

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2024)10-3157-14

论文引用格式: Ding W L, Liao W Y, Zhu W, Wang C N, Zhu X Q and Zhu H B. 2024. Artifact detection of low-magnification pathology whole-slide images. *Journal of Image and Graphics*, 29(10):3157-3170(丁维龙, 廖婉茵, 朱伟, 汪春年, 祝行琴, 朱红波. 2024. 低倍率病理全切片图像伪影检测. *中国图象图形学报*, 29(10):3157-3170)[DOI:10.11834/jig.230647]

低倍率病理全切片图像伪影检测

丁维龙¹, 廖婉茵¹, 朱伟¹, 汪春年^{2*}, 祝行琴², 朱红波³

1. 浙江工业大学计算机科学技术学院, 杭州 310023; 2. 宁波市临床病理诊断中心, 宁波 315021;

3. 上海市浦东医院暨复旦大学附属浦东医院病理科, 上海 201399

摘要: 目的 高质量的病理切片对人工诊断和计算机辅助诊断至关重要。当前基于图像块的伪影检测方法存在着计算资源消耗巨大以及伪影检测过程的完整性缺失问题。为此, 本文提出了一种适用于低倍率病理全切片图像的伪影检测算法 WRC_Net(window-row-col_net)。方法 首先, 将低倍率的全切片图像输入到 ResNet50(residual neural network)网络中, 以提取图像的低级特征。随后, 这些低级特征被传入特征融合模块, 用于聚合来自不同深度和方向的特征。此外, 在特征提取模块中, 引入了 WRC 模块, 包括 WRC 注意力和多尺度扩张模块, 其能够同时捕捉全局和局部信息, 提取多尺度特征, 从而增强了特征的表达能力。最后, 将融合后的特征传入单一检测头, 以获取最终的检测结果。结果 在 SPDPSD(Shanghai Pudong department of pathology slide dataset)和 NCPDCSD(Ningbo clinical pathology diagnosis center slide dataset)两个数据集上, 所提方法的平均精度(mean average precision, mAP)分别达到了 63.1% 和 55.0%, 与目前主流的目标检测算法相比具有一定竞争力。结论 本文提出的病理切片伪影检测算法能够准确识别数字病理切片中的不同种类伪影, 为病理图像质量评估提供了一种有效的技术解决方案。

关键词: 数字病理学; 数字病理切片; 伪影检测; 多尺度; 特征融合

Artifact detection of low-magnification pathology whole-slide images

Ding Weilong¹, Liao Wanyin¹, Zhu Wei¹, Wang Chunnian^{2*}, Zhu Xingqin², Zhu Hongbo³

1. College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China;

2. Ningbo Diagnostic Pathology Center, Ningbo 315021, China;

3. Department of Pathology, Shanghai Pudong Hospital, Fudan University Affiliated Pudong Medical Center, Shanghai 201399, China

Abstract: Objective High-quality pathological slides are crucial for manual diagnosis and computer-aided diagnosis. However, pathology slides may contain artifacts that can affect their quality and consequently influence expert diagnostic judgment. Currently, the assessment of pathology slide quality often depends on manual sampling, which is time consuming and cannot encompass all slides. Additionally, despite relatively standardized detection criteria across different institutions, different quality control personnel may have varying interpretations, leading to subjective differences. These limitations restrict the traditional quality control process. Current methods typically involve analyzing image patches cropped at high magnifications, resulting in considerable consumption of computational resources. Moreover, certain artifacts, such as hole and tremor, often have larger dimensions and are better suited for learning on low-resolution images. However,

收稿日期: 2023-09-28; 修回日期: 2023-12-25; 预印本日期: 2024-01-02

* 通信作者: 汪春年 niana217@126.com

基金项目: 国家自然科学基金项目(32271983); 浙江省基础公益研究计划项目(TGY24F020014, LTGY23F020005)

Supported by: National Natural Science Foundation of China (32271983); Zhejiang Public Welfare Technology Research Plan Project (TGY24F020014, LTGY23F020005)

cropping image blocks at high magnifications may exclude artifacts that coexist in tissue and background regions, such as ink and bubble, which compromise the integrity of artifact detection and analysis. To tackle these challenges, the process of assessing the quality of pathology slides must be digitized to enhance efficiency and accuracy. Therefore, this study introduces a novel algorithm, Window-Row-Col_Net (WRC_Net), designed for the detection of artifacts in digital pathology slides, specifically tailored for low-magnification pathological whole slide images. **Method** This study primarily consists of four components: pathological slide preprocessing, feature extraction module, feature fusion module, and single detection head. First, we tackle the problem of artifacts in pathological slides by performing preprocessing, which involves transforming the tissue pathological slides into lower-pixel-resolution versions of whole-slide images (thumbnails). This step helps reduce computational resource consumption and is well-suited for addressing larger-sized artifacts, such as holes and wrinkles. Afterward, these thumbnails are input into the feature extraction network to acquire the low-level feature representation of the slides. Moreover, to integrate effectively feature information from different levels, we devised a feature fusion module. This module plays a crucial role in the entire model, facilitating the aggregation of features from varying depths and directions. Additionally, we introduce the Window-Row-Col (WRC) module, which encompasses two vital components: the WRC attention module and the multiscale dilated module (MSDM). The WRC attention module dynamically aggregates features within square windows, spanning horizontal and vertical directions, capturing global and local contextual information, and establishing long-range dependencies. The MSDM incorporates dilated convolutions with varying dilation rates to extract multiscale information and enhance the model's perception of artifacts of different sizes. Through this feature fusion, we boost the expressive capability of features, thus giving the model with a competitive advantage in object detection tasks. Finally, the fused features are directed into a single detection head to produce the ultimate detection outcomes. The design of a single detection head structure streamlines the model's architecture, rendering it more succinct and efficient and also diminishing storage and computational burdens, thereby expediting the detection process. **Result** We created two new datasets for pathology slide artifacts, known as the Shanghai Pudong Department of Pathology slide dataset (SPDPSD) and the Ningbo Clinical Pathology Diagnosis Center slide dataset (NCPDCSD). The SPDPSD dataset originates from the Pathology Department of Pudong Hospital in Shanghai and comprises two types of tissue artifacts: ink and bubble. The NCPDCSD dataset is obtained from the Clinical Pathological Diagnosis Center in Ningbo and encompasses six types of tissue artifacts, including hole, tremor, incompleteness, ink, appearance, and bubble. Through comprehensive experiments on these two datasets, we conducted an extensive comparison between our proposed WRC_Net algorithm and the current state-of-the-art object detection methods. The WRC_Net algorithm demonstrated superior detection performance on the SPDPSD dataset, yielding remarkable improvements in metrics, such as mean average precision (mAP), $\text{mAP@IoU} = 0.75$ (mAP75), small object mAP (mAPs), medium object mAP (mAPm), and large object mAP (mAPl). Additionally, on the NCPDCSD dataset, our method displayed competitive performance. The results indicate that compared with other methods, WRC_Net achieved remarkable enhancements in multiple performance metrics while maintaining low computational complexity, striking a balance between speed and accuracy. **Conclusion** The proposed algorithm for detecting pathology slide artifacts accurately identifies various types of artifacts in digital pathology slides, offering a more comprehensive and dependable quality assessment tool for the domain of medical image analysis. By striking a balance between efficiency and accuracy, our algorithm is anticipated to enhance the pathology diagnosis process and usher in a more efficient and precise workflow within pathology departments.

