中图法分类号:TP391.41　　**文献标识码**:A　　**文章编号**:1006-8961(XXXX)XX-0001-19

论文引用格式:Huang Zhenjie, Li Bingbing, Wang Huadeng, Lan Rushi, Xiao Qiuxiang, Fang Weilan, He Fen, Liu Liyuan, Guo Zhenqiang, Chen Longjiao, Liang Li. Clustering refinement and difficulty-aware contrastive learning for mitosis detection in breast histopathology images[J/OL]. Journal of Image and Graphics, XXXX:1-19. DOI: 10.11834/jig.260604. (黄振杰, 黎兵兵, 汪华登, 蓝如师, 肖秋香, 方伟兰, 何芬, 刘力源, 郭振强, 陈龙娇, 梁莉. 聚类细化与难度感知对比学习的乳腺病理图像有丝分裂核检测[J/OL]. 中国图象图形学报,XXXX:1-19. DOI: 10.11834/

聚类细化与难度感知对比学习的乳腺病理图像有丝分裂核检测

黄振杰[1],黎兵兵[2,3,4],汪华登[1],蓝如师[1],肖秋香[5],方伟兰[6],何芬[5],刘力源[7,8,9],郭振强[6],陈龙娇[6],梁莉[2,3,4]

1. 桂林电子科技大学计算机与信息安全学院,桂林　541004; 2. 南方医科大学南方医院病理科,广州　510515; 3. 南方医科大学基础医学院病理学系,广州　510515; 4. 广东省分子肿瘤病理学重点实验室,广州　510515; 5. 赣南医科大学第一附属医院病理科,赣州　341000; 6. 广东省人民医院赣州医院(赣州市立医院)病理科,赣州　341000; 7. 赣州市人民医院病理科,赣州　341000; 8. 南昌大学附属赣州医院病理科,赣州　341000; 9. 南方医科大学南方医院赣州医院病理科,赣州　341000; (#为通信作者)

摘　要:**目的**　乳腺癌病理图像中的有丝分裂核计数是肿瘤组织学分级和预后评估的重要依据,但人工计数耗时、主观性强,自动检测又受到标注数据有限、类别失衡、形态异质性及困难负样本干扰。为提高复杂病理场景下的检测精度,构建一种基于难度感知对比学习与特征优化的渐进式检测框架。**方法**　首先,设计混合上下文聚合模块,通过内容相关的通道选择动态融合局部形态细节与全局组织语义;其次,利用多层级特征、位置编码和软聚类构建多特征聚类细化模块,对类激活图进行聚类感知重建,抑制背景误激活并改善目标定位;最后,以细化后的高置信度区域生成伪标签,结合预热策略、像素级正负样本选择和困难负样本挖掘实施难度感知对比学习,并联合优化分类损失、定位损失和对比损失。**结果**　所提方法在TUPAC16数据集上的精确率、召回率和F1值分别为0.790、0.756和0.773;在临床数据集GZMH-V2的Test1和Test2上,F1值分别为0.578和0.558;在MIDOG22数据集上取得0.693的F1值。临床验证中,病理医师与真实标注之间的组内相关系数由0.224提高到0.323,相对提升44.2%;单张高倍视野图像平均阅片时间缩短9.937s,差异具有统计学意义(p=0.015)。**结论**　混合上下文聚合、多特征聚类细化与难度感知对比学习能够从特征表征、空间定位和困难样本判别三个层面协同提升有丝分裂核检测性能,并具有辅助病理医师提高判读一致性与效率的应用潜力。

关键词:乳腺癌病理图像;有丝分裂核检测;难度感知对比学习;混合上下文聚合;多特征聚类细化;类激活图

Clustering refinement and difficulty-aware contrastive learning for mitosis detection in breast histopathology images

Huang Zhenjie[1], Li Bingbing[2,3,4], Wang Huadeng[1], Lan Rushi[1], Xiao Qiuxiang[5], Fang Weilan[6], He Fen[5], Liu Liyuan[7,8,9], Guo Zhenqiang[6], Chen Longjiao[6], Liang Li[2,3,4]

1. School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, 541004, China; 2. Department of Pathology, Nanfang Hospital, Southern Medical University, Guangzhou, 510515, China. ; 3. Department of Pathology, School of

收稿日期:2026-09-21;**修回日期**:2026-11-17

基金项目:广西科技重大专项(桂科AA24263033);广西自然科学基金(2024GXNSFAA010342);国家自然科学基金(62362014,62672143)

Supported by: This work was partially supported by the Guangxi Science and Technology Major Program (GuikeAA24263033), the Guangxi Natural Science Foundation (Grant No. 2024GXNSFAA010342), and the National Natural Science Foundation of China (Grant No. 62362014, Grant No. 62672143).

© 中国图象图形学报版权所有

Basic Medical Sciences, Southern Medical University, Guangzhou, 510515, China.; 4. Guangdong Province Key Laboratory of Molecular Tumor Pathology, Guangzhou, 510515, China.; 5. Department of Pathology, The First Affiliated Hospital of Gannan Medical University, Ganzhou, 341000, China; 6. Department of Pathology, Ganzhou Hospital of Guangdong Provincial People's Hospital, Ganzhou Municipal Hospital, Ganzhou, 341000, China; 7. Department of Pathology, Ganzhou People's Hospital, Ganzhou, 341000, China; 8. Department of Pathology, The Affiliated Ganzhou Hospital of Nanchang University, Ganzhou, 341000, China; 9. Department of Pathology, Ganzhou Hospital-Nanfang Hospital, Southern Medical University, Ganzhou, 341000, China

