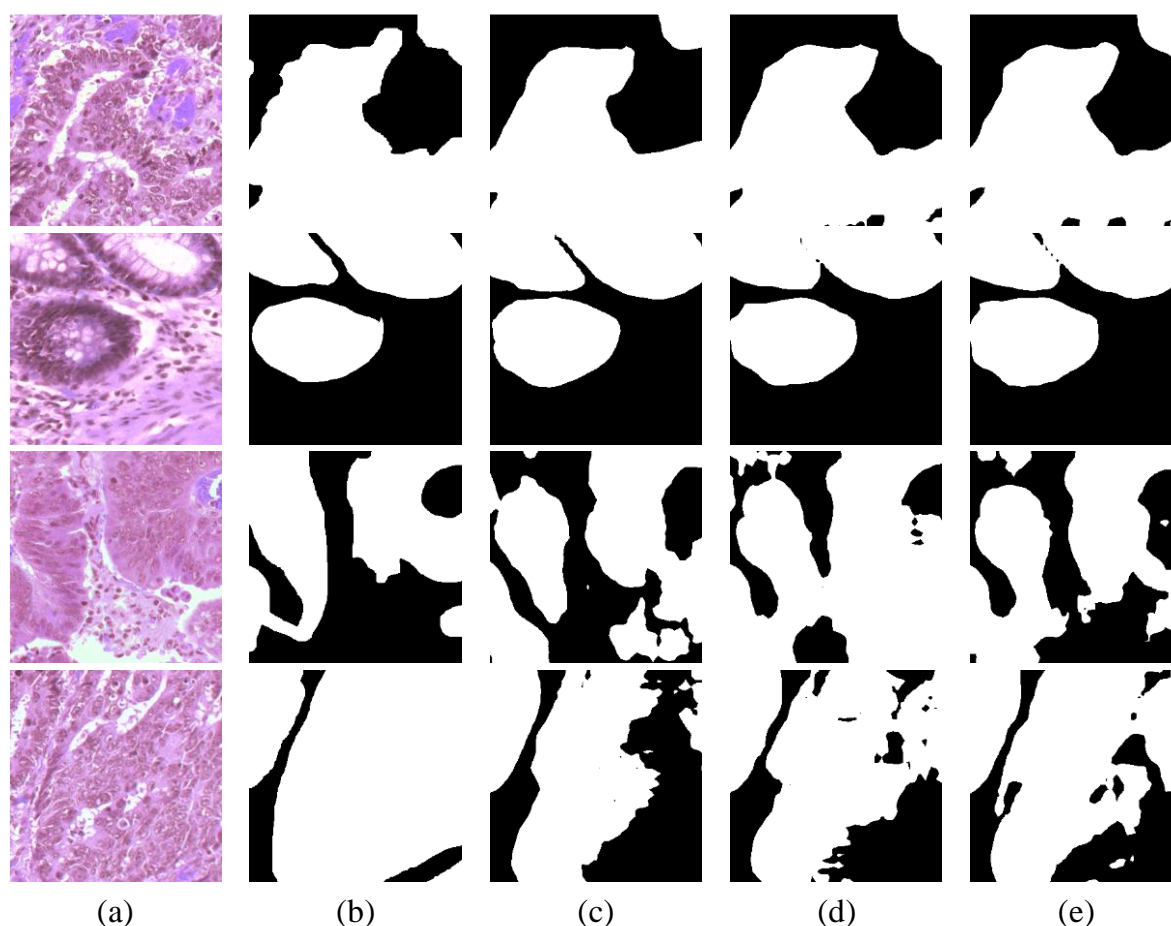


图 4.9 本文先验知识约束的多任务网络模型的腺体结构自动分割结果对比。(a)列为病理图像样本；(b)列为对应的腺体结构掩膜；(c)、(d)、(e)列分别为 ST-Seg 模型、MT-no-P 模型和本文 MT-P 模型的腺体结构自动分割结果。