Key words: digital pathology; whole slide images; artifact detection; multi-scale; feature fusion

0 引言

高质量的病理切片应能满足病理医生的阅片需求,并能清晰地从切片中分辨出细胞有丝分裂、细胞凋亡和细胞核碎片等细节(Deng等,2020)。然而,

病理切片制片过程中的任何不当操作都可能导致组织结构的变化(本文称之为伪影),从而影响切片的质量(Woods和Ellis,1994)。伪影可以分为以下3类:1)组织级伪影,如颤痕、撕裂;2)切片级伪影,如气泡、墨水;3)扫描级伪影,如失焦(Kanwal等,2022)。这些伪影的存在可能掩盖或误导病理医生

对真实病变的观察和判断,导致误诊或漏诊,从而影响病患的治疗和预后。此外,不同的医生对伪影的感知和解释可能不同,导致对同一病例的诊断结果存在差异,降低了诊断的可靠性。因此,对制作完成的病理切片进行质量检测至关重要。

目前病理切片质量检测主要依赖于两种方法:人工检测和计算机辅助诊断(computer aided diagnosis, CAD)。病理切片质量评估的人工检测通常需要质检人员仔细检查切片厚度、清晰度以及其他可能影响诊断结果的因素。采用人工质检方式存在若干问题:1)人工抽检无法覆盖所有切片,导致遗漏和漏检;2)质检过程耗时耗力,且不同质检人员间存在主观差异性;3)不同的医学机构、实验室以及研究项目可能有不同的需求和标准,导致切片质量控制通常没有统一的检测标准。这些问题使得传统的质量控制流程存在较大的局限性。

针对上述问题,国内外学者提出了多种计算机辅助诊断的方法来自动筛查病理切片质量。早期的病理切片伪影检测通常采用传统机器学习算法。Bautista 和 Yagi(2010)提出了一种具有固定阈值的基于 RGB 颜色的方法来检测组织褶皱。Kothari 等人(2013)提出利用组织结构的颜色和组织基元的连通性,并采用自适应阈值方法来识别组织折叠区域。一般地,不同医院或同一医院不同批次制作的病理切片,其染色过程可能存在差异,导致图像的颜色特征有所不同。上述方法可能难以处理这种差异性,无法达到预期效果。

基于深度学习的方法能够自动从原始数据中学习 to 更高级别的抽象特征表示,无需手动设计特征(赵樱莉等,2023)。诸多学者已将卷积神经网络(convolution neural network, CNN)应用于病理切片伪影识别。Foucart 等人(2018)采用基于像素的分类方法,在数据有限的情境下开发了伪影区域分割模型,成功实现了数字病理图像中伪影组织区域的分类和分割,模型表现良好且在无监督信息下效果显著。Hossain 等人(2023)利用分割算法成功应用于组织折叠和气泡两种伪影的分割。Maleki 等人(2020)通过不同的骨干网络实现了病理切片中墨迹区域的分割,并通过生成的掩码有效消除墨迹区域,保持较高准确率,但需要大量标注数据。Hemmatirad 等人(2022)运用 YOLOv4 (you only look once)(Bochkovskiy 等,2020)模型检测载玻片制备中可能

出现的边缘、组织褶皱以及墨水标记的组织玻片。Babaie 和 Tizhoosh(2019)通过比较不同深度的预训练卷积神经网络作为特征提取器,成功区分褶皱组织与正常组织。

以上这些方法在高倍率下使用滑动窗口或随机窗口来处理全切片图像,通常将其裁剪成多个尺寸为 256×256 或 512×512 像素的图像块。然而,对全切片进行裁剪,再分别对每个小图像块进行性质判断,然后把这些结果整合起来合成整个切片的性质的判定,将会导致大量的计算资源和内存消耗。此外,在高倍率下裁剪图像块,还可能将某些伪影目标排除在裁剪范围之外,使得伪影检测和分析过程失去完整性。比如,墨水和气泡可能同时存在于组织和背景区域;而空洞、组织不完整等伪影与背景区域外形相似。

病理切片具有金字塔结构(Tong 等,2019),其中高倍率层级图像呈现更多细节,适用于较小目标,如组织褶皱等伪影,但其文件较大且需要更多计算资源。相比之下,低倍率层级的图像提供更广阔的视野,适用于较大目标,如墨水等伪影,其文件较小,节省了存储和传输资源。在病理切片中,存在一些伪影目标,如空洞、颤痕、组织不完整、墨水、外观以及气泡等,其几何尺寸较大,现有方法基于高倍率进行,其缺点是显而易见的。由于高分辨率图像尺寸较大,无法直接送入模型进行训练,需将其裁剪成多个小图像块。而这些伪影尺寸较大,裁剪时可能导致单个伪影被分割成多个图像块,从而引发伪影全局性缺失问题。此外,低倍率图像通常包含较少的像素和信息,这有助于提升算法的计算效率。事实上,这些伪影更适合在低倍率图像上进行学习。

然而,将病理切片从高倍率转换为低倍率时,图像细节信息被压缩,切片中伪影的细节无法清晰地展现。同时,低倍率图像中的伪影通常较小且形状多变,难以区分和检测。可见,低倍率图像中的伪影特点与高倍率图像不同,使得现有目标检测方法在低倍率图像上性能不佳。

为了解决上述问题,本文提出了一个专门针对低倍率图像的病理切片缺陷伪影检测算法 WRC_Net(window-row-col_net)。通过在低倍率图像(即缩略图)上进行学习,本文方法能够全面检测和定位病理切片中的缺陷伪影目标,包括在高倍率下裁剪图像块时容易被忽略的缺陷伪影。具体来说,

首先提取低倍率图像的低级特征,然后引入特征融合模块,以聚合不同方向和深度的特征信息。这有助于增强模型对不同大小和形状的伪影的感知能力。WRC_Net采用单一检测头的设计,简化了模型结构,使其更适应低倍率图像中伪影的特点,从而实现高效准确地检测不同大小和形状的伪影。

另外,经文献检索发现,目前国内外关于伪影的数据集非常有限。现有研究中,除了使用自制数据集(Hemmatirad等,2022),研究人员还从癌症基因组谱(the cancer genome atlas program, TCGA)中挑选含伪影的病理切片,并对其进行标注以构建相应的伪影数据集(Maleki等,2020;Shakarami等,2023)。然而,上述的自制数据集均未公开。当前仅有的针对伪影检测的公开数据集 HistoQCRepo(Janowczyk等,2019),其数据量有限,覆盖的伪影种类不够全面,且不含标注信息。若直接在该数据集上进行检测,需要对其进行标注,且数据量不足,可能无法取得良好的效果。为解决这一问题,本文构建了两个新的组织病理切片伪影数据集。

