

doi:10.3969/j.issn.1005-3697.2024.07.003

❖ 基础研究 ❖

# 基于胃组织病理图像数据集的卷积神经网络模型对胃癌的早期预测价值

孙伟, 史航, 黄臻, 法良玲

(青岛市胶州中心医院病理科, 山东 青岛 266300)

**【摘要】目的:** 探究胃组织病理图像数据集的卷积神经网络 (CNN) 模型对胃癌 (GC) 的早期预测价值, 开发并验证 GC 早期预测模型。**方法:** 将 154 例 GC 患者按照分期不同分为早期组 ( $n=87$ ) 和中晚期组 ( $n=67$ )。采用 Logistic 回归分析临床协变量; 使用卷积神经网络 (CNN) 特征提取模型, 搭建 CNN 预测模型; 受试者工作特征 (ROC) 曲线评估区分度, 校准曲线评估准确度。**结果:** 年龄、基础疾病、幽门螺旋菌感染、红细胞计数 (RBC)、白细胞计数 (WBC) 是 GC 的独立危险因素。最佳的 CNN 特征提取模型为 3 个卷积层、2 个池化层和 1 个全连接层。CNN 的各项指标均优于其他模型; 校准曲线分析, CNN 模型的拟合效果显著。**结论:** 基于胃组织病理图像数据集的 CNN 模型具有良好的预测性能, 临床可行性较好。

**【关键词】** 胃癌; 胃组织病理图像; 卷积神经网络模型; 影像组学

**【中图分类号】** R735.2 **【文献标志码】** A

## Early prediction value of convolutional neural network model based on gastric histopathological image dataset for gastric cancer

SUN Wei, SHI Hang, HUANG Zhen, FA Liang-ling

(Department of Pathology, Jiaozhou Central Hospital of Qingdao, Qingdao 266300, Shandong, China)

**【Abstract】Objective:** To explore the early prediction value of convolutional neural network (CNN) model for gastric histological image dataset for gastric cancer (GC), and the development and validation of GC early prediction model. **Methods:** 154 patients with GC were selected and divided into early stage group ( $n=87$ ) and middle stage group ( $n=67$ ) according to different stages. Logistic regression was used to analyze the clinical covariates, and using the CNN feature extraction model, the CNN prediction model was built. The ROC evaluates the degree of differentiation and the accuracy of the calibration curve evaluation. **Results:** Age, underlying disease, helicobacter pylori infection, red blood cell count (RBC) and white blood cell count (WBC) were independent risk factors for GC. The optimal CNN feature extraction model consists of 3 convolution layers, 2 pooling layers and 1 fully connected layer. The index of CNN was better than other models. Calibration curve analysis showed that the fitting effect of CNN model was remarkable. **Conclusion:** The CNN model based on gastric histopathological image dataset has good predictive performance and good clinical feasibility.

**【Key words】** Gastric cancer; Gastric histopathological image; Convolutional neural network model; Imagomics

胃癌 (Gastric cancer, GC) 是第五大常见癌症, 占癌症死亡总数的 18.0%, 预后相对较差, 对全球健康构成严重威胁<sup>[1]</sup>。因此, 早期发现、诊疗对于提高患者的预后极其重要<sup>[2]</sup>。然而, 全国和全球解剖病理学家严重短缺, 造成病理诊断的超负荷劳动, 从而影响了诊断的准确性<sup>[3]</sup>。越来越多的病理学实验室应用组织病理图像进行 GC 的常规诊断<sup>[4]</sup>。图像分类在计算机辅助诊断中起着重要作用<sup>[5]</sup>, 在区分良恶性肿瘤、肿瘤分化阶段、癌症亚型等领域可作为医生临床诊断的重要参考<sup>[6-7]</sup>, 且医学图像技术高精度和高抗干扰, 发展迅速<sup>[8]</sup>。