Abstract: Objective Mitotic figure counting in hematoxylin-and-eosin-stained breast cancer histopathology images is an important component of tumor grading and prognostic assessment. Reliable counting is difficult because mitotic figures occupy only a very small proportion of nuclei within a high-power field and exhibit substantial morphological variation across the phases of cell division. Late mitotic figures may appear as two connected components, while apoptotic nuclei, lymphocytes, hyperchromatic nuclei, and densely packed cells may have a similar appearance. Manual assessment is consequently time-consuming and depends strongly on the experience of the pathologist, which can produce observer variability and limited repeatability. Automatic methods can reduce this burden, but their performance is restricted by scarce point-level annotations, severe foreground-background imbalance, coarse localization responses, and the dominance of easy negatives during training. This study aims to improve mitotic figure detection under these conditions by developing a progressive framework that combines adaptive context modeling, class activation map refinement, and difficulty-aware metric learning. **Method** The framework contains three coordinated components: hybrid context aggregation (HCA), multi-feature clustering refinement (MFCR), and difficulty-aware contrastive learning (DACL). HCA is embedded in the encoder stages of a ResNet-based baseline. A channel selector first estimates channel importance and assigns content-dependent decision scores. The feature channels are then routed to a local branch or a global branch rather than being divided with fixed proportions. The local branch applies depthwise convolution to retain nuclear boundaries, staining intensity, and membrane morphology. The global branch reshapes the selected features into multiple heads, derives query, key, and value representations, and aggregates long-range contextual information through attention. The outputs of the local and global branches are concatenated with the retained head feature to form the representation forwarded to the next encoder stage. MFCR addresses the fragmented response, background activation, and center displacement that commonly occur in class activation maps. Activation maps from different network depths are combined so that shallow features contribute spatial detail and deep features contribute semantic discrimination. A reduced guidance feature is concatenated with two-dimensional positional encoding, and a clustering projection head produces pixel-wise soft assignments. Each soft region is weighted by its mean activation response, normalized to reduce the bias caused by unequal region areas, and used to reconstruct a cluster-aware activation map. Min-max normalization places the refined response in a consistent range. A differentiable localization constraint further measures the Euclidean distance between the spatial expectation of the predicted response and the pathologist-annotated centroid. The refined activation map then provides pseudo-labels for DACL. For each anchor, positive samples share the same pseudo-label, while negative candidates are ranked by similarity to the anchor. Only the most confusing proportion of negative samples is retained together with all valid positives to construct a restricted contrastive set. This design concentrates the repulsive gradient on morphologically similar nuclei instead of allowing numerous easy background pixels to dominate optimization. Because early activation maps may be noisy, the contrastive objective is activated only after a warm-up period. The final objective jointly optimizes image-level classification loss, centroid localization loss, and difficulty-aware contrastive loss. Experiments use the public TUPAC16 and MIDOG22 benchmarks and the clinically derived GZMH-V2 dataset, which contains 2322 high-power field patches with a resolution of 2084 by 2084 pixels and 5291 annotated mitotic nuclei. All images are converted to RGB and cropped or resized to 500 by 500 pixels. Training uses random flipping, rotation, color perturbation, and normalization. A prediction is considered correct when its center lies within 30 pixels of a ground-truth centroid, with each annotation matched to at most one prediction. Precision, recall, and F1 score are used for detection evaluation. A clinical validation additionally compares pathologist counts with reference annotations using intraclass correlation coefficients and evaluates reading time with a paired-sample t test. **Result** On TUPAC16, the proposed framework achieves a precision of 0.790, a recall of 0.756, and an F1

© 中国图象图形学报版权所有

score of 0.773. Its recall and F1 score exceed those of the two-stage Uninformed Teacher-Student method by 6.3% and 3.1%, respectively, while retaining high precision. On the more challenging GZMH-V2 dataset, the method obtains balanced precision and recall of 0.583 and 0.573 on Test1, corresponding to an F1 score of 0.578; the Test2 F1 score is 0.558. On MIDOG22, the end-to-end framework reaches an F1 score of 0.693, close to the 0.696 result of the two-stage Uninformed Teacher-Student method and higher than the reported results of the reproducible comparison models used in the study. Stepwise ablation on TUPAC16 shows that HCA, MFCR, and DACL each improve the baseline and provide complementary gains when combined. The full configuration yields the best overall performance; internal ablations also support dynamic channel selection, multi-level feature fusion, a 15-epoch warm-up, and ResNet-18 as the backbone under the reported setting. The t-distributed stochastic neighbor embedding visualization shows a clearer boundary between mitotic features and confusing background features. In clinical validation, the intraclass correlation coefficient (ICC) between the four pathologists′ mean counts per high-power field and reference counts increases from 0.224 to 0.323 with model assistance, an absolute increase of 0.099. The corresponding coefficient for less-experienced pathologists rises from 0.279 to 0.450, whereas that for more-experienced pathologists rises from 0.122 to 0.167. The mean reduction in assessment time is 9.937 seconds per high-power field, and the paired-sample test gives p=0.015. **Conclusion** The results indicate that adaptive local-global context aggregation supplies a stronger feature basis, clustering-guided activation refinement improves spatial localization, and difficulty-aware contrastive learning increases separation from visually deceptive negatives. Their combination provides competitive performance across public, challenge, and clinical datasets and can improve both the consistency and efficiency of pathologist assessment. The remaining limitations are mainly associated with inaccurate initial activation maps in densely cellular, unevenly stained, or structurally complex regions and with the dependence of clustering refinement on those maps. Future work should therefore include multi-center validation and more robust suppression of activation noise to strengthen generalization and stability.

Key words: breast cancer histopathology images; mitotic figure detection; difficulty-aware contrastive learning; hybrid context aggregation; multi-feature clustering refinement; class activation map

论文引用格式:[DOI:10.11834/jig.260604]

0 引　言

乳腺癌是全球女性癌症相关死亡的第二大主要原因。其发生发展是一个涉及多种细胞类型的多步骤过程,且实现有效的预防仍是当前全球面临的重大挑战(Sun 等,2017)。在乳腺癌的临床分类标准中,肿瘤细胞核的有丝分裂计数是最为关键的指标之一。病理医师通常通过评估经苏木精-伊红(hematoxylin and eosin,HE)染色的组织病理学切片,对有丝分裂核进行人工计数(Tellez 等,2018)。有丝分裂计数不仅反映了肿瘤的增殖活性,更在判定肿瘤恶性程度及评估患者预后方面发挥着至关重要的作用。然而,这一人工流程耗时费力,且高度依赖病理医师的专业经验,往往导致诊断结果存在主观性、不一致性以及较低的可重复性。

近年来,深度学习技术的飞速发展引发了医学影像分析领域的深刻变革。自动化计算方法已被广泛应用于乳腺癌组织病理学中的有丝分裂核检测。深度学习算法在计算病理学领域取得了令人瞩目的成就,特别是在将有丝分裂核检测建模为分类(Beevi 等,2019;Cireşan 等,2013;Zerhouni 等,2017;Sohail 等,2021)、分割(Li 等,2019;Wang 等,2023)或目标检测(Wollmann 和 Rohr,2021;Gonzalez 等,2023)问题的监督学习任务中。近期,为进一步提升检测精度,研究人员提出了多种多阶段方法(Wang 等,2024a;Wang 等,2025;Tang 和 Cai,2024;Li 等,2018;Fernandez-Martín 等,2024),此类方法通常先生成候选区域,随后进行更精细化的分类与定位。尽管取得了上述进展,有丝分裂核检测依然极具挑战性,其主要受制于以下问题:标注数据受限且为弱标注、类别严重不平衡、各分裂期形态的多样性,以及困难样本的存在。

标注数据的匮乏构成了自动化有丝分裂核检测所面临的最严峻挑战之一。对有丝分裂核进行精确标注是一项耗时且费力的工作,通常需要多名病理医师达成共识(Fernandez-Martín 等,2024)。人工标注不仅需要投入大量精力,还极易受到观察者间差异的影响。公开的有丝分裂核数据集通常仅提供基

© 中国图象图形学报版权所有

于质心的标签或粗糙的像素级标注(Veta等,2019)。此外,类别不平衡是该任务固有的问题,在单个高倍视野(high-power field, HPF)内,与背景细胞核相比,有丝分裂核的数量极低。进一步而言,如图1第一行所示(第二行为非有丝分裂核),有丝分裂核会经历多个阶段,每个阶段均呈现出截然不同的形态特征,这显著加剧了识别的复杂性。例如,如图1第一行第二列的切片所示,处于末期的细胞核在视觉上已分裂为两部分,但仍应被计为一个分裂期细胞,这种现象极易引发潜在的误分类。而且,密集排列的细胞以及形态上与有丝分裂核高度相似的淋巴细胞极易导致假阳性(Barlow,1977)。在标注过程中,病理医师们常对那些与有丝分裂核极为相似的非分裂期细胞核产生分歧;这类模棱两可的实例被统称为困难负样本。尽管困难样本能够促进模型学习到复杂的特征表示,但它们同时也增加了模型获取具有强判别力特征的难度。