本文的主要工作包含两个部分:1)构建了两个病理切片伪影数据集,即 SPDPSD(Shanghai Pudong department of pathology slide dataset)和 NCPDCSD(Ningbo clinical pathology diagnosis center slide dataset)。前者来源于上海市浦东医院病理科,包含墨水和气泡两种伪影。后者来源于宁波市临床病理诊断中心,包括空洞、颤痕、组织不完整、墨水、外观和气泡6种伪影。2)提出一种新的数字病理切片伪影

检测算法 WRC_Net。该算法能在低倍率下自动检测病理切片中的空洞、颤痕、组织不完整、外观、气泡和墨水等伪影,并给出相应的预测结果。在 SPDPSD 数据集和 NCPDCSD 数据集上的实验结果表明,本文方法实现了精度和速度的平衡,可为后续病理切片伪影检测研究提供技术参考。

1 私有数据集的建立

本文构建的 SPDPSD 数据集来源于上海市浦东医院病理科。它包含 363 幅全切片图像,其中涵盖了墨水和气泡两种伪影,如图 1 所示。

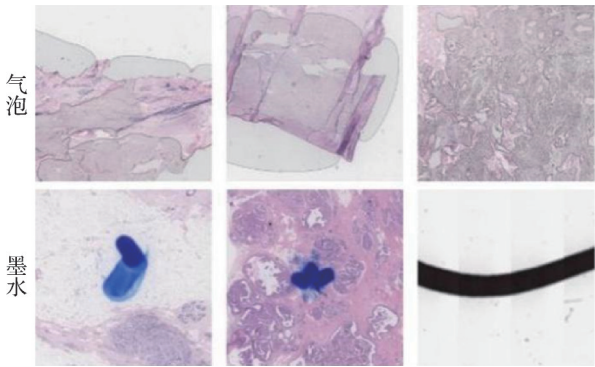


图1 SPDPSD数据集

Fig. 1 SPDPSD dataset

NCPDCSD数据集来自宁波市临床病理诊断中心,包含 526 幅全切片图像,涵盖了空洞、颤痕、组织不完整、墨水、外观和气泡等 6 种伪影,如图 2 所示。

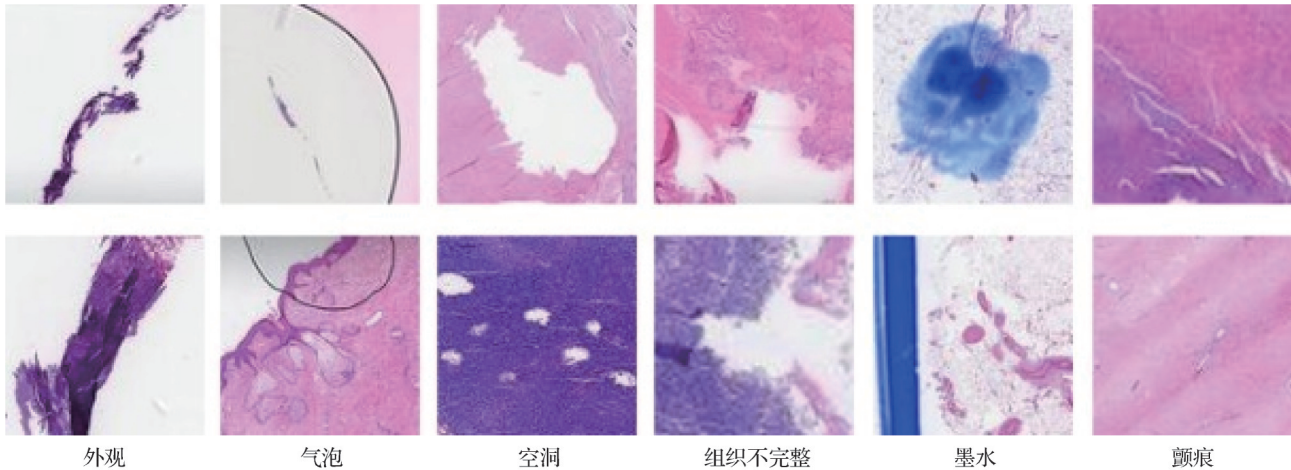


图2 NCPDCSD数据集

Fig. 2 NCPDCSD dataset

数据集制作包括采集、扫描和标注这3个步骤。以NCPDCSD数据集的制作为例,首先,从宁波市临床病理诊断中心的《病理常规制片每日总质控表》中挑选经过质检的切片,并将它们整合到《缺陷表格》中。接下来,将这些信息添加到《病理档案资料借阅单》中,并提交借阅申请。经借阅科室负责人和病理档案科负责人审核后,协同工作人员到病理切片档案室中挑选目标切片。其次,填写《数字病理扫描申请表》,经医教科负责人和病理档案科负责人审批后,由病理档案科经办人使用显微扫描仪,以20倍的放大倍率将切片扫描成病理数字切片。最后,采用“2+1”的标注策略,成立切片缺陷标注小组,由1名组长(科室主任)和2名组员(科室工作人员)组成。在标注过程中,每位小组成员会独立对每张切片进行标注。随后,标注结果会经过组长的审核,以确保标注质量的准确性和一致性。

本文将所有全切片图像转换成尺寸为 $1\,024 \times 1\,024$ 像素的缩略图。具体操作如下:对于构建的SPDPSD和NCPDCSD数据集,其图像数据均源自全切片图像的原始放大倍率,即20倍率。本文从每个全切片图像中提取出尺寸为 $1\,024 \times 1\,024$ 像素的缩略图图像。经实验证明,在低倍率下,气泡、墨水等伪影仍有较好的识别精度。构建伪影数据集需要大量的切片和标注信息。然而,现实生活中,采集大量病理切片并进行手工标注是困难且耗时的。为了解决伪影数据集规模有限的问题,采用了类似CutMix(Yun等,2019)的数据增强技术来扩充数据集。具体做法是将伪影目标框裁剪出来,然后经随机缩放操作后填充至其余图像的任意区域中。经数据预处理和数据增强操作后,SPDPSD数据集包含6 190幅缩略图。NCPDCSD数据集包含3 830幅缩略图。

2 方 法

本文充分考虑低倍率图像的特点,提出一个针对低倍率图像的数字病理切片伪影检测算法WRC_Net。为了克服低倍率图像细节信息的丢失,首先使用特征提取模块从低倍率图像中提取低级图像特征。同时,考虑到低倍率图像中伪影较小且形状多变的问题,引入特征融合模块。该模块能聚合不同方向和深度的特征信息,从而增强模型对不同大小和形状伪影的感知能力。此外,WRC_Net采用

单检测头的设计,简化了模型结构和参数量。通过这些改进,WRC_Net能够更好地适应低倍率图像中伪影的特点,实现高效准确地检测不同大小和形状的伪影,从而为病理科室提供了有效的伪影检测解决方案。WRC_Net的整体结构如图3所示,“Win_Row_Col Module”表示WRC模块,其结构如图4所示。为清晰表达图4中的多尺度扩张模块(multi-scale dilated module, MSDM)的结构,单独绘制了图5。