4.5 本章小结

本文提出了一种基于多任务学习和先验知识的结直肠腺癌良恶自动评估分级方法。该方法首先利用特征学习骨干网络进行特征提取，提取到的特征分别被送入分割分支和分类分支进行腺体结构的自动分割和腺癌图像良恶自动分级；同时，分割分支的腺体结构特征层预测作为先验知识被编码为空间注意力以约束分级分支的推理偏好。

本文使用 Warwick-QU 结直肠腺癌公开数据集对该方法进行验证。不同方法的对比结果表明，相比于有监督学习方法 DL-Baseline、无监督学习算法 SSAE 和传统的病理组学特征方法 Pathomics, 本文提出的先验知识约束的多任务模型具有更出色的性能表现。tSNE 特征降维可视化也表明本文方法在该数据上具有更好的区分性。此外，消融实验分析、类激活热图、特征图谱可视化等评估结果都验证了本文提出的先验知识约束思想和多任务网络框架的有效性。

尽管如此，本文方法仍有一些局限性。本文的多任务网络模型中提供先验信息约束的分割分支性能稍逊色于单任务分割模型，后续可以通过调节多任务损失的平衡因子进行优化，同时可以对分割分支设计更有效的网络结构以提高自动分割性能。此外，本文方法仅在单一公开数据集上进行实验评估，尚未在更多公开数据集乃至医院临床数据上验证，未来可以尝试获取更多单位的数据进行外部验证、算法的更新优化以及全景病理切片预测。

总之，本文提出了一种先验知识约束的多任务网络模型，用于实现结直肠腺癌病理图像的良恶自动评估。该多任务模型可以基于腺体结构约束实现良恶预测，具有可解释性。而且该模型可以输出腺体结构分割结果，这使得其更有潜力成为临床辅助诊断工具。

第五章 总结与展望

5.1 本文工作内容及创新之处

在临床病理诊断中，腺体组织结构的分化程度是病理学家确定腺癌恶性程度乃至决策治疗方案的决定性因素。因此，本文从腺体结构分化程度这一病理诊断标准出发，分别从手工特征设计和先验知识建模两个角度实现腺癌恶性程度的自动评估分级。

手工特征设计角度，本文提出了基于同源性分布统计表示的前列腺癌自动格里森分级方法。该方法基于同源性分布算法获取病理图像的同源性分量，定量描述了腺体周围细胞核间的拓扑排列以表征腺体结构的分化程度。

先验知识建模角度，本文提出了基于多任务学习和先验知识的结直肠腺癌良恶自动评估方法。该方法将腺体结构自动分割和病理图像良恶分级诊断目标融合到一个端对端的网络模型中。而且，分割分支的腺体结构预测作为先验知识被编码为空间注意力融合到分类分支中以约束分级的推理偏好。

本文的主要贡献和创新之处在于：

1. 本文在病理图像分析领域提出了同源性分布统计表示这一计算病理学特征，该特征能够捕获腺体周围细胞核的拓扑排列，实现对腺体结构分化程度的定量化描述。
2. 本文基于同源性分布统计表示特征运用机器学习算法实现了前列腺癌的自动格里森分级，该方法应用到全景病理切片上进行等级预测同样具有可行性。
3. 本文对腺体结构这一先验知识进行建模，构建了基于先验知识约束的多任务学习卷积网络模型，该模型不仅实现了结直肠腺癌良恶自动评估分级，而且实现了腺体结构的自动分割。
4. 本文基于腺体结构这一病理诊断标准进行建模，通过病理图像分析方法实现了腺癌恶性程度的自动评估分级，为病理医生提供了可解释的辅助诊断工具。

5.2 今后工作的展望

本文工作虽然取得了一些成果，但仍有许多不足之处，未来可以从以下几个方面继续开展研究：

1. 同源性分布统计表示特征可以进一步优化提速，从而纳入病理组学以更好地适用于计算病理领域的图像分析计算任务。

2. 基于先验知识约束的多任务学习卷积神经网络模型目前仅在良恶类别腺癌数据上进行实验，后续可以在结直肠腺癌的高中低等级全景病理图像数据进行建模预测。
3. 未来可以采集来自不同机构、不同扫描设备以及不同器官腺癌的病理图像数据以验证本文方法在不同数据中心的泛化性，同时基于更多数据进行算法的更新优化、外部验证以及临床部署应用。