卷积神经网络 (convolutional neural network, CNN) 已广泛应用于图像识别任务<sup>[9]</sup>。输入层、隐藏层和输出层是 CNN 的典型结构, 原始图像被馈送到输入层, 并进一步传输到隐藏层以执行特征提取任务。隐藏层由卷积层、池化层和全连接层组成。卷积层和池化层中的特征提取是 CNN 模型的主要特征。卷积层可通过使用固定尺寸滤波器的滑动窗口对原始图像进行卷积, 从而自动从原始图像中提取特征。池化层是在保留相关特征的同时降低特征图的维数。全连接层是 CNN 模型的最后一层, 对特征图进行展平以执行分类<sup>[10]</sup>。

作者简介: 孙伟 (1972 -), 男, 副主任技师。E-mail: 13793262066@163.com

通讯作者: 法良玲。E-mail: faliangling879@163.com

目前,探究组织病理图像数据集的 CNN 模型对 GC 的研究在国内尚未见报道。故本研究基于组织病理图像数据集的 CNN 模型,建立对 GC 的早期预测模型,分析比较模型引入前、后模型的拟合优度和准确率变化,有效、准确地预测 GC 病情,以期以低成本提高 GC 的治疗决策。

## 1 资料与方法

### 1.1 一般资料

选取 2020 年 5 月至 2023 年 5 月青岛市胶州中心医院收治的 154 例 GC 患者为研究对象,按照分期不同分为早期组 ( $n = 87$ ) 和中晚期组 ( $n = 67$ )。所有患者中,男性 80 例,女性 74 例,年龄 ( $56.32 \pm 7.20$ ) 岁。纳入标准:(1)符合美国国立综合癌症网络 (National Comprehensive Cancer Network, NCCN) 更新的 2022 版《胃癌临床实践指南》<sup>[11]</sup>,病理证实为 GC;(2)年龄 18 ~ 75 岁;(3)术前未曾接受抗肿瘤治疗;(4)资料完整,病理齐全,图像质量较高。排除标准:(1)术前治疗或术后放疗;(2)非肿瘤相关死亡原因;(3)其他恶性肿瘤;(4)医疗记录不完整;(5)成像质量差。

### 1.2 观察指标

1.2.1 一般资料收集 收集研究对象的年龄、性别、体重、体质量指数 (BMI)、身高、教育情况、居住地等资料。

1.2.2 血液学指标检测 清晨空腹状态下,抽取患者静脉血 5 mL,离心 10 min, 3 500 r/min 离心后取上层血清,当日完成检测。包括患者白细胞计数 (WBC)、血小板分布宽度 (PDW)、红细胞计数 (RBC) 等指标。

### 1.3 特征提取方法

利用 DeepLab v3 架构的 CNN 进行二元图像分割方法,从而实现像素级癌症检测。全视野数字切片 (whole slide image, WSI) 及其相应的注释以  $\times 20$  倍 ( $0.476 \mu\text{m}/\text{像素}$ ) 的放大倍率分成  $320 \times 320$  像素的补丁后,馈送到网络进行训练。在训练期间执行精心设计的数据增强。由于组织病理学载玻片没有特定的方向,故对训练贴片进行 90、180 和 270° 的随机旋转和随机翻转 (水平和垂直)。为了提高从不同医院收集并从各种扫描仪数字化的 WSI 的模型稳定性,在亮度 (0.0 ~ 0.2)、饱和度 (0.0 ~ 0.25)、对比度 (0.0 ~ 0.2) 和色调 (0.0 ~ 0.04) 中应用了高斯和运动模糊以及颜色抖动。在训练过程中,将“质量差”视为“忽略”,并忽略来自“忽略”类的损失。在推理阶段,每个像素都被训练后的模型分配了恶性的概率。通过对所有像素级预测的概率