2.1 特征提取模块

对于每张全切片缩略图,采用在ImageNet上预训练的ResNet50(He等,2016)模型来提取包含局部细节信息的低级特征,如边缘和纹理。ResNet50是一种流行的深度学习模型,广泛应用于分类、检测和分割等任务中。它通过残差连接解决了梯度消失和梯度爆炸问题,这使得其在学习复杂特征表示方面具有出色的性能。为了降低计算成本和加快训练时间,只使用ResNet50模型的第3层输出作为特征表示,这样能捕捉到图像的局部和全局信息。同时,相较于完整的ResNet50模型,只使用第3层输出可以减少模型的参数量和计算复杂度。

2.2 特征融合模块

由于病理切片上伪影的尺寸和形状各异,单一的检测头往往难以全面捕捉所有尺度的目标。为了克服这一挑战,本文引入了一个WRC模块,以增强模型对于不同尺寸和形状的伪影目标的感知能力,并为后续的伪影检测任务提供更全面、准确的特征表示。在2.2.1节中,本文将详细介绍该模块。而在目标检测任务中,不同尺度和层级的特征对于检测不同大小的目标都是重要的。受此启发,本文提出了基于WRC模块的特征融合模块。该模块通过在通道维度上拼接不同WRC层的特征,有效地提升了模型对不同大小目标的感知能力。此外,为了进一步区分不同通道特征的重要性,本文还引入了ECA(efficient channel attention)(Wang等,2020)模块。该模块能够自适应地调整通道权重,增强特征的表达能力。

2.2.1 WRC模块

如图3所示,作为特征提取模块的核心,WRC模块由LN(layer norm)层、多尺度扩张模块以及WRC注意力(Win-Row-Col attention)3部分组成。MSDM负责提取多尺度信息,并引入上下文信息,以辅助检

为特征图的宽和高。首先,将特征沿通道维度进行拆分,得到方形特征($f_w \in \mathbf{R}^{B \times C \times H \times W}$)、水平特征($f_r \in \mathbf{R}^{B \times C \times 1 \times W}$)和垂直特征($f_c \in \mathbf{R}^{B \times C \times H \times 1}$)。其中,水平特征和垂直特征的提取是通过水平和垂直注意力机制(Ho等,2019)实现的,前者用于将特征的每一行聚合到目标特征的每一行上,以捕获行之间的关系;后者则用于捕获特征的列之间的关系。引入水平和垂直注意力有助于模型在处理二维矩阵数据时捕获行或列之间的关系,从而提升对图像内容的理解和处理能力。通过使用局部的窗口注意力来捕获图像特征的局部注意力,可以减少全局上下文的干扰,使模型可以更专注于局部区域的细节。为了更好地融合这3种不同的注意力,参考CEFF(channel equalization and feature fusion)模块(Noman等,2023),对3种注意力重新加权,并进行特征融合。

首先,将特征 f_c 、 f_w 和 f_r 相加,然后进行全局平均池化(global average pooling, GAP)得到全局平均特征 f_{gap} ,具体为

$$f_{gap} = \text{GAP}(f_c \oplus f_w \oplus f_r) \quad (1)$$

式中,GAP为全局平均池化, $\oplus$ 为按元素相加。将特征 f_{gap} 输入到ConvL层,以减少通道数。之后,再将特征分别输入3个ConvS层,以分别获得3种注意力的权重,具体为

$$\begin{aligned} f'_{gap} &= \text{LeakyReLU}(\text{Conv}^{1 \times 1}(f_{gap})) \\ w_c &= \text{softmax}(\text{Conv}^{1 \times 1}(f'_{gap})) \\ w_w &= \text{softmax}(\text{Conv}^{1 \times 1}(f'_{gap})) \\ w_r &= \text{softmax}(\text{Conv}^{1 \times 1}(f'_{gap})) \end{aligned} \quad (2)$$

式中, w_c 、 w_w 、 w_r 为卷积权重,LeakyReLU和softmax为激活函数, $\text{Conv}^{1 \times 1}$ 为 1×1 卷积。最后,分别使用权重 w_c 、 w_w 、 w_r 对特征 f_c 、 f_w 、 f_r 进行通道重新加权,然后相加进行特征融合,得到增强后的特征 f_e ,如下所示:

$$f_e = w_c \otimes f_c \oplus w_w \otimes f_w \oplus w_r \otimes f_r \quad (3)$$

式中, $\otimes$ 为按元素相乘。这样可以增强与语义变化相关的特征,并抑制噪声的影响。通过同时使用局部特征、水平特征和垂直特征,WRC模块能够更全面地捕捉图像中的全局和局部信息,提高模型对复杂形态的伪影的检测能力。

2.2.3 多尺度扩张模块

由于伪影的尺寸变化较大,可能出现极大或极小等极端情况,这给伪影检测带来了极大的困难。针对此问题,本文在WRC模块中引入了多尺度结构MSDM,旨在捕获更多的上下文信息。如图5所示,MSDM的核心思想是使用不同扩张率的膨胀卷积操作来提取多尺度信息。MSDM由一个膨胀卷积分支和一个SEBlock(Hu等,2018)分支组成。膨胀卷积分支由4个 3×3 卷积组成,每个卷积具有不同扩张率($r = \{1, 2, 3, 4\}$)。为了降低参数量,本文在膨胀卷积分支的前后分别添加了 1×1 卷积层,实现降维和升维的作用。膨胀卷积在保持参数量不变的情况下增大卷积的感受野,提取多尺度信息,进一步提升不同尺度目标的检测性能。SEBlock分支的输出权重通过乘法逐通道加权到膨胀卷积分支的输出特征上,实现在通道维度上对特征进行重加权,增强重要特征的表示能力。

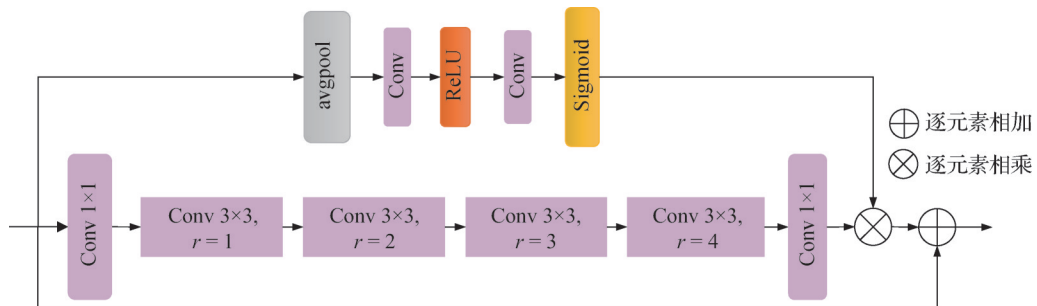


图5 MSDM模块

Fig. 5 MSDM module

2.3 单检测头

多检测头方法在同一模型中使用多个独立的“头部”来处理不同的子任务,然后将它们的输出进

行整合以生成最终的结果。与多检测头的方法不同,单检测头的方法将目标检测任务视为一个回归问题,即通过一个单一的检测头来同时预测目标的

位置和类别。相较于多检测头,单检测头能够简化模型结构和参数量,减少存储和计算开销。其次,采用单检测头的设计还降低了模型的复杂度,模型只需学习一个检测头便可完成整个目标检测任务。因此,出于这些考虑,本文采用单检测头的设计,以简化模型、提高效率、减少计算开销并降低模型复杂度,从而提供了一种更为合适的伪影检测方案。