致 谢

在本篇论文完成之际，谨向给予我帮助的老师、同学、家人和朋友表示由衷的感谢。

首先我要感谢我的导师徐军教授，在我硕士阶段对我谆谆教导，不断给予我帮助，让我在科研道路上不断前行。感谢徐老师对本论文呕心沥血的指导，无论是选题、研究方向、研究内容，还是后期对本论文的修改，都凝聚着恩师的心血。徐老师在学术上的成就以及对待科研的精神深深影响着我，使我养成了严谨认真、一丝不苟的学习态度。徐老师在这几年的时间教会我做科研的方法和思维逻辑，这将使我受益终生。同时我要感谢日本大阪大学的 Kazuaki Nakane 教授和美国凯斯西储大学的 Anant Madabhushi 教授，在合作的过程中，他们给了我很多科研上的建议和指导。尤其是 Kazuaki Nakane 教授邀请我去东京明治大学交流访问，这将使我受益终生。

感谢研究生三年来同学们的帮助和照顾。感谢鲁浩达、帅惠、蔡程飞、徐海俊、马伟、谢嘉伟、赵增瑞、李宝明等众位师兄和王荃师姐，你们德才兼备、幽默风趣、亲切和善，每当我偏离学习方向的时候，你们给我警醒，把我拉回正确的轨道。你们总能指导我一些做人做事的方法，使我受益良多。感谢同届的顾松、王雷、徐春燕，你们让我的三年时光充实而美好，这将是我一生的珍藏。同时也感谢我的师弟师妹们，在科研、体育、娱乐等众多领域，伴我同行。

感谢我的父母、亲人，是你们给了我良好的成长环境，也给我最大的支持和鼓励。不管何时何地，你们总提醒我去追求更广阔的人生，这也是我不断奋斗的意义。父亲母亲，作为你们的孩子，我很幸福。

感谢我的姐姐，感谢你从小到大对我的关心、照顾和爱护，你总带给我温暖和快乐。

感谢我的女朋友邵晓雯，我们走过春夏秋冬，未来我们一路携手同行。

感谢江苏省大数据分析技术重点实验室的各位老师和同学，不仅为我提供了很好的科研平台和资源，也让我感受到大家庭的温暖。

感谢母校南京信息工程大学，您教会我做人，将我培养成一个有知识、讲文明的好青年。同时也要感谢自动化学院，感谢各位老师的帮助和支持，祝愿各位老师在南信大工作顺心，也祝愿母校明天更加辉煌。

最后，感谢论文的评审老师在百忙之中对我的论文进行批评和指正。

参考文献

- [1] Sung H, Ferlay J, Siegel R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA: a cancer journal for clinicians, 2021.
- [2] Adenocarcinoma. Available from: www.webmd.com/cancer/what-is-adenocarcinoma.
- [3] Gibson P R, Anderson R P, Mariadason J M, et al. Protective role of the epithelium of the small intestine and colon[J]. Inflammatory bowel diseases, 1996, 2(4): 279-302.
- [4] Humphries A, Wright N A. Colonic crypt organization and tumorigenesis[J]. Nature Reviews Cancer, 2008, 8(6): 415-424.
- [5] Barker N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration[J]. Nature reviews Molecular cell biology, 2014, 15(1): 19-33.
- [6] Washington M K, Berlin J, Branton P, et al. Protocol for the examination of specimens from patients with primary carcinoma of the colon and rectum[J]. Archives of pathology & laboratory medicine, 2009, 133(10): 1539-1551.
- [7] Derwinger K, Kodeda K, Bexe-Lindskog E, et al. Tumour differentiation grade is associated with TNM staging and the risk of node metastasis in colorectal cancer[J]. Acta Oncologica, 2010, 49(1): 57-62.
- [8] Bosman F T, Carneiro F, Hruban R H, et al. WHO classification of tumours of the digestive system[M]. World Health Organization, 2010: 7-10.
- [9] Gleason D F, Mellinger G T. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging[J]. The Journal of urology, 1974, 111(1): 58-64.
- [10] Humphrey P A. Gleason grading and prognostic factors in carcinoma of the prostate[J]. Modern pathology, 2004, 17(3): 292-306.
- [11] 中国人工智能行业综述[J]. 科技中国, 2019(01):63-77.
- [12] Senior A W, Evans R, Jumper J, et al. Improved protein structure prediction using potentials from deep learning[J]. Nature, 2020, 577(7792): 706-710.
- [13] Prewitt J M S. Intelligent microscopes: recent and near-future advances[C]//Recent and Future Developments in Medical Imaging II. International Society for Optics and Photonics, 1979, 206: 206-219.