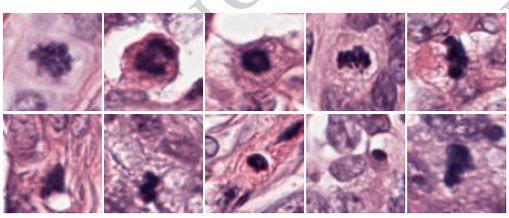

图1 有丝分裂核与非有丝分裂核示意图

Fig. 1 Samples of mitotic and non-mitotic nuclei

针对定位响应粗糙与困难负样本混淆相互影响的问题,本文构建"类激活图(class activation map, CAM)细化—伪标签构建—困难负样本学习"的检测框架。混合上下文聚合提供局部形态与上下文特征,多特征聚类细化将融合CAM重建为区域一致的响应,难度感知对比学习再依据细化响应生成的伪标签优化像素特征。该设计将定位细化结果用于困难样本选择,而非仅将定位和判别作为彼此独立的优化任务。

本文在TUPAC16(Veta 等, 2019)、MIDOG22(Aubreville 等,2024)和临床数据集GZMH-V2上进行检测实验,并开展病理医师辅助阅片评估。检测结果用于分析框架的有效性,临床结果用于探索其辅助应用潜力。

本工作的主要贡献可总结如下:

1)提出了一种新颖的混合上下文聚合模块(hybrid context aggregation, HCA),该模块有效融合了局部与全局的上下文线索,保持了有效的信息传递。

2)开发了一种由软特征聚类引导的CAM细化机制,多特征聚类细化(multi-feature clustering refinement, MFCR),提升了类激活图的精度与空间一致性。

3)引入了一种由CAM引导的对比学习策略,难度感知对比学习(difficulty-aware contrastive learning, DACL),该策略结合了困难负样本的加权与挖掘。此方法缓解了困难负样本造成的干扰,并在细粒度层面上增强了模型的判别能力。

4)在公开的TUPAC16和MIDOG22数据集,以及本文作者构建和公开的临床数据集GZMH-V2上进行了实验。实验结果表明,所提方法在有丝分裂核检测任务中具有一定有效性和应用潜力。

1 相关工作

1.1 有丝分裂核检测

随着深度学习技术的快速发展,计算机视觉及其在医学图像分析中的应用取得了显著进展,特别是在有丝分裂核检测任务中。传统的有丝分裂核检测方法主要依赖于图像处理技术,如阈值分割、形态学特征提取以及几何分析(Beevi 等,2016a;Beevi 等,2016b)。例如,Khan等人(2013)采用了一种伽马-高斯(Gamma-Gaussian)混合模型,基于强度分布的差异来区分分裂期与非分裂期细胞;然而,手工设计特征的泛化能力依然有限。Irshad等人(2013)提出了一种有丝分裂核检测方法,该方法利用从多个成像通道中提取的纹理、尺度不变特征变换(scale-invariant feature transform,SIFT)、强度及形态学特征来实现自动化的有丝分裂核检测。尽管如此,当面对有丝分裂核在视觉外观上存在巨大差异且类间相似度较高的数据集时,这些方法往往表现不佳,从而无法满足现代病理图像分析的需求。

近年来,卷积神经网络(convolutional neural network,CNN)已成为医学图像分析领域的主要工具,特别是在有丝分裂核检测方面。CNN强大的特征学习能力得到了广泛应用,许多研究采用单阶段网络模型直接进行细胞核分类(Beevi 等,2019;Cireşan 等,2013;Zerhouni 等,2017;Sohail 等,2021;Maroof 等,2020;Dodballapur 等,2019)。例如,Beevi

© 中国图象图形学报版权所有

等人(2019)通过将预训练的CNN与随机森林分类器及初始全连接层相结合来增强网络,从而即使在训练数据有限的情况下,也能实现有效且准确的细胞核分类。此外,基于语义分割和目标检测的深度学习方法已被广泛应用于有丝分裂核的定位与检测(Li 等,2019; Wang 等,2023; Wollmann 和 Rohr, 2021; Gonzalez 等,2023; Lei 等,2021; Sebai 等,2020; Ronneberger 等,2015; Khanam 和 Hussain, 2024; He 等,2017)。Li等人(2019)使用结合了同心损失的全卷积网络(fully convolutional network, FCN),来解决仅有质心标注的病理图像中的有丝分裂核检测问题,并通过后处理滤波器提高了准确率。Wang等人(2023)将检测任务重新建模为语义分割任务,应用了基于U-Net的多尺度通道注意力机制以及基于傅里叶变换的数据增强方法。基于Mask R-CNN(He 等,2017)构建MaskMitosis(Sebai 等,2020)模型,在2012ICPR数据集的乳腺癌组织学切片自动核分裂检测中,取得了0.863的F1分数。然而,由于类别严重不平衡、外观形态多变,以及将自然图像检测模型迁移至病理图像存在困难等问题,这些方法依然面临一定的局限性,并导致假阴性率增加。

为了克服单阶段方法的局限性,许多研究探索了多阶段混合策略,以进一步改进细胞核检测。这些方法通常先进行粗略的候选区域检测,随后通过细粒度的分类或回归来优化结果(Wang 等,2024a; Wang 等,2025; Tang 和 Cai, 2024; Li 等,2018; Fernandez-Martín 等,2024)。林凡超等人(2023)利用外观与运动模式感知增强细胞状态判别,但其时序显微图像设置不同于本文的静态病理图像,运动信息不能直接用于本任务。例如,Wang等人(2024a)提出了两阶段级联网络FoCasNet,它将粗检测与精细分类相结合,以实现乳腺癌病理图像中有丝分裂核的精确检测。Fernandez-Martín 等人(2024)引入了无先验师生框架(uninformed teacher-student, UTS),其中经过强数据增强训练的教师模型基于质心距离识别困难样本,而经过弱数据增强训练的学生模型则从蒸馏后的干净数据集中进行学习。Nateghi 等人(2021)利用深度预处理和两步困难负样本挖掘流程解决了核分裂检测的挑战,从而能够在全切片图像中可靠地识别具有高核分裂活性的区域。Alom 等人(2020)开发了一个端到端的多任务系统MitosisNet,该系统结合了分割、检测和分类模块,随后采用IRRCNN来区分分裂期与非分裂期细胞。Lei 等人(2019)采用带有区域提议网络的VGG-16,并结合具有双分类分支的改进版R-CNN模块以提高检测精度。尽管现有的多阶段方法具有诸多优势,但它们仍然面临效率有限和细粒度优化不足的问题,这使得它们难以充分应对细胞核形态的多样性以及医学影像标注数据的稀缺性。