YOLOX(Ge等,2021)中的解耦头结构考虑到目标检测任务中分类和定位的关注点不同,采用不同的分支来分别对它们进行计算,改善了检测效果。如图3所示,解耦头将任务分解为3个部分:1) $\text{Reg}(h, w, 4)$ 分支,用于计算每个特征点的回归参数,通过调整这些参数,可以获得预测框的位置和尺寸,其中, h 和 w 分别为特征图的高和宽;2) $\text{Obj}(h, w, 1)$ 分支,用于判断每个特征点是否包含目标;3) $\text{Cls}(h, w, c)$ 分支,用于判断每个特征点包含的目标种类,其中 c 为目标的类别数。解耦头将3个预测结果分别计算,然后将它们堆叠在一起,输出 $\text{out}(h, w, 4 + 1 + c)$ 。这种结构的设计在保持高性能的同时,提高了计算效率。因此,本文采用解耦头作为检测头,以更好地考虑到伪影检测中的分类和定位需求的差异,从而在任务中取得更好的综合表现。

2.4 损失函数

目标检测损失函数用于衡量模型在预测目标和真实目标之间的差异。它由3个组成部分组成:交并比损失(iou_loss)、目标置信度损失(obj_loss)和目标类别损失(cls_loss)。 iou_loss 用于评估预测框与实际框之间的重叠度。 obj_loss 用于评估预测目标的置信度。 cls_loss 则用于评估预测目标的类别(Ge等,2021)。目标检测损失函数为

$$\begin{aligned} L_{\text{det}} &= \alpha L_{\text{iou}} + \beta L_{\text{obj}} + \gamma L_{\text{cls}} \\ L_{\text{iou}} &= \sum_{i=0}^{S^2} \sum_{j=0}^B \left[C_i^j \left((x_i^j - \hat{x}_i^j)^2 + (w_i^j - \hat{w}_i^j)^2 + (h_i^j - \hat{h}_i^j)^2 \right) \right] \\ L_{\text{obj}} &= - \sum_{i=0}^{S^2} \left[C_i^j \times \log(\hat{C}_i^j) + (1 - C_i^j) \times \log(1 - \hat{C}_i^j) \right] \\ L_{\text{cls}} &= - \sum_{i=0}^{S^2} \sum_{j=0}^B \sum_{c=0}^C \left[C_i^j \times \log(\hat{p}_i^j(c)) \right] \end{aligned} \quad (4)$$

式中, L_{iou} 表示 iou_loss , L_{obj} 表示 obj_loss , L_{cls} 表示 cls_loss 。 α, β, γ 分别为3种损失的权重系数。 S 是网格单元的数量, B 是每个网格单元预测的边界框

数量, C_i^j 表示第 i 个网格单元中的第 j 个边界框是否包含物体的真实标签(0或1), $\hat{C}_i^j$ 表示模型的预测概率,表示第 i 个网格单元中的第 j 个边界框是否包含物体。 x_i^j 和 y_i^j 表示第 i 个网格单元中的第 j 个边界框的中心坐标, w_i^j 和 h_i^j 表示第 i 个网格单元中的第 j 个边界框的宽度和高度, $\hat{x}_i^j, \hat{y}_i^j, \hat{w}_i^j, \hat{h}_i^j$ 表示模型的预测值。 C 是类别的数量, $\hat{p}_i^j(c)$ 表示第 i 个网格单元中的第 j 个边界框预测为第 c 个类别的概率。与本工作类似, FastestDet(Ma, 2022)也是基于 anchor-free 的目标检测方法,且仅含有单个轻量级检测头。因此,本文在设计目标检测损失函数权重时,选择了借鉴 FastestDet 的权重设置,其中 $\alpha = 8, \beta = 16, \gamma = 1$ 。

3 实验及结果分析

3.1 实验设置

实验硬件环境为 Intel i9-10940x CPU, 两张 NVIDIA 3090 显卡,操作系统为 Linux。本文设计的网络模型基于 PyTorch 深度学习框架实现。使用 SGD 优化器和 MultiStepLR 学习率衰减策略。其中优化器的权值衰减为 0.0005, 动量系数(momentum)设为 0.949。模型的数据批次大小为 32, 初始学习速率为 0.001, 并进行 200 个 epoch 的训练。文中使用的评价指标主要包括 mAP、mAP05、mAP75、mAPs、mAPm 和 mAPl。

平均精度(mean average precision, mAP)(Everingham等, 2010)是目标检测模型中常用的评价指标,它综合考虑了所有类别下不同召回率水平下的平均精确率,能够全面评估目标检测模型在不同类别上的性能。

mAP05 和 mAP75 分别表示在 IoU 阈值为 0.5 和 0.75 时的平均精确度。而 mAPs、mAPm 和 mAPl 则分别表示在小目标、中等目标和大目标上的平均精确度。通过这些衍生指标,可以更细致地了解模型在不同情况下的性能表现。

3.2 实验结果

3.2.1 对比实验

为了验证 WRC_Net 的有效性,针对 SPDPSD 数据集和 NCPDCSD 数据集,将 WRC_Net 与近几年的目标检测方法进行了对比。为保持一致性,本文将所有模型的输入尺寸设置为 352×352 像

素,并且所有模型在训练过程中均不使用预训练权重。

在 SPDPSD 数据集上,将提出的 WRD_Net 与主流的目标检测算法进行比较。FastestDet 是一种实时 anchor-free 目标检测算法,其在速度和参数量上优于其他算法,但精度相对较低。YOLOv5(Ultralytics, 2020)、YOLOv7(Wang 等, 2023)以及 YOLOX(Ge 等, 2021)为 YOLO 系列中的不同版本。这些算法在目标检测领域有着显著的影响。CenterNet2(Zhou 等, 2021)是基于单阶段检测器构造的概率性两阶段检测器。相较于单阶段检测器,它具有更高的速度和准确率。另外, CornerNet-Lite(Law 等, 2020)是在 CornerNet(Law 和 Deng, 2018)的基础上进行改进和优化的轻量级版本,其目标是在保持高检测性能的

同时减少模型的计算和内存消耗。如表 1 所示,相比于其他算法, WRC_Net 在 SPDPSD 数据集上表现出更高的检测性能,在 mAP、mAP75、mAPs、mAPm、mAPl 等指标上均取得了一定提升。同时, WRC_Net 实现了精度和速度上的平衡,即在保持较高检测性能的同时具有相对较高的计算效率。通过与其他算法的比较, WRC_Net 在病理图像伪影检测方面的优越性得到了验证。

在 NCPDCSD 数据集上,本文验证了 WRC_Net 的通用性。如表 2 所示,与其他算法相比,本文方法仍展现出一定的竞争力。通过在该数据集上的验证, WRC_Net 在检测更多种类伪影的情况下依然保持较高的性能,这表明本文的模型具有较好的泛化能力和适应性。