- [14] 郑众喜. 拥抱数字病理时代[J]. 实用医院临床杂志, 2017, 14(005):6-9.
- [15] Fuchs T J, Buhmann J M. Computational pathology: challenges and promises for tissue analysis[J]. Computerized Medical Imaging and Graphics, 2011, 35(7-8): 515-530.
- [16] Hamilton P W, Bankhead P, Wang Y, et al. Digital pathology and image analysis in tissue biomarker research[J]. Methods, 2014, 70(1): 59-73.
- [17] Bera K, Schalper K A, Rimm D L, et al. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology[J]. Nature reviews Clinical oncology, 2019, 16(11): 703-715.
- [18] Cui M, Zhang D Y. Artificial intelligence and computational pathology[J]. Laboratory Investigation, 2021: 1-11.
- [19] Quay W B. Integument and the environment glandular composition, function, and evolution[J]. American Zoologist, 1972, 12(1): 95-108.
- [20] Pantanowitz L, Valenstein P N, Evans A J, et al. Review of the current state of whole slide imaging in pathology[J]. Journal of pathology informatics, 2011, 2.
- [21] Srinidhi C L, Ciga O, Martel A L. Deep neural network models for computational histopathology: A survey[J]. Medical Image Analysis, 2020: 101813.
- [22] Doyle S, Feldman M, Tomaszewski J, et al. A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies[J]. IEEE transactions on biomedical engineering, 2010, 59(5): 1205-1218.
- [23] Monaco J P, Madabhushi A. Class-specific weighting for Markov random field estimation: Application to medical image segmentation[J]. Medical image analysis, 2012, 16(8): 1477-1489.
- [24] Farooq M T, Shaukat A, Akram U, et al. Automatic gleason grading of prostate cancer using Gabor filter and local binary patterns[C]//2017 40th International Conference on Telecommunications and Signal Processing (TSP). IEEE, 2017: 642-645.
- [25] Li W, Li J, Sarma K V, et al. Path R-CNN for prostate cancer diagnosis and gleason grading of histological images[J]. IEEE transactions on medical imaging, 2018, 38(4): 945-954.
- [26] Arvaniti E, Fricker K S, Moret M, et al. Automated Gleason grading of prostate cancer tissue microarrays via deep learning[J]. Scientific reports, 2018, 8(1): 1-11.

- [27] Nagpal K, Foote D, Liu Y, et al. Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer[J]. NPJ digital medicine, 2019, 2(1): 1-10.
- [28] 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.
- [29] Poojitha U P, Sharma S L. Hybrid unified deep learning network for highly precise Gleason grading of prostate cancer[C]//2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2019: 899-903.
- [30] 龚磊, 徐军, 王冠皓, 等. 基于多特征描述的乳腺癌肿瘤病理自动分级[J]. 计算机应用, 2015, 35(12):3570-3575+3580.
- [31] 何雪英, 韩忠义, 魏本征. 基于深度学习的乳腺癌病理图像自动分类[J]. 计算机工程与应用, 2018, 54(012):121-125.
- [32] Strom P, Kartasalo K, Olsson H, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study[J]. The Lancet Oncology, 2020, 21(2): 222-232.
- [33] Bulten W, Pinckaers H, van Boven H, et al. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study[J]. The Lancet Oncology, 2020, 21(2): 233-241.
- [34] Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases[J]. Journal of pathology informatics, 2016, 7.
- [35] 田娟秀, 刘国才, 谷珊珊, 等. 医学图像分析深度学习研究方法研究与挑战[J]. 自动化学报, 2018, 44(03):401-424.
- [36] Lu C, Koyuncu C, Corredor G, et al. Feature-driven local cell graph (FLoCK): New computational pathology-based descriptors for prognosis of lung cancer and HPV status of oropharyngeal cancers[J]. Medical Image Analysis, 2021, 68: 101903.
- [37] Diao J A, Wang J K, Chui W F, et al. Human-interpretable image features derived from densely mapped cancer pathology slides predict diverse molecular phenotypes[J]. Nature Communications, 2021, 12(1): 1-15.