1.2 困难样本处理

处理困难样本是提升细胞核检测模型性能的关键。在医学图像分析中,由于细胞核形态的高度变异性以及不同病理背景下细胞核的相似性,困难样本的出现是不可避免的。例如,处于有丝分裂后期且尚未完全分离的细胞核极易与凋亡细胞相混淆,从而成为典型的困难样本。Cireşan 等人(2013)设计了一种分阶段的困难样本挖掘策略,即先使用初始模型过滤出难以区分的非有丝分裂样本,然后构建一个优化的训练集,以增强模型区分困难样本的能力。Lafarge 和 Koelzer(2022)通过一种重采样策略解决了非有丝分裂样本被误分类的问题。他们首先通过随机采样训练模型,随后剔除容易分类的样本,并重新采样困难负样本,从而强化模型的判别能力。Lafarge 和 Koelzer(2023)在此基础上进一步改进,通过自动从无标注区域中选择负样本作为“困难病例”并将其加入训练集,从而在训练数据有限的情况下,增强了模型的泛化能力并提升了其性能。刘蓉等(2026)通过自适应上下文聚合增强目标特征,并在图像、前景和类别层面进行域对抗对齐,以缓解跨设备域偏移及有丝分裂核与困难负样本的混淆。

在近期的研究中,C. Fernandez-Martín 等人(2024)提出了一种使用师生蒸馏机制的弱监督框架,该机制帮助学生模型聚焦于教师模型预测中难以分类的样本,从而间接提升了对困难样本的识别能力。Jahanifar 等人(2024)引入了一种两阶段模型:第一阶段执行快速的候选区域分割,而第二阶段则专门针对难以区分的样本进行分类细化,从而减少了假阳性和漏检现象。

尽管现有方法在处理细胞核检测中的困难样本时,已在数据过滤和模型架构等领域进行了有效的尝试,但依然存在诸多挑战,包括样本过滤准确率不足、特征区分过于粗糙,以及对标注或多阶段冗余设计的依赖。这些局限性阻碍了它们完全适应医学图

© 中国图象图形学报版权所有

像复杂多变的形态与背景，从而为通过像素级细粒度区分进行优化提供了空间。在本研究中，本文旨在通过利用像素级细粒度区分，来增强模型处理困难样本的能力。

2 方 法

图2详细展示了本文所提方法的整体网络框架。具体而言，该系统主要由三个核心组件协同构成：混合上下文聚合（HCA）模块、多特征聚类细化（MFCR）模块，以及难度感知对比学习（DACL）机制。在特征提取阶段，网络首先利用HCA模块在各个网络层中有效融合局部细节与全局上下文，从而捕获更为丰富的多尺度特征信息，并据此生成初步的多层级类激活图（CAM）。随后，为了克服传统CAM定位粗糙的局限性，本文使用MFCR方法对这些多层级CAM进行进一步的细化，以提升其空间一致性与激活精度。最后，基于细化后CAM提供的精准定位线索，模型从中动态提取具有高度代表性的正样本与困难负样本，进而执行难度感知对比学习，以此在像素层面上大幅增强模型对复杂细胞形态和易混淆实例的细粒度判别能力。

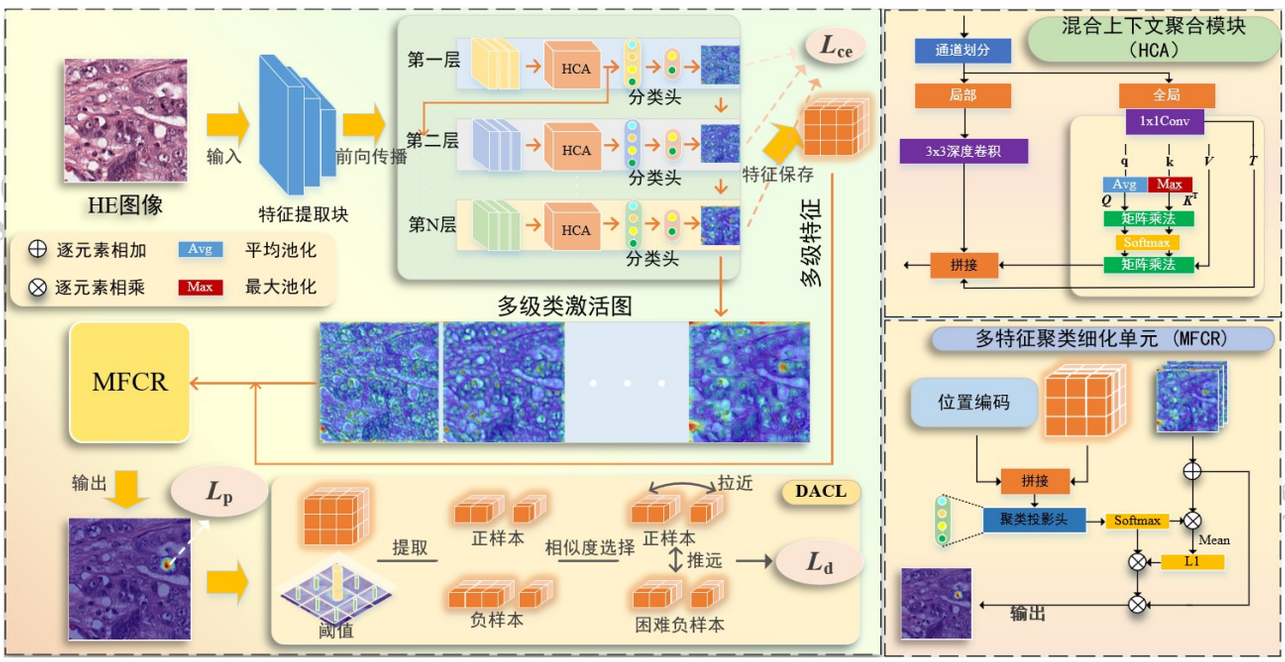

图2 所提方法整体框架图。

Fig. 2 Overall framework of the proposed method

2.1 混合上下文聚合模块

在有丝分裂核检测中，目标尺寸小、形态变化大，且易与淋巴细胞、凋亡细胞及深染细胞核等困难负样本混淆。因此，模型既需要保留细胞核边缘、染色强度和核膜形态等局部细节，也需要利用周围组织结构、细胞密度和邻域上下文等全局信息。传统局部-全局融合方法通常采用多尺度卷积并联、全局池化拼接或固定权重注意力重标定，融合方式相对静态，难以根据不同病理区域的形态复杂度自适应调整。

为此，本文设计了混合上下文聚合模块（hybrid context aggregation，HCA）。如图2右上角所示，HCA嵌入baseline编码器的各个特征层中，以该层输出特征作为输入，经局部形态分支和上下文建模分支聚合后生成融合特征，并将其作为下一编码层的输入。

首先，通道选择器利用通道注意力机制计算每个通道的权重$\boldsymbol{W}$。随后，基于重要性权重$\boldsymbol{W}$，通道选择器为每个通道分配一个决策分数$\boldsymbol{S}$，如公式（1）所示。式中，*LayerNorm*表示层归一化操作，Conv表示卷积操作。决策分数$\boldsymbol{S}$用于生成全局掩码$\boldsymbol{M}_g$和局部掩码$\boldsymbol{M}_l$，如公式（2）所示。式中，下标g和l分别表示全局和局部分支；θ为通道划分阈值。基于掩码的分配，本文将通道划分为全局通道和局部通道，它们