表 1 SPDPSD 数据集上对比实验
Table 1 Comparative experiments on SPDPSD dataset

方法	mAP	mAP05	mAP75	mAPs	mAPm	mAPl	参数量/M	计算量/GFlops
FastestDet	0.511 441	0.750 893	0.581 395	0.401 485	0.633 397	0.810 657	0.24	0.13
YOLOX-s	0.595 196	0.785 986	—	—	—	—	8.94	26.77
YOLOv5s	0.627	0.843	—	—	—	—	7.03	16.0
YOLOv7-tiny	0.587	0.802	—	—	—	—	6.03	13.2
CornerNet-Lite	0.597	0.637	0.622	0.202	0.622	0.834	116.86	157.51
CenterNet2	0.455 882	0.641 002	0.522 349	0.289 82	0.461 265	0.572 584	45.26	53.28
WRC_Net(本文)	0.630 869	0.817 162	0.703 747	0.547 372	0.738 799	0.853 862	30.14	13.65

注:加粗字体表示各列最优结果,“—”表示对比实验中没有该项指标。

表 2 NCPDCSD 数据集上对比实验
Table 2 Comparative experiments on NCPDCSD dataset

方法	mAP	mAP05	mAP75	mAPs	mAPm	mAPl	参数量/M	计算量/GFlops
FastestDet	0.334 264	0.565 022	0.365 91	0.190 553	0.406 063	0.584 418	0.24	0.13
YOLOX-s	0.519 827	0.752 11	—	—	—	—	8.94	26.77
YOLOv5s	0.511	0.762	—	—	—	—	7.03	16.0
YOLOv7-tiny	0.394	0.594	—	—	—	—	6.03	13.2
CornerNet-Lite	0.556	0.621	0.590	0.207	0.523	0.676	116.86	157.51
CenterNet2	0.341	0.533	0.361	0.137	0.298	0.395	45.26	53.28
WRC_Net(本文)	0.550 147	0.688 046	0.623 537	0.371 297	0.663 119	0.776 049	30.14	13.65

注:加粗字体表示各列最优结果,“—”表示对比实验中没有该项指标。

3. 2. 2 消融实验

本文在 SPDPSD 数据集上对模型结构进行了消融实验,来验证模型的有效性。

1) Backbone 选取。首先对不同模型的特征提取能力进行了比较,包括 DenseNet161(Huang 等, 2017)、ShuffleNetV2(Ma 等, 2018)、EfficientNetB0

(Tan 和 Le, 2020)、MobileNetV2(Sandler 等, 2018)、ResNet50 和 InceptionV3(Szegedy 等, 2016)。如表 3 所示, ResNet50 在所有评价指标上的性能均优于其他模型。ResNet50 是 ResNet 系列中的一个经典模型, 具有强大的特征提取能力、性能与计算复杂度的平衡, 同时其具有较强的迁移学习能力。在目标检测任务中, 准确的目标定位和识别是关键, 而高质量的特征能够提供更好的目标表示。因此, 使用

ResNet50 作为本文的特征提取网络。

此外, 还对 ResNet50 的不同层进行了实验。具体来说, 选取了 ResNet50 中的不同层输出作为特征进行实验。如表 4 所示, 选取 ResNet 的第 3 层特征时, 模型性能最佳。第 3 层特征能够提供最佳的表征能力, 从而更好地捕捉病理切片上伪影的特征信息。因此, 本文在后续实验中均使用 ResNet50 的第 3 层特征作为特征提取模块预先提取的特征。

表 3 不同模型的特征提取能力比较
Table 3 Comparison of feature extraction capabilities of different models

模型	mAP	mAP05	mAP75	mAPs	mAPm	mAPI
DenseNet161	0.546 27	0.748 818	0.632 691	0.443 643	0.661 832	0.787 376
ShuffleNetV2	0.464 43	0.723 571	0.523 417	0.362 023	0.578 193	0.786 302
EfficientNetB0	0.537 509	0.743 031	0.620 543	0.436 087	0.657 848	0.786 302
MobileNetV2	0.519 941	0.742 424	0.598 346	0.417 373	0.639 996	0.809 449
ResNet50	0.590 916	0.778 550	0.683 883	0.497 146	0.697 771	0.887 888
InceptionV3	0.439 679	0.679 772	0.491 426	0.370 44	0.567 614	0.660 787

注: 加粗字体表示各列最优结果。

表 4 ResNet50 中 Layer 层数选取
Table 4 Layer selection in ResNet50

层	mAP	mAP05	mAP75	mAPs	mAPm	mAPI
L1	0.529 228	0.750 021	0.593 834	0.448 994	0.645 087	0.706 681
L2	0.590 916	0.791 544	0.656 273	0.516 298	0.702 959	0.761 176
L3	0.603 751	0.795 137	0.676 065	0.521 396	0.701 243	0.844 274
L4	0.590 916	0.778 55	0.683 883	0.497 146	0.697 771	0.887 888

注: 加粗字体表示各列最优结果。

(1) WRC 模块。为了验证提出的 WRC_Att (WRC attention) 和 MSDM 对模型性能的影响, 本文分别对 WRC 模块中仅包含 WRC_Att 模块、仅包含 MSDM 模块以及同时包含两个模块的情况进行了比较。如表 5 所示, 当同时使用 WRC_Att 和 MSDM 两个模块时, 模型性能最佳。WRC_Att 模块能够更好地捕获全局和局部信息, 而 MSDM 模块通过增大卷积核的感受野来引入更多的上下文信息, 以辅助检测任务。两个模块的结合能降低由于伪影形状变化而引起的性能下降, 从而提升整体模型的性能。

(2) 不同 WRC 模块层融合。研究了融合不同层的特征对模型性能的影响。如表 6 所示, 将不同层的特征进行拼接, 相较于仅使用最后一层 WRC 模块

层特征, 模型性能有了显著提升。这表明了通过拼接不同层特征, 能够充分利用不同层级的信息, 从而提升模型的感知能力和特征表达能力。可以发现, 即使仅使用单个检测头, 模型仍能有效地检测不同尺度的目标。

2) WRC 注意力模块。在 WRC 注意力模块中, 本文探究了 Col Attention 与 Row Attention 以及 CEFF 结构对模型性能的影响。如表 7 所示, Window_Att 为仅使用窗口注意力 Win_Col_Row_Att 为同时使用窗口注意力、水平注意力和垂直注意力。由于单独使用垂直注意力或水平注意力无法充分融合全局信息, 在这里仅测试了同时使用 Col Attention 和 Row Attention 的情况, 没有对单独使用其中一种注意力

表 5 WRC 模块对模型性能的影响

Table 5 The impact of the WRC module on model performance

WRC 模块	mAP	mAP05	mAP75	mAPs	mAPm	mAPl
WRC_Att	0.612 449	0.814 684	0.689 55	0.529 065	0.721 311	0.821 171
MSDM	0.621 78	0.808 233	0.702 083	0.535 218	0.725 235	0.857 327
WRC_Att+ MSDM	0.630 869	0.817 162	0.703 747	0.547 372	0.738 799	0.853 862

注:加粗字体表示各列最优结果。

表 6 WRC 模块层融合

Table 6 WRC module layer fusion

是否融合	mAP	mAP05	mAP75	mAPs	mAPm	mAPl
不融合	0.616 965	0.792 879	0.692 807	0.529 479	0.726 244	0.870 007
融合	0.630 869	0.817 162	0.703 747	0.547 372	0.738 799	0.853 862