- [38] Mosquera-Lopez C, Agaian S, Velez-Hoyos A, et al. Computer-aided prostate cancer diagnosis from digitized histopathology: a review on texture-based systems[J]. *IEEE reviews in biomedical engineering*, 2014, 8: 98-113.
- [39] Ali S, Veltri R, Epstein J I, et al. Selective invocation of shape priors for deformable segmentation and morphologic classification of prostate cancer tissue microarrays[J]. *Computerized medical imaging and graphics*, 2015, 41: 3-13.
- [40] Niazi M K K, Yao K, Zynger D L, et al. Visually meaningful histopathological features for automatic grading of prostate cancer[J]. *IEEE journal of biomedical and health informatics*, 2016, 21(4): 1027-1038.
- [41] Zhou Y, Graham S, Alemi Koohbanani N, et al. Cgc-net: Cell graph convolutional network for grading of colorectal cancer histology images[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. 2019: 0-0.
- [42] Xu J, Sparks R, Janowczyk A, et al. High-throughput prostate cancer gland detection, segmentation, and classification from digitized needle core biopsies[C]//*International Workshop on Prostate Cancer Imaging*. Springer, Berlin, Heidelberg, 2010: 77-88.
- [43] Lee G, Sparks R, Ali S, et al. Co-occurring gland angularity in localized subgraphs: predicting biochemical recurrence in intermediate-risk prostate cancer patients[J]. *PloS one*, 2014, 9(5): e97954.
- [44] Awan R, Sirinukunwattana K, Epstein D, et al. Glandular morphometrics for objective grading of colorectal adenocarcinoma histology images[J]. *Scientific reports*, 2017, 7(1): 1-12.
- [45] Edelsbrunner H, Harer J. Persistent homology-a survey[J]. *Contemporary mathematics*, 2008, 453: 257-282.
- [46] Otter N, Porter M A, Tillmann U, et al. A roadmap for the computation of persistent homology[J]. *EPJ Data Science*, 2017, 6: 1-38.
- [47] Zhang Y, Yang Q. A survey on multi-task learning[J]. *arXiv preprint arXiv:1707.08114*, 2017.
- [48] Vandenhende S, Georgoulis S, Proesmans M, et al. Revisiting multi-task learning in the deep learning era[J]. *arXiv preprint arXiv:2004.13379*, 2020.
- [49] Ruifrok A C, Johnston D A. Quantification of histochemical staining by color deconvolution[J]. *Analytical and quantitative cytology and histology*, 2001, 23(4): 291-299.

- [50] Macenko M, Niethammer M, Marron J S, et al. A method for normalizing histology slides for quantitative analysis[C]//2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, 2009: 1107-1110.
- [51] Qaiser T, Sirinukunwattana K, Nakane K, et al. Persistent homology for fast tumor segmentation in whole slide histology images[J]. *Procedia Computer Science*, 2016, 90: 119-124.
- [52] Epstein J I. A new contemporary prostate cancer grading system[C]//*Annales de pathologie*. 2015, 35(6): 474-476.
- [53] Rohatgi V K, Saleh A K M E. An introduction to probability and statistics[M]. John Wiley & Sons, 2015: 101-103.
- [54] Xu J, Xiang L, Liu Q, et al. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images[J]. *IEEE transactions on medical imaging*, 2015, 35(1): 119-130.
- [55] Van Leenders G J L H, Kweldam C F, Hollemans E, et al. Improved prostate cancer biopsy grading by incorporation of invasive cribriform and intraductal carcinoma in the 2014 grade groups[J]. *European urology*, 2020, 77(2): 191-198.
- [56] Xu J, Lu H, Li H, et al. Computerized spermatogenesis staging (CSS) of mouse testis sections via quantitative histomorphological analysis[J]. *Medical Image Analysis*, 2020: 101835.
- [57] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 770-778.
- [58] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017: 2881-2890.
- [59] Sirinukunwattana K, Snead D R J, Rajpoot N M. A stochastic polygons model for glandular structures in colon histology images[J]. *IEEE transactions on medical imaging*, 2015, 34(11): 2366-2378.
- [60] Reinhard E, Ashikhmin M, Gooch B, et al. Color transfer between images[J]. *IEEE Computer Graphics & Applications*, 2002, 21(5):34-41.
- [61] Altaf F, Islam S M S, Akhtar N, et al. Going deep in medical image analysis: concepts, methods, challenges, and future directions[J]. *IEEE Access*, 2019, 7: 99540-99572.