© 中国图象图形学报版权所有

分别用于捕获全局信息和局部细节。

$$\boldsymbol{S} = LayerNorm(Conv(\boldsymbol{W})) \tag{1}$$

$$\boldsymbol{M}_{\mathrm{g}} = I(\boldsymbol{S} \geqslant \theta), \boldsymbol{M}_{l} = 1 - \boldsymbol{M}_{\mathrm{g}} \tag{2}$$

在完成通道划分之后，局部与全局分支将分别对通道进行独立处理。对于局部通道，本文首先应用深度卷积来提取局部特征，如公式(3)所示。式中，$DWConv$表示深度卷积操作。该深度卷积操作仅独立作用于每个通道的特征图，而不改变通道的数量。

$$\boldsymbol{X}_{\mathrm{l}}' = DWConv(\boldsymbol{X}_{\mathrm{l}}) \tag{3}$$

对于全局通道，本文首先利用卷积层结合重塑(reshape)操作将其映射为多个头。随后，通过从这多个头中进行选择，提取出查询($\boldsymbol{Q}$)、键($\boldsymbol{K}$)和值($\boldsymbol{V}$)。整个过程如公式(4)和(5)所示，$\boldsymbol{X}_{\mathrm{h}} \in \mathbf{R}^{B \times n_{\mathrm{h}} \times d \times H \times W}$且$\boldsymbol{X}_{\mathrm{g}} \in \mathbf{R}^{B \times (C/2) \times H \times W}$，$B$为批量大小；$n_{\mathrm{h}}$为特征组数；$d$为每组通道数；$C$为模块输入通道数；$H$和$W$分别为特征图高度和宽度；下标h表示多头分组；$i$为组索引；$\boldsymbol{T}$表示保留特征。

$$\boldsymbol{X}_{\mathrm{h}} = reshape(Conv(\boldsymbol{X}_{\mathrm{g}})) \tag{4}$$

$$\begin{aligned} \boldsymbol{Q} &= \sum_{i=1}^{n_{\mathrm{h}}-3} \boldsymbol{X}_{\mathrm{h},i}, \quad \boldsymbol{K} = \boldsymbol{X}_{\mathrm{h},n_{\mathrm{h}}-2}, \\ \boldsymbol{V} &= flatten(\boldsymbol{X}_{\mathrm{h},n_{\mathrm{h}}-1}), \quad \boldsymbol{T} = \boldsymbol{X}_{\mathrm{h},n_{\mathrm{h}}} \end{aligned} \tag{5}$$

接着，通过$\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}$计算查询与键之间的相似度，随后执行Softmax操作以获得注意力权重$\boldsymbol{G}$。计算过程如公式(6)所示：

$$\boldsymbol{G} = Softmax\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_{\mathrm{k}}}}\right) \tag{6}$$

式中，d_{k}是键$\boldsymbol{K}$的维度。随后，获得的注意力权重$\boldsymbol{G}$被用于对值$\boldsymbol{V}$进行加权，从而得到聚合后的全局特征表示，如公式(7)所示：

$$\boldsymbol{X}_{\mathrm{g}}' = \mathbf{G}\mathbf{V} \tag{7}$$

最后，HCA模块将局部特征、全局特征以及保留特征$\boldsymbol{T}$进行拼接，从而获得最终的输出特征图。最终输出如公式(8)所示：

$$\boldsymbol{X}' = concat(\boldsymbol{X}_{\mathrm{l}}', \boldsymbol{X}_{\mathrm{g}}', \boldsymbol{T}) \tag{8}$$

与常见局部-全局融合模块相比，HCA并非简单拼接多尺度特征或采用固定权重加和，而是通过通道选择机制进行内容相关的分支分配，使局部形态建模和全局上下文建模能够随输入区域自适应调整。该设计有助于在保留有丝分裂核细粒度形态特征的同时，引入邻域组织上下文以减少相似细胞核造成的误检，并为后续的细化提供更可靠的特征基础。

2.2 多特征聚类细化单元

传统类激活图(CAM)通常只能突出最具判别性的局部区域，容易出现响应破碎、背景误激活和目标中心偏移等问题。为此，本文设计了多特征聚类细化模块(MFCR)。MFCR以融合CAM和空间引导特征为输入。该模块利用多层级CAM和浅层空间特征生成像素级软区域分配，并根据各区域内的CAM响应强度估计其前景相关性，从而对原始CAM进行聚类感知细化。细化后的CAM进一步用于目标定位和困难样本挖掘。

从网络不同阶段提取的类激活图蕴含着高度互补的信息：浅层特征能够保留至关重要的细粒度空间细节，而深层特征则提供了更强的语义线索。为了充分整合这两种优势信息，本文对多层级的类激活图进行了融合，方法如公式$\boldsymbol{M} = \frac{1}{N-1}(\boldsymbol{M}_1 + \cdots + \boldsymbol{M}_{N-1}) + \boldsymbol{M}_N$所示。式中，$\boldsymbol{M}$为融合后的类激活图；$N$为参与融合的特征层数；$\boldsymbol{M}_i$为第$i$层经空间尺寸对齐后的类激活图。在此融合机制中，浅层特征主要致力于提供高精度的局部定位，而深层特征则负责确保全局的语义一致性。这种多尺度特征的结合，有效地弥补了单一网络层感受野的局限性，为后续的精细化处理奠定了基础。

为了进一步对融合后的类激活图进行精细化处理，本文引入了多层级特征结合位置编码的策略，以指导空间软聚类过程。给定引导特征图$\boldsymbol{F} \in \mathbf{R}^{B \times C \times H \times W}$。本文首先应用$1 \times 1$卷积对其进行降维操作，随后将其与二维位置编码进行拼接，得到多层级融合特征$\boldsymbol{F}_{\mathrm{a}}$。

随后，利用卷积操作(Conv)生成像素级的聚类逻辑值，并沿聚类维度执行Softmax操作，从而获得软聚类分配$\boldsymbol{A}$，如公式(9)所示。

$$\boldsymbol{A} = Softmax(Conv(\boldsymbol{F}_{\mathrm{a}})) \tag{9}$$

该分配图并不直接表示前景或背景类别，而是将图像划分为若干具有相似空间和外观属性的软区域。随后，MFCR根据每个软区域内的CAM平均响应计算区域权重，用于衡量该区域与有丝分裂核前景的相关程度。为了消除由聚类面积差异引起的系统性偏差，本文将聚类与前景的关联度重新定义为各聚类内部CAM的加权平均强度。具体而言，本文

© 中国图象图形学报版权所有

计算加权得分s_k,随后应用L1归一化以获得聚类权重w_k。如公式(10)所示。式中,s_k为第k个聚类的CAM加权平均响应;H和W分别为特征图高度和宽度;u和v分别为行、列索引;$A_{k,u,v}$为位置(u,v)属于第k个聚类的软分配值;$M_{u,v}$为融合类激活图在该位置的响应值;w_k为归一化后的聚类权重;K为聚类数;j为聚类索引,s_j为第j个聚类的加权平均响应。

$$s_k=\frac{\sum_{u=1}^{H}\sum_{v=1}^{W}A_{k,u,v}M_{u,v}}{\sum_{u=1}^{H}\sum_{v=1}^{W}A_{k,u,v}},\quad w_k=\frac{s_k}{\sum_{j=1}^{K}s_j} \tag{10}$$