进行排序来获得幻灯片级预测。

使用 VGG-19、VGG-16、Xception 等最先进的深度学习模型从图像中进行深度特征提取,以提取更好和增强的特征,实现更好的学习和图像分类。CNN 的工作原理采用灰度输入图像 ( $28 \times 28 \times 1$ ) 并执行卷积和池化以进行特征提取,以及使用 ReLU 激活、dropout 和 softmax 激活进行基于神经网络的全连接分类。CNN 作为特征提取器的工作原理,以 RGB 图像 ( $224 \times 224 \times 3$ ) 为输入,并执行 5 次卷积和 5 次池化操作,以增强图像的表达,从而更好地进行特征提取和基于 FCNN 的分类。见图 1 及图 2。

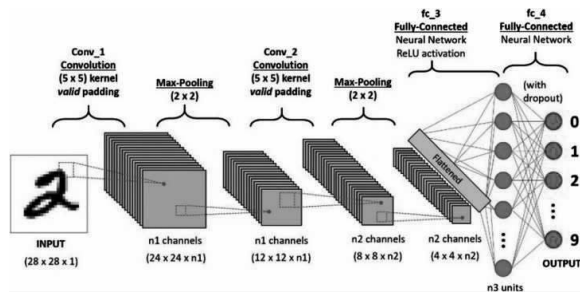


图 1 CNN 工作原理

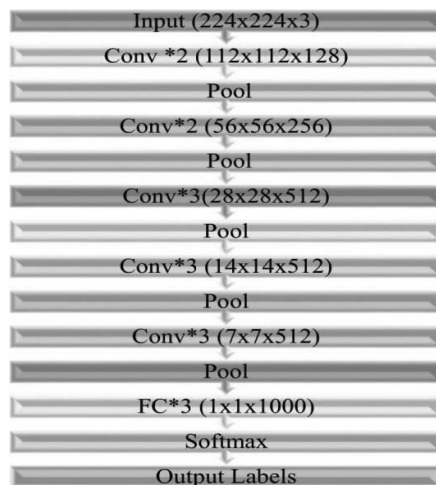


图 2 CNN 特征提取

### 1.4 统计学分析

利用 SPSS 23.0 统计软件对数据进行统计分析。计量资料以  $(\bar{x} \pm s)$  表示,两组间比较行独立样本  $t$  检验,多组间比较行  $F$  检验;计数资料以  $[n (\%)]$  表示,组间比较行  $\chi^2$  检验或 Fisher 确切概率法检验;采用 Logistic 回归分析影响患者的临床危险因素 GC;区分度由受试者工作特征 (ROC) 曲线下面积 (AUC) 进行评估;准确度由校准曲线 Hosmer-Lemeshow 检验进行评估。 $P < 0.05$  为差异有统计学意义。

## 2 结果

### 2.1 早期组和中晚期组患者临床指标比较

训练集共 6 个影响因素,分别为年龄、基础疾病、幽门螺旋菌感染、RBC、WBC、PDW,差异均有统计学意义( $P < 0.05$ )。见表 1。

表 1 早期组和中晚期组训练集患者临床资料的比较 [ $\bar{x} \pm s, n(\%)$ ]