[62] Van der Maaten L, Hinton G. Visualizing data using t-SNE[J]. Journal of machine learning research, 2008, 9(11).

[63] Zhou B, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2921-2929.

作者简介

个人简历

闫朝阳，男，1995年11月生，汉族，江苏连云港人。2014年9月至2018年6月，就读于南京信息工程大学信息与控制学院测控技术与仪器专业；2018年9月至2021年6月，就读于南京信息工程大学自动化学院控制科学与工程专业。硕士期间，所学课程共17门，总学分31分，学位课程学分21分。

硕士期间发表论文情况

- [1] **Chaoyang Yan**, Jun Xu, Jiawei Xie, Chengfei Cai, Haoda Lu. Prior-aware CNN with multi-task learning for colon images analysis[C]//2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE, 2020: 254-257. (Published)
- [2] **Chaoyang Yan**, Kazuaki Nakane, Xiangxue Wang, Yao Fu, Haoda Lu, Xiangshan Fan, Michael D. Feldman, Anant Madabhushi, Jun Xu. Automated Gleason grading on prostate biopsy slides by statistical representations of homology profile[J]. Computer Methods and Programs in Biomedicine, 2020, 194: 105528. (Published)
- [3] Jun Xu, Haoda Lu, Haixin Li, **Chaoyang Yan**, Xiangxue Wang, Min Zang, Dirk G. de Rooij, Anant Madabhushi, Eugene Yujun Xu. Computerized Spermatogenesis Staging (CSS) of Mouse Testis Sections via Quantitative Histomorphological Analysis[J]. Medical Image Analysis, 2021,70: 101835. (Published)
- [4] **Chaoyang Yan**, Jing-Jing Lu, Kang Chen, Lei Wang, Haoda Lu, Mengyan Sun, Jun Xu. Scale- and Slice- aware Net (S²aNet) for 3D segmentation of organs and musculoskeletal structures in pelvic MRI[J]. Magnetic Resonance in Medicine, 2021: 21847. (Minor Revision)

硕士期间竞赛获奖情况

- [1] 2018, “华为杯”第十五届中国研究生数学建模竞赛三等奖（第二完成人）
- [2] 2019, 第三届图像计算与数字医学国际研讨会（ISICDM2019）骨盆多类骨骼肌肉分割竞赛冠军（第一完成人）

硕士期间申请专利情况

- [1] 闫朝阳, 徐军. 基于先验感知和多任务学习的腺癌病理图像分析方法, 发明专利, 公开。
- [2] 闫朝阳, 徐军, 鲁浩达. 基于尺度和序列关系的骨盆核磁共振图像肌肉骨骼分割方法, 发明专利, 公开。
- [3] 徐军, 闫朝阳, 鲁浩达. 前列腺癌穿刺病理图像的格里森自动分级方法和装置, 发明专利, 公开。
- [4] 鲁浩达, 徐军, 闫朝阳. 一种基于深度学习的小鼠睾丸病理切片内多种组织的自动分割方法, 发明专利, 公开。
- [5] 闫朝阳, 邵晓雯. 引导式可计数引体向上单杠, 实用新型专利, 授权。