最后,通过聚类加权融合对CAM进行细化,如公式(11)所示。式中,$\boldsymbol{M}_{\mathrm{r}}$为细化后的类激活图;$\odot$表示逐元素相乘。

$$\boldsymbol{M}_{\mathrm{r}}=\sum_{k=1}^{K}w_k\boldsymbol{A}_k\odot\boldsymbol{M} \tag{11}$$

在此过程中,像素级的聚类分配与聚类级别的重要性共同指导了CAM的重建。这种公式化的方法实现了对聚类感知前景语义的显式建模,并生成了更为准确且具有结构一致性的激活图。

在经过聚类加权融合之后,本文应用最大最小归一化,如公式(12)所示,将值域重新缩放至[0,1]范围内,以保持不同样本间后续阈值处理的一致性。

$$\boldsymbol{M}_{\mathrm{n}}=\frac{\boldsymbol{M}_{\mathrm{r}}-m_{\min}}{m_{\max}-m_{\min}} \tag{12}$$

式中,$m_{\min}$和$m_{\max}$分别表示$\boldsymbol{M}_{\mathrm{r}}$中的最小和最大激活值;$\boldsymbol{M}_{\mathrm{n}}$为归一化后的类激活图。

为了在保留细胞核形态的同时过滤掉困难负样本,本文引入了一种协同约束机制,该机制结合了多特征聚类细化(MFCR)、特征归一化以及定位损失L_{p}。仅仅依赖MFCR会缺乏空间边界约束,从而带来将附近困难负样本包含在内的风险;而如果仅优化L_p,则会导致激活区域收缩至单一像素。这些组件共同构成了一个几何驱动的排除机制。具体来说,如果一个困难负样本被错误地归入前景中,它会使得预测的质心$\boldsymbol{C}_{\mathrm{p}}$偏离真实的质心$\boldsymbol{C}_{\mathrm{g}}$。为了对这种偏差进行惩罚,本文构建了一种基于空间期望的定位约束机制。该方法保证了在反向传播过程中的可微性。$\boldsymbol{C}_{\mathrm{p}}$的空间坐标是直接在归一化后的类激活图上计算得出的,如公式(13)所示:

$$\begin{cases}x_{\mathrm{p}}=\dfrac{\sum_{u=1}^{H}\sum_{v=1}^{W}vM_{\mathrm{n},u,v}}{\sum_{u=1}^{H}\sum_{v=1}^{W}M_{\mathrm{n},u,v}}\\ y_{\mathrm{p}}=\dfrac{\sum_{u=1}^{H}\sum_{v=1}^{W}uM_{\mathrm{n},u,v}}{\sum_{u=1}^{H}\sum_{v=1}^{W}M_{\mathrm{n},u,v}}\end{cases} \tag{13}$$

式中,x_{p}和y_{p}分别表示预测质心的横、纵坐标;H和W分别表示为激活图的高度和宽度;u和v分别为行、列索引;$M_{\mathrm{n},u,v}$为激活图在(u,v)处的响应值。最终的定位损失被定义为预测的空间期望中心与病理学家标注的真实质心之间的欧氏距离,如公式(14)所示。式中,L_{p}为定位损失。

$$\begin{aligned}L_{\mathrm{p}}&=\|\boldsymbol{C}_{\mathrm{p}}-\boldsymbol{C}_{\mathrm{g}}\|_2\\&=\sqrt{(x_{\mathrm{p}}-x_{\mathrm{g}})^2+(y_{\mathrm{p}}-y_{\mathrm{g}})^2}\end{aligned} \tag{14}$$

2.3 难度感知对比学习方法

在深度学习任务中,特别是在细粒度分类与弱监督检测领域,有效地选择正负样本以构建特征表示,对于增强模型的判别能力至关重要。传统的对比损失往往会被海量易于区分的负样本所主导。这种主导作用稀释了来自困难样本的梯度,从而限制了模型对具有视觉相似性的细粒度特征的区分能力。为了应对这一问题,本文提出了一种难度感知对比学习(DACL)方法。DACL以MFCR的细化响应构建像素级伪标签,再据此组织正负样本关系。该方法将伪标签引导的采样与严格的困难负样本挖掘机制相结合,以显式地提升网络在特征空间中对困难样本的识别能力。首先,本文将由MFCR模块生成的高质量细化类激活图$\boldsymbol{M}_{\mathrm{n}}$进行二值化处理,从而生成指示潜在目标区域的伪标签。基于这些伪标签,本文为特征像素分配正极性和负极性标签。针对每一个锚点(anchor)特征,本文严格地筛选出其对应的困难负样本,并利用它们来构建难度感知监督对比损失。

给定一个批次的特征向量$\{\boldsymbol{f}_t\in\mathbf{R}^C\}_{t=1}^{N}$及其对应的伪标签$\{y_t\}_{t=1}^{N}$,本文首先计算每个锚点i与所有其他样本j之间的成对相似度,如公式(15)所示:

$$s_{ij}=\frac{\boldsymbol{f}_i^{\mathrm{T}}\boldsymbol{f}_j}{\tau} \tag{15}$$

式中,τ是用于缩放相似度的温度系数。

对于每个锚点,本文将正样本集$\boldsymbol{P}_i$定义为与该锚点共享相同标签的所有样本(不包括其自身),相

© 中国图象图形学报版权所有

应地可以得出标准的普通负样本集$\boldsymbol{N}_i$定义。如公式(16)所示：

$$\boldsymbol{P}_i=\{j\neq i\,|\,y_j=y_i\},\boldsymbol{N}_i=\{j\neq i\,|\,y_j\neq y_i\} \quad (16)$$

为了严格强调难以区分的负样本，本文在负样本集内执行困难负样本挖掘。设$|\boldsymbol{N}_i|=n_i$表示锚点i的负样本总数，并令$r\in(0,1]$为预定义的困难负样本比例。按比例r确定保留数量k_i，当负样本集合非空时，至少保留1个。

随后，提取与锚点相似度最高的k_i个负样本以构成困难负样本集$\boldsymbol{H}_i$。对于每个锚点，本文构建一个有效的、受限的对比集，该集合仅由所有正样本和挖掘出的困难负样本组成，整个过程分别如公式(17)和公式(18)所示。

$$\boldsymbol{H}_i=TopK_{j\in\boldsymbol{N}_i}(s_{i,j},k_i) \quad (17)$$

$$\boldsymbol{D}_i=\boldsymbol{P}_i\cup\boldsymbol{H}_i \quad (18)$$

式中，$\boldsymbol{H}_i$为锚点的困难负样本索引集；$TopK$表示从负样本集合$\boldsymbol{N}_i$中选取相似度$s_{i,j}$最高的k_i个样本，并返回其索引；$\mathbf{D}_i$为由正样本集合$\boldsymbol{P}_i$与困难负样本集合$\boldsymbol{H}_i$合并得到的对比集。在这个受限集合上，本文通过对相似度执行Softmax操作来定义归一化概率，如公式(19)所示。

$$p_{i,j}=\frac{\exp(s_{i,j})}{\sum_{k\in\boldsymbol{D}_i}\exp(s_{i,k})},\quad j\in\boldsymbol{D}_i \quad (19)$$