注:加粗字体表示各列最优结果。

的情况进行实验。第 3 个实验 Win_Col_Row_Att+CEFF 则是在第 2 个实验的基础上引入了 CEFF 结构。如表 7 所示,在窗口注意力的基础上加上水平注意力和垂直注意力时,mAP50、mAP75、mAPs 和 mAPm 均有所提升。这证明了同时提取局部信息

和水平、垂直方向上的信息更好地适应不同形状伪影的检测。在此基础上再引入 CEFF 结构,mAP50、mAP75 以及 mAPm 均提升了约 1%,进一步验证了对 3 种注意力重新加权并进行特征融合的有效性。

表 7 WRC 注意力模块对模型性能的影响

Table 7 The impact of the WRC attention module on model performance

WRC 注意力模块	mAP	mAP05	mAP75	mAPs	mAPm	mAPl
Window_Att	0.622 089	0.803 966	0.701 178	0.539 158	0.725 586	0.862 055
Win_Col_Row_Att	0.621 193	0.807 937	0.698 743	0.540 648	0.728 358	0.843 842
Win_Col_Row_Att+CEFF	0.630 869	0.817 162	0.703 747	0.547 372	0.738 799	0.853 862

注:加粗字体表示各列最优结果。

3)MSDM 模块。本文研究了 MSDM 中膨胀卷积的不同连接方式对模型性能的影响,并分别对并联、串联、并联 + SE 以及串联 + SE 这 4 种方式进行了验证。如表 8 所示,串联连接方式的性能要优于并联。通过串联连接的方式,膨胀卷积可以在不同扩张率

下逐层提取特征,使模型能够更好地捕捉多尺度信息。并且在引入 SE 模块后,无论是并联和串联连接方式,性能均有了一定的提升。进一步表明了本文采用的串联+SE 方式增强了模型在捕获多尺度信息方面的优势。

表 8 MSDM 模块不同连接方式对模型性能的影响

Table 8 The impact of different connection modes of MSDM module on model performance

连接方式	mAP	mAP05	mAP75	mAPs	mAPm	mAPl
并联	0.585 522	0.791 261	0.657 897	0.498 353	0.715 589	0.728 468
串联	0.624 011	0.808 978	0.700 984	0.538 498	0.729 872	0.880 195
并联 + SE	0.585 161	0.802 212	0.663 522	0.502 174	0.702 06	0.778 145
串联 + SE	0.630 869	0.817 162	0.703 747	0.547 372	0.738 799	0.853 862

注:加粗字体表示各列最优结果。

4)ECABlock 模块。本文研究了在特征融合模块最后加 ECA 模块对模型性能的影响。如表 9 所示,加入 ECA 模块后模型在 mAP、mAP05、mAP75、mAPs 和 mAPm 上性能均有了一定的提

升。在本文中,ECA 模块对不同层特征图在通道维度上赋予了不同的权重,使模型能更准确地捕捉到重要特征,进而提高了模型的性能与泛化能力。

表 9 ECA 模块对模型性能的影响
Table 9 The impact of the ECA module on the performance of the model

ECA 模块	mAP	mAP05	mAP75	mAPs	mAPm	mAPl
未使用	0.623 290	0.807 64	0.699 555	0.531 241	0.736 435	0.865 222
使用	0.630 869	0.817 162	0.703 747	0.547 372	0.738 799	0.853 862

注:加粗字体表示各列最优结果。

5)图像处理速度比较。为了验证使用缩略图带来的速度优势,进行了对比实验。将同一幅病理切片分别通过两种不同策略处理,即裁剪成图像块和转换成缩略图。采用了 10 折交叉验证来进行全面的评估。如表 10 所示,切图(521×512)表示将病理切片裁剪成尺寸 512×512 像素的图像块;切图(256×256)表示将病理切片裁剪成 256×256 像素尺寸的图像块;缩略图(1 024×1 024)表示将切片转换成尺寸为 1 024×1 024 像素的缩略图。结果表明,直接将切片转换成缩略图的速度明显优于裁剪成图像块。

表 10 图像处理速度比较
Table 10 Comparison of image processing speed

病理切片/像素	时间/s
切图(512×512)	29.46
切图(256×256)	11.88
缩略图(1 024×1 024)	0.55

注:加粗字体表示最优结果。

4 结 论

本文提出了一种新的数字病理切片伪影检测算法 WRC_Net。该网络仅含有单个检测头,但通过汇集不同感受野特征,使其具备适应不同尺度物体的检测能力。此外,还提出了 WRC 模块,旨在解决病理切片中不同尺度和形状的伪影问题。该模块由两部分组成:1)能够同时捕捉图像中全局和局部信息的注意力模块;2)能提取多尺度信息的多尺度扩张模块。还建立了两个数字病理切片伪影数据集

(SPDPSD 数据集和 NCPDCSD 数据集),并在这两个数据集上验证了 WRC_Net 的性能,证明其在精度和速度上取得了平衡。值得强调的是,这两个数据集来源于不同机构,并且包含不同种类的伪影。在这两个数据集上的实验,进一步验证了本文方法在切片伪影检测上的有效性。未来的工作可以考虑构建规模更大、包含更多伪影种类的数据集。通过在这些数据集上验证 WRC_Net 的性能,将有助于提高模型的泛化能力,并推动数字病理切片质量控制的发展。

参考文献(References)

Babaie M and Tizhoosh H R. 2019. Deep features for tissue-fold detection in histopathology images//Proceedings of the 15th European Congress on Digital Pathology. Warwick, UK: Springer International Publishing: 125-132 [DOI: 10.1007/978-3-030-23937-4_15]

Bautista P A and Yagi Y. 2010. Improving the visualization and detection of tissue folds in whole slide images through color enhancement. Journal of Pathology Informatics, 1(1): #25 [DOI: 10.4103/2153-3539.73320]

Bochkovskiy A, Wang C Y and Liao H Y M. 2020. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. [2023-09-28]. <https://arxiv.org/pdf/2004.10934.pdf>

Deng Y, Feng M, Jiang Y, Zhou Y Y, Qin H Y, Xiang F, Wang Y Z, Bu H and Bao J. 2020. Development of pathological reconstructed high-resolution images using artificial intelligence based on whole slide image. MedComm, 1(3): 410-417 [DOI: 10.1002/mco.2.39]

Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X H, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J and Hounsby N. 2021. An image is worth 16×16 words: Transformers for image recognition at scale [EB/OL].