资料	早期组( $n=87$ )	中晚期组( $n=67$ )	$t/\chi^2$ 值	$P$ 值
年龄(岁)	52.36 ± 6.69	57.98 ± 7.13	5.022	<0.001
性别			0.653	0.419
男	46(52.87)	34(50.75)		
女	41(47.13)	33(49.25)		
体重(kg)	63.22 ± 15.73	65.01 ± 15.35	0.707	0.480
BMI(kg/m <sup>2</sup> )	23.11 ± 2.13	23.09 ± 2.22	0.057	0.955
身高(cm)	167.59 ± 5.25	167.61 ± 5.27	0.023	0.981
教育情况			2.450	0.118
高中以下	40(54.02)	33(49.25)		
高中及以上	47(54.02)	34(50.75)		
居住地			0.462	0.497
农村	45(51.72)	36(53.73)		
城市	42(48.28)	31(46.27)		
体温(℃)	36.86 ± 0.21	36.88 ± 0.18	0.623	0.534
呼吸频率(次/min)	18.47 ± 0.88	18.41 ± 0.91	0.413	0.680
心率(次/分)	89.64 ± 9.08	88.58 ± 9.45	0.706	0.482
医疗付费形式			3.122	0.077
医疗保险	38(43.68)	33(49.25)		
自费	49(56.32)	34(50.75)		
基础疾病			4.000	0.046
有	27(31.03)	40		
无	60(68.97)	27		
饮食习惯			0.551	0.458
刺激	46(52.87)	48(71.64)		
清淡	41(47.13)	19(28.36)		
饮酒史			3.122	0.077
有	38(43.68)	33(49.25)		
无	49(56.32)	34(50.75)		
吸烟史			0.174	0.677
有	43(49.06)	48(71.64)		
无	44(50.94)	19(28.36)		
幽门螺旋菌感染			5.000	0.025
是	57(65.52)	50(74.63)		
否	30(34.48)	17(25.37)		
RBC( $\times 10^9/L$ )	4.12 ± 0.35	4.97 ± 0.49	12.552	<0.001
WBC( $\times 10^9/L$ )	6.03 ± 1.14	7.81 ± 1.55	8.212	<0.001
PDW(fl)	11.09 ± 1.12	12.92 ± 2.66	5.789	<0.001

### 2.2 GC 的临床独立因素

将训练集 6 个影响因素纳入 Logistic 回归分析,最终确定年龄( $OR = 3.014, 95\% CI: 2.142 \sim 3.619, P < 0.05$ )、基础疾病( $OR = 1.424, 95\% CI: 1.135 \sim 1.868, P < 0.05$ )、幽门螺旋菌感染( $OR = 2.775,$

$95\% CI: 1.625 \sim 3.728$ )、RBC( $OR = 2.242, 95\% CI: 1.625 \sim 3.728, P < 0.05$ )、WBC( $OR = 3.503, 95\% CI: 2.014 \sim 4.175, P < 0.05$ )是 GC 的独立危险因素。见表 2。

表 2 GC 患者的临床独立因素分析

因素	多因素分析					
	$\beta$ 值	SE 值	Wald 值	OR 值	95% CI	$P$ 值
年龄	1.103	0.512	4.643	3.014	2.142 ~ 3.619	0.012
基础疾病	0.353	0.493	0.514	1.424	1.135 ~ 1.868	0.016
幽门螺旋菌感染	1.021	0.533	3.667	2.775	1.625 ~ 3.728	0.022
RBC	0.807	0.468	2.976	2.242	1.402 ~ 3.335	0.015
WBC	1.254	0.490	6.545	3.503	2.014 ~ 4.175	0.005
PDW	0.182	0.506	0.130	1.200	0.804 ~ 1.673	0.021

### 2.3 卷积特征提取模型

随着迭代次数 epoch 的增加,模型准确率呈现出先提升后逐渐平稳的趋势。当 epoch 从 1 增加至 20 时,模型准确率提升;而当 epoch 从 20 增加至 50 时,模型变化逐渐平稳。在 epoch 为 20 时,模型已经取得了最佳准确率,为 86.37%。同时,损失函数的变化趋势为:当 epoch 在 20 左右时,损失函数趋于稳定,表明模型具有良好的收敛性。因此本研究将 epoch 参数设定为 20。见图 3 及图 4。

卷积层作为 CNN 的核心组成部分,负责提取输入数据的局部特征。卷积层的数量以及卷积核的尺寸对模型的性能具有重要影响。为了深入研究卷积层参数的影响,对比实验分别设置卷积层数为 1、2、3、4,并统一使用 20 的卷积核尺寸进行对比。此外,卷积层在提取 GC 特征方面表现出色,能够全面而有效地提取 GC 的关键特征。见表 3 及表 4。

进一步分析 CNN 特征提取模型网络深度的选择。无法充分捕获关键特征可能由于过浅的网络结构导致,训练难度增加和过拟合问题可能由于网络过深造成。综合实验结果得到最佳的网络结构配置:包含 3 个卷积层、2 个池化层和 1 个全连接层。