式中，$p_{i,j}$为样本j在锚点i的对比集内的归一化概率；exp表示指数函数；k为对比集$\boldsymbol{D}_i$中的样本索引。

锚点对损失的贡献由其正样本的平均负对数概率给出，如公式(20)所示：

$$\ell_i=-\frac{\tau}{\tau_0}\cdot\frac{1}{|\boldsymbol{P}_i|}\sum_{j\in\boldsymbol{P}_i}\log p_{i,j} \quad (20)$$

式中，ℓ_i为锚点i的对比损失；τ为温度系数；τ_0是用于归一化的基础温度系数；$|\boldsymbol{P}_i|$为锚点i的正样本数量；log表示自然对数函数。最后，整个批次的总损失计算如公式(21)所示：

$$L_{\mathrm{d}}=\frac{1}{N'}\sum_{i\in\boldsymbol{I}}\ell_i \quad (21)$$

式中，L_d为批次对比损失，下标d表示对比学习项；$\boldsymbol{I}$表示至少拥有一个正样本(即$|P_i|>0$)的锚点集合；N'为有效锚点数量且$N'=|\boldsymbol{I}|$。

为了确保特征空间优化的稳定性，本文采用了一种受课程学习启发的预热(warm-up)策略。在训练的初始阶段，伪标签不可避免地包含噪声。过早地强制执行对比学习会排斥被错误分配的负样本，从而导致严重的确认偏差和特征空间崩溃。因此，DACL仅在预定义的预热期之后才会被激活，从而允许网络在执行度量学习之前先建立起可靠的初始定位。与标准的监督对比损失相比，本文的DACL损失保留了对正样本的拉近效应，同时通过$\boldsymbol{H}_i$集合将推开效应显式地集中在一小部分极具欺骗性的困难负样本上。通过将分母限制为$\boldsymbol{D}_i$，DACL有效地防止了简单负样本主导梯度，从而深度增强了特征空间中的类间可分性，并提升了模型在细粒度和易混淆样本上的识别性能。

2.4 损失函数

本文所提模型的总损失函数由难度感知对比损失L_d、交叉熵分类损失L_{ce}和定位损失L_p联合构成，具体如公式(22)所示：

$$L=\alpha L_{\mathrm{d}}+\beta L_{\mathrm{ce}}+\gamma L_{\mathrm{p}} \quad (22)$$

经过实际网络训练验证，本文将各项损失的超参数分别设定为α=0.03、β=1以及γ=0.01。因为各损失函数值域尺度不同，从而导致超参数设定差异较大。

3 实　验

3.1 实验数据与实验设置

3.1.1 数据集

实验采用TUPAC16(Veta等，2019)、MIDOG22(Aubreville等，2024)及作者团队构建并公开发布的GZMH-V2数据集。GZMH-V2基于GZMH数据集(Wang等，2024b)，一共包含2322张HPF切片，总计5291个有丝分裂细胞核标注。

3.1.2 实验设置

实验基于PyTorch，采用ResNet-18骨干网络，运行于配备NVIDIA RTX 5080 GPU和128 GB内存的工作站。DACL在预热后启用，以降低训练初期CAM噪声的影响。

所有图像转换为RGB格式，并裁剪或缩放为500×500像素。训练采用随机翻转、旋转、颜色扰动和归一化处理。以标注质心为中心构建正样本区域，距任一标注质心超过匹配半径的候选区域作为负样本。预测中心距标注中心小于30像素时判为真正例，每个标注最多匹配一个预测。所有模型采

© 中国图象图形学报版权所有

用相同的数据划分。

训练联合优化分类、定位和DACL损失，保留阴性patch及仅含单个标注点的阳性patch，剔除多目标阳性patch，以避免单质心监督歧义，但可能增加密集目标的漏检风险。推理不限制目标数量，概率图阈值化后，各连通区域质心分别作为预测点。

3.2 实验结果与可视化分析

3.2.1 定量对比结果

如表1所示，所提方法在所列方法中取得最高F1值0.773，精确率和召回率分别为0.790和0.756。相较UTS，召回率和F1值分别提高0.045和0.023。未公开且无法按原协议复现的指标不予补填。与UTS和Sohail等人等两阶段模型相比，本方法在保持高精确率（Precision）的同时提升了召回率（Recall），从而降低了漏检率。得益于本方法对困难负样本进行的细粒度判别，其取得了第二高的精确率。尽管UTS同样整合了困难样本过滤机制并获得了最高的精确率，但与之相比，本方法在召回率和F1分数上分别提升了6.3%和3.1%。

表1 TUPAC16数据集上不同方法的定量对比结果

Table 1 Quantitative comparison of different methods on the TUPAC16 dataset

模型	技术类型	精确率	召回率	F1值
U-Net（Ronneberger等，2015）	End-to-end segmentation	0.649	0.707	0.677
Zerhouni等人（2017）	Single CNN	0.675	0.623	0.648
Paeng等人（2017）	Multi stages CNN	-	-	0.652
Lafarge等人（2017）	Domain-adversarial neural networks	-	-	0.620
Tellez等人（2018）	Multi CNN	-	-	0.480
Akram等人（2018）	Multi stages CNN	0.613	0.671	0.640
SegMitos（Li等，2019）	End-to-end segmentation	0.640	0.700	0.669
TL-Mit-Seg（Wahab等，2019）	Two stages	0.770	0.660	0.713
DeepConsensus（Wollmann和Rohr，2021）	End-to-end detection	-	-	0.470
Sohail等人（2021）	Two stages	0.710	**0.760**	0.750
Nateghi R等人（2021）	Multi stages CNN	0.764	0.714	0.738
FMDet（Wang等，2023）	End-to-end segmentation	0.713	0.735	0.724
UTS（Fernandez-Martín等，2024）	Two stages	**0.795**	0.711	0.750
YOLOV11（Khanam和Hussain，2024）	End-to-end detection	0.715	0.723	0.719
FoCasNet（Wang等，2024a）	Two stages	0.687	0.724	0.705
OURS	**End-to-end detection**	<u>0.790</u>	<u>0.756</u>	**0.773**

注：最优结果以粗体标注；次优结果以下划线标注；“-”表示原文未报告该项指标。

GZMH-V2数据集来自于医院真实的临床数据。该数据集比现有的公开数据集更具挑战性，具有两个不同的测试集（Test1和Test2）。表2和表3展示了在该数据集的不同测试集上不同方法的对比结果。与几种可复现的基准模型相比，所提方法在两个测试集上均取得了最佳的综合表现。

在Test1测试集上，所提方法取得了0.578的F1分数，超过了排名第二的方法UTS（Fernandez-Martín等，2024）达5.5%，同时保持了平衡的精确率和召回率，分别为0.583和0.573。

在Test2测试集上，该方法再次获得最佳结果，其中，F1分数达到0.558，超过了排名第二的FMDet（Wang等，2023）达3.7%。

如表4所示，提出的方法在MIDOG22数据集上取得了0.693的F1分数，仅略低于UTS（Fernandez-Martín等，2024）（0.696），但优于FMDet（Wang等，2023）、FoCasNet（Wang等，2024a）和YOLOV11（Khanam和Hussain，2024）。尽管本方法的表现与UTS