- [2023-09-28]. <https://arxiv.org/pdf/2010.11929.pdf>
- Everingham M, Van Gool L, Williams C K I, Winn J and Zisserman A. 2010. The Pascal visual object classes (VOC) challenge. *International Journal of Computer Vision*, 88 (2) : 303-338 [DOI: 10.1007/s11263-009-0275-4]
- Foucort A, Debeir O and Decaestecker C. 2018. Artifact identification in digital pathology from weak and noisy supervision with deep residual networks//*Proceedings of the 4th International Conference on Cloud Computing Technologies and Applications*. Brussels, Belgium: IEEE: 1-6 [DOI: 10.1109/CloudTech.2018.8713350]
- Ge Z, Liu S T, Wang F, Li Z M and Sun J. YOLOX: exceeding YOLO series in 2021 [EB/OL]. [2023-09-28]. <https://arxiv.org/pdf/2107.08430.pdf>
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hemmatirad K, Babaie M, Afshari M, Maleki D, Saiadi M and Tizhoosh H R. 2022. Quality control of whole slide images using the YOLO concept//*Proceedings of the 10th IEEE International Conference on Healthcare Informatics (ICHI)*. Rochester, USA: IEEE: 282-287 [DOI: 10.1109/ICHI54592.2022.00049]
- Ho J, Kalchbrenner N, Weissenborn D and Salimans T. 2019. Axial attention in multidimensional transformers[EB/OL]. [2023-09-28]. <http://arxiv.org/pdf/1912.12180.pdf>, 2019-12-20
- Hossain M S, Shahriar G M, Syeed M M M, Uddin M F, Hasan M, Hossain M S and Bari R. 2023. Tissue artifact segmentation and severity assessment for automatic analysis using WSI. *IEEE Access*, 11: 21977-21991 [DOI: 10.1109/ACCESS.2023.3250556]
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]
- Huang G, Liu Z, van der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE: 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Janowczyk A, Zuo R, Gilmore H, Feldman M and Madabhushi A. 2019. HistoQC: an open-source quality control tool for digital pathology slides. *JCO Clinical Cancer Informatics*, 3 (3) : 1-7 [DOI: 10.1200/CCI.18.00157]
- Kanwal N, Pérez-Bueno F, Schmidt A, Engan K and Molina R. 2022. The devil is in the details: whole slide image acquisition and processing for artifacts detection, color variation, and data augmentation: a review. *IEEE Access*, 10: 58821-58844 [DOI: 10.1109/ACCESS.2022.3176091]
- Kothari S, Phan J H and Wang M D. 2013. Eliminating tissue-fold artifacts in histopathological whole-slide images for improved image-based prediction of cancer grade. *Journal of Pathology Informatics*, 4(1): #22 [DOI: 10.4103/2153-3539.117448]
- Law H and Deng J. 2018. CornerNet: detecting objects as paired keypoints//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer: 734-750 [DOI: 10.1007/978-3-030-01264-9_45]
- Law H, Teng Y, Russakovsky O and Deng J. 2020. CornerNet-Lite: efficient keypoint based object detection [EB/OL]. [2023-09-28]. <https://arxiv.org/pdf/1904.08900.pdf>
- Ma N N, Zhang X Y, Zheng H T and Sun J. 2018. ShuffleNet v2: practical guidelines for efficient CNN architecture design//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer: 122-138 [DOI: 10.1007/978-3-030-01264-9_8]
- Ma X H. 2022. FastestDet: ultra lightweight anchor-free real-time object detection algorithm [CP/OL]. [2023-09-28]. <https://github.com/dog-qiuqiu/FastestDet>
- Maleki D, Afshari M, Babaie M and Tizhoosh H R. 2020. Ink marker segmentation in histopathology images using deep learning//*15th International Symposium on Visual Computing*. San Diego, USA: Springer: 359-368 [DOI: 10.1007/978-3-030-64556-4_28]
- Noman M, Fiaz M, Cholakkal H, Narayan S, Anwer R M, Khan S and Khan F S. 2023. Remote sensing change detection with transformers trained from scratch [EB/OL]. [2023-09-28]. <https://arxiv.org/pdf/2304.06710.pdf>
- Sandler M, Howard A, Zhu M L, Zhmoginov A and Chen L C. 2018. MobileNetV2: inverted residuals and linear bottlenecks//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE: 4510-4520 [DOI: 10.1109/CVPR.2018.00474]
- Shakarami A, Nicolè L, Terreran M, Tos A P D and Ghidoni S. 2023. TCNN: a transformer convolutional neural network for artifact classification in whole slide images. *Biomedical Signal Processing and Control*, 84: #104812 [DOI: <https://doi.org/10.1016/j.bspc.2023.104812>]
- Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z. 2016. Rethinking the inception architecture for computer vision//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE: 2818-2826 [DOI: 10.1109/CVPR.2016.308]
- Tan M X and Le Q V. 2020. EfficientNet: rethinking model scaling for convolutional neural networks [EB/OL]. [2023-09-28]. <https://arxiv.org/pdf/1905.11946.pdf>
- Tong L, Sha Y and Wang M D. 2019. Improving classification of breast cancer by utilizing the image pyramids of whole-slide imaging and multi-scale convolutional neural networks//*Proceedings of the 43rd IEEE Annual Computer Software and Applications Conference (COMPSAC)*. Milwaukee, USA: IEEE: 696-703 [DOI: 10.1109/COMPSAC.2019.00105]
- Ultralytics. 2020. YOLOv5 [EB/OL]. [2023-09-23].

<https://github.com/ultralytics/yolov5>

Wang C Y, Bochkovskiy A and Liao H Y M. 2023. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE: 7464-7475 [DOI: 10.1109/CVPR52729.2023.00721]

Wang Q L, Wu B G, Zhu P F, Li P H, Zuo W M and Hu Q H. 2020. ECA-Net: efficient channel attention for deep convolutional neural networks//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 11531-11539 [DOI: 10.1109/CVPR42600.2020.01155]

Woods A E and Ellis R C. 1994. Laboratory Histopathology: A Complete Reference. Edinburgh: Churchill Livingstone

Yun S, Han D, Chun S, Oh S J, Yoo Y and Choe J. 2019. CutMix: regularization strategy to train strong classifiers with localizable features//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE: 6022-6031 [DOI: 10.1109/ICCV.2019.00612]

Zhao Y L, Ding W L, You Q H, Zhu F L, Zhu X J, Zheng K and Liu D D. 2023. Classification of whole slide images of breast histopathology based on spatial correlation characteristics. Journal of Image and Graphics, 28(4): 1134-1145 (赵樱莉, 丁维龙, 游庆华,

朱峰龙, 朱筱婕, 郑魁, 刘丹丹. 2023. 融合空间相关性特征的乳腺组织病理全切片分类. 中国图象图形学报, 28(4): 1134-1145 [DOI: 10.11834/jig.211133]

Zhou X Y, Koltun V and Krähenbühl P. 2021. Probabilistic two-stage detection [EB/OL]. [2023-09-28].
<https://arxiv.org/pdf/2103.07461.pdf>

作者简介

丁维龙,男,教授,博士生导师,主要研究方向为医疗人工智能与虚拟仿真。E-mail: wlding@zjut.edu.cn

汪春年,通信作者,女,主任医师,硕士生导师,主要研究方向为人工智能技术在肿瘤病理诊断中的应用。

E-mail: niana217@126.com

廖婉茵,女,硕士研究生,主要研究方向为数字病理图像处理。E-mail: wyliao1216@qq.com

朱伟,男,硕士研究生,主要研究方向为数字病理图像处理。

E-mail: zhuvuee@outlook.com

祝行琴,女,主管技师,主要研究方向为病理技术HE染色制片质量控制。E-mail: zhu20061050223@163.com

朱红波,女,副主任医师,主要研究方向为乳腺肿瘤病理诊断。E-mail: hzbzhu2001@126.com