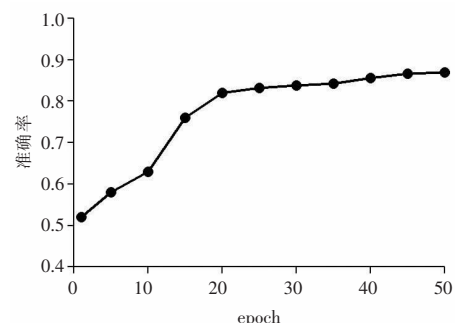


图 3 epoch 与准确率的关系

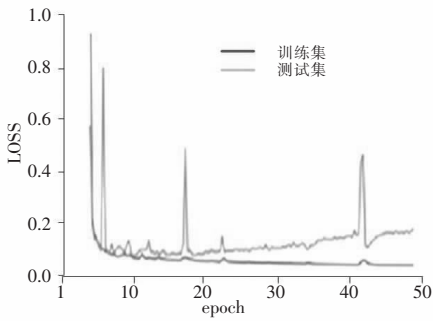


图 4 训练集与测试集损失函数

表 3 卷积层数量与准确率的关系

卷积层数量	卷积核数量	准确率
1	20	0.711
2	20	0.834
3	20	0.816
4	20	0.864

表 4 卷积核尺寸大小与准确率关系

卷积层 1	卷积层 2	卷积层 3	准确率
5 × 5	5 × 5	3 × 3	0.810
5 × 5	5 × 5	5 × 5	0.892
5 × 5	7 × 7	3 × 3	0.882
5 × 5	7 × 7	5 × 5	0.886
7 × 7	3 × 3	3 × 3	0.854
7 × 7	3 × 3	3 × 3	0.868
7 × 7	3 × 3	5 × 5	0.886
7 × 7	5 × 5	3 × 3	0.885

## 2.4 CNN 模型对 GC 的预测结果

相对于其他常用对比方法如 K 近邻 (K-nearest neighbor, KNN)、支持向量机 (support vector machine, SVM)、随机森林 (random forest, RF), CNN 模型具有较高性能, CNN 模型的评价指标提升, 优越性能良好。见表 5。

表 5 不同预测方法的比较

预测方法	准确率	查准率	查全率	AUC 值
RF	0.812 (125/154)	0.813	0.765	0.88
SVM	0.818 (126/154)	0.854	0.744	0.87
KNN	0.883 (136/154)	0.863	0.762	0.90
CNN	0.896 (138/154)	0.867	0.853	0.91

## 2.5 验证模型的 ROC 曲线

在训练集和验证集中, 3 个模型的 AUC 均为 0.7 ~ 0.9, 模型的准确性较高。在训练集中, CNN (AUC = 0.798, 95% CI: 0.734 ~ 0.862) 的敏感度为 81.67%, 特异度为 85.03%, 准确度为 87.37%, 均高于另外 3 种模型。在验证集中, CNN 敏感度为 81.52%, 特异度为 81.91%, 准确度为 86.77%。在评估训练集和验证集时, CNN 模型有更好的判别性

能, 能够更好的预测 GC。经过校准曲线分析, CNN 模型的拟合效果显著, 其预测性能良好。见表 6 及图 5。

表 6 训练集和验证集中三种模型的诊断能力

模型	AUC(95% CI)	敏感度 (%)	特异度 (%)	准确度 (%)
训练集 (n = 154)				
RF	0.785 (0.718 ~ 0.852)	78.54	74.14	75.57
SVM	0.776 (0.708 ~ 0.844)	73.68	76.12	77.01
KNN	0.718 (0.644 ~ 0.792)	81.67	85.03	87.37
CNN	0.798 (0.734 ~ 0.862)	83.57	86.54	88.35
验证集 (n = 94)				
RF	0.750 (0.679 ~ 0.821)	77.09	72.64	74.32
SVM	0.713 (0.634 ~ 0.791)	72.56	72.32	78.61
KNN	0.633 (0.550 ~ 0.716)	81.52	81.91	82.35
CNN	0.785 (0.718 ~ 0.852)	82.35	83.54	86.77

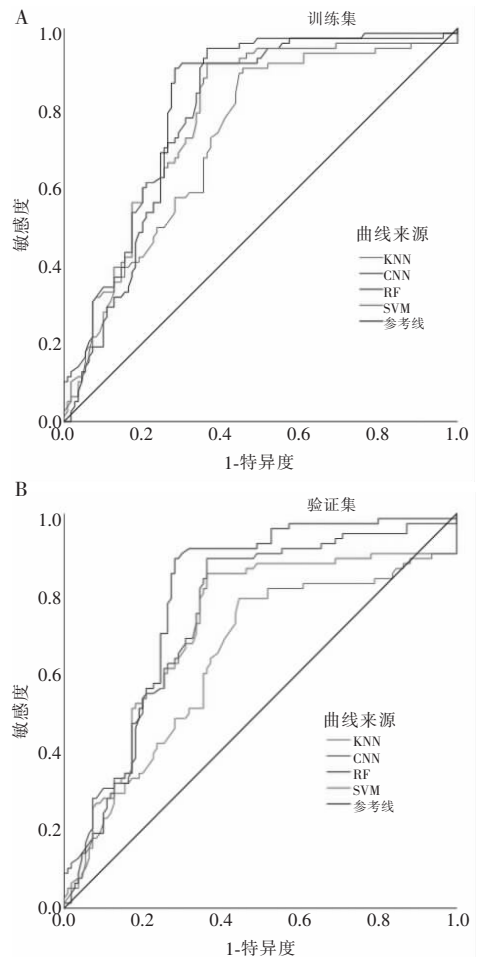


图 5 训练集和验证集中 4 种模型的 ROC 曲线

A. 训练集; B. 验证集。

## 3 讨论

GC 是全球第五大常见癌症, 也是癌症死亡的第三大原因<sup>[12-13]</sup>, 发病率存在很大地域差异, 东亚人群中发病率最高<sup>[14]</sup>。在中国, 2015 年诊断出 498 000 例新发 GC 病例, 是癌症相关死亡的第二大

原因<sup>[15]</sup>。早期发现、准确诊断和手术干预是降低GC死亡率的关键因素,因此强大且有效的病理学检测必不可少。目前,由经验丰富的病理学家对疑似癌性组织制成的石蜡切片进行显微镜检查仍是GC组织病理诊断的金标准,但可能存在以下问题:诊断结果主观性较强且难以定量描述,工作量大,工作时长,易忽略切片信息,无法客观使用大数据<sup>[11]</sup>。从显微镜到图像的转变在病理学中使用人工智能(artificial intelligence, AI)辅助系统克服人类局限性和减少诊断错误提供了契机<sup>[16]</sup>。通过深度学习实现人工智能能够标识潜在可疑区域的算法,从而引导病理学家在高倍率显微镜下对组织样本进行更为详尽的审查,或在必要时进行免疫组织化学(IHC)研究并做出更为准确的诊断<sup>[17]</sup>。

随着计算机视觉技术的飞速发展,特别是医学图像分类的出现,可快速有效地检查每一张电子显微照片,为解决与GC诊断相关的问题提供了机遇<sup>[18]</sup>。CNN模型具有更深层次的架构,可分析图像中比传统神经网络(neural network, NN)更详细的特征<sup>[19]</sup>。本研究指出,年龄、基础疾病、幽门螺旋菌感染、RBC和WBC是GC的独立危险因素。此外,本研究开发并验证的CNN特征提取模型展现出良好的区分度和较高的准确性,将年龄、基础疾病、幽门螺旋菌感染、RBC和WBC参数多学科结合,为预测早期GC提供了一种更精确的联合方法。

综上,基于胃组织病理图像数据集的CNN模型各项指标均优于传统模型,临床可行性较好,具有改善预后和指导GC个性化治疗的潜力。

#### 参考文献

[1] Sung H, Ferlay J, Siegel RL, *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, 71(3): 209–249.

[2] 石小凤,李学锋,周卫华. 胃癌诊断相关血清学标志物的研究进展[J]. *现代医药卫生*, 2023, 39(23): 4084–4089.

[3] Temprana-Salvador J, López-García P, Castellví Vives J, *et al.* DigiPatICS: digital pathology transformation of the Catalan health institute network of 8 hospitals—planning, implementation, and preliminary results[J]. *Diagnostics*, 2022, 12(4): 852.

[4] 吴玮,黄杰,黄宇华,等. 应用计算机辅助诊断技术对口腔鳞状细胞癌组织病理图像进行自动检测的价值[J]. *现代肿瘤医*

学, 2023, 31(3): 459–463.

- [5] 郑雨欣,方梦捷,李海林,等. 影像组学在胃癌诊疗中的应用[J]. *中国体视学与图像分析*, 2022, 27(3): 209–219.
- [6] 王泽朋,李坤鹏,周玉,等. 基于机器学习的胃癌关键基因筛选及预测模型构建[J]. *中国医学物理学杂志*, 2024, 41(1): 115–124.
- [7] 蒋芙蓉,赵静文,刘翔,等. 基于病理图像的计算机辅助诊断进展[J]. *中国医学物理学杂志*, 2022, 39(3): 384–389.
- [8] 徐贵璇,王阳,张杨杨,等. 深度学习在肿瘤组织病理图像分析中的应用[J]. *临床与病理杂志*, 2021, 41(6): 1454–1462.
- [9] 王继仙,桂坤,陈炳宪,等. 基于卷积神经网络的病理活检胃癌诊断模型[J]. *协和医学杂志*, 2022, 13(4): 597–604.
- [10] Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: a review[J]. *Briefings in Bioinformatics*, 2022, 23(2): bbab569.
- [11] Ajani JA, D'Amico TA, Bentrem DJ, *et al.* Gastric cancer, version 2. 2022, NCCN clinical practice guidelines in oncology[J]. *Journal of the National Comprehensive Cancer Network*, 2022, 20(2): 167–192.
- [12] Xuan M, Gu X, Liu Y, *et al.* Intratumoral microorganisms in tumors of the digestive system[J]. *Cell Communication and Signaling*, 2024, 22(1): 69.
- [13] Cheng PL, Loh EW, Chen JT, *et al.* Effects of preoperative oral carbohydrate on postoperative discomfort in patients undergoing elective surgery: a meta-analysis of randomized controlled trials[J]. *Langenbeck's Archives of Surgery*, 2021, 406(4): 993–1005.
- [14] López MJ, Carbajal J, Alfaro AL, *et al.* Characteristics of gastric cancer around the world[J]. *Critical Reviews in Oncology/Hematology*, 2023, 181: 103841.
- [15] Li Q, Yin LK. Comprehensive analysis of disulfidptosis related genes and prognosis of gastric cancer[J]. *World Journal of Clinical Oncology*, 2023, 14(10): 373–399.
- [16] Moscalu M, Moscalu R, Dascălu CG, *et al.* Histopathological images analysis and predictive modeling implemented in digital pathology—current affairs and perspectives[J]. *Diagnostics*, 2023, 13(14): 2379.
- [17] Valente J, António J, Mora C, *et al.* Developments in image processing using deep learning and reinforcement learning[J]. *Journal of Imaging*, 2023, 9(10): 207.
- [18] Zhang J, Hua Z, Yan K, *et al.* Joint fully convolutional and graph convolutional networks for weakly-supervised segmentation of pathology images[J]. *Medical Image Analysis*, 2021, 73: 102183.
- [19] Bajić F, Orel O, Habijan M. A multi-purpose shallow convolutional neural network for chart images[J]. *Sensors (Basel, Switzerland)*, 2022, 22(20): 7695.

(收稿日期: 2024-01-13

修回日期: 2024-04-21)