© 中国图象图形学报版权所有

表2　GZMH-V2数据集(Test1)上不同方法的对比结果

Table 2　Comparison of different methods on GZMH-V2 Test1

模型	精确率	召回率	F1值
U-Net(Ronneberger等,2015)	0.462	0.447	0.454
FMDet(Wang等,2023)	0.563	0.496	0.527
UTS(Fernandez-Martín等,2024)	**0.584**	0.516	0.548
YOLOV11(Khanam和Hussain,2024)	0.519	0.529	0.524
FoCasNet(Wang等,2024a)	0.513	0.499	0.506
OURS	0.583	**0.573**	**0.578**

注:最优结果以粗体标注;次优结果以下划线标注。

表3　GZMH-V2数据集(Test2)上不同方法的对比结果

Table 3　Comparison of different methods on GZMH-V2 Test2

模型	精确率	召回率	F1值
U-Net(Ronneberger等,2015)	0.476	0.483	0.479
FMDet(Wang等,2023)	0.515	**0.564**	0.538
UTS(Fernandez-Martín等,2024)	0.528	0.521	0.525
YOLOV11(Khanam和Hussain,2024)	0.551	0.510	0.530
FoCasNet(Wang等,2024a)	0.497	0.518	0.507
OURS	**0.566**	0.550	**0.558**

注:最优结果以粗体标注;次优结果以下划线标注。

非常接近,但值得注意的是,UTS采用的是两阶段策略,而本方法是完全端到端的架构。此外,相对较高的精确率也表明,本方法能够产生更稳定的置信度预测,并在抑制误报方面表现出更强的能力。

综上所述,该方法在三个复杂病理数据集上均取得了领先或极具竞争力的综合表现,充分证明了其在有丝分裂核检测任务中卓越的能力与潜力。

表4　MIDOG22数据集上不同方法的定量对比结果

Table 4　Comparison on the MIDOG22 dataset

模型	精确率	召回率	F1值
U-Net(Ronneberger等,2015)	0.650	0.655	0.652
FMDet(Wang等,2023)	0.682	**0.685**	0.683
UTS(Fernandez-Martín等,2024)	**0.710**	0.682	**0.696**
YOLOV11(Khanam和Hussain,2024)	0.676	0.680	0.678
FoCasNet(Wang等,2024a)	0.650	0.680	0.665
OURS	0.704	0.682	0.693

注:最优结果以粗体标注;次优结果以下划线标注。

3.2.2　定性可视化分析

有丝分裂核在组织学上的形态极其多变,且常与凋亡细胞、极度挤压变形的非分裂核等在视觉上高度相似,这构成了该检测任务在特征提取上的核心难点。

图3呈现了不同方法在三个数据集上的可视化对比结果,其中绿色圆圈代表真阳性(TP),黄色代表假阳性(FP),蓝色代表假阴性(FN)。如图所示,本文模型在全部三个数据集上均稳定地展现出更优的视觉检测效果。在TUPAC16和MIDOG22数据集上,本方法实现了更高的精确率,且未出现任何漏检现象。

尽管UTS(Fernandez-Martín等,2024)在TUPAC16和MIDOG22数据集上取得了与本方法相媲美的可视化结果,但本方法在更具挑战性的GZMH-V2数据集上展现出了显著的优势。

综上所述,本方法在多个数据集上均取得了显著的性能提升。这些实验结果为本方法在有丝分裂核检测任务中的有效性提供了强有力的证据。

3.3　消融实验

为了验证所提各组件的有效性,本文在TUPAC16数据集上进行了消融实验。基线模型(Base)采用了配备定位损失的ResNet-18网络。定量结果如表5所示。

模型A将F1分数从0.697提升至0.727,获得了4.3%的性能增益。HCA模块通过动态通道划分与多头注意力机制,有效地整合了局部与全局信息,增强了模型捕获细胞核上下文特征的能力,从而有效减少了因特征表征不完整而导致的漏检。

模型B的F1分数达到了0.728(提升4.4%)。通过执行多特征软聚类来细化类激活图(CAM),MFCR提升了定位精度,并减少了由粗略CAM区域引发的假阳性(误检)。

模型C取得了0.717的F1分数,提升了2.8%。通过挖掘困难负样本,DACL机制强化了模型区分疑难样本(例如视觉上与有丝分裂核高度相似的淋巴细胞)的能力。然而,其性能增益略低于前两个模块,这表明困难样本的学习高度依赖于高质量的初始特征表征。

模型D的F1分数达到0.766。HCA提供的丰富

© 中国图象图形学报版权所有

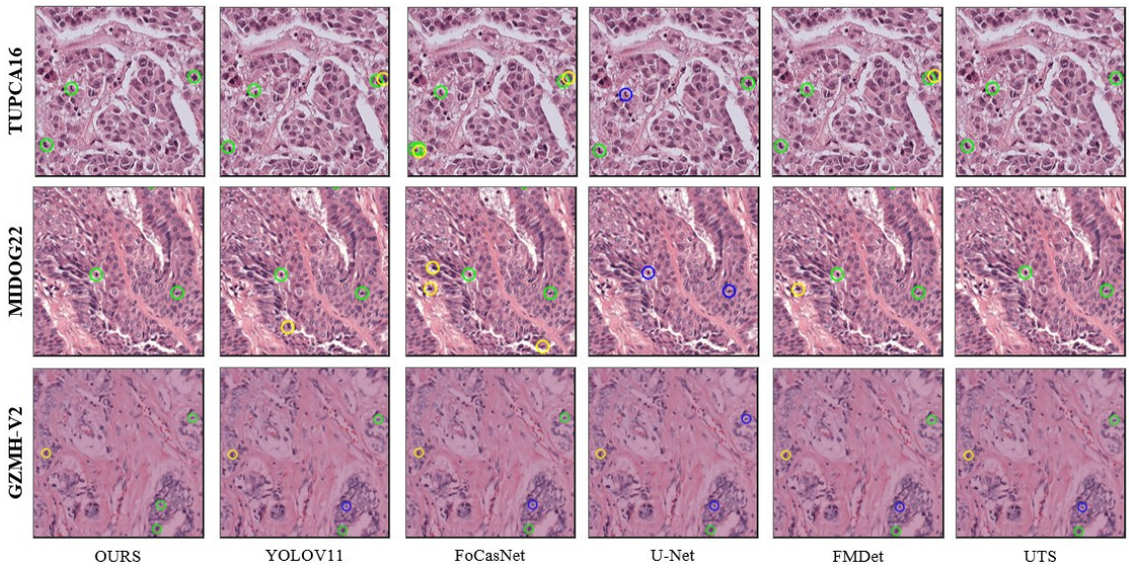

图3　不同方法在3个数据集上的可视化结果对比

Fig. 3　Visual comparison of different methods on three datasets

表5　TUPAC16数据集上逐步消融实验对比结果

Table 5　Stepwise ablation on TUPAC16

模型	HCA	MFCR	DACL	精确率	召回率	F1值
Base	—	—	—	0.685	0.710	0.697
A	√	—	—	0.715	0.740	0.727
B	—	√	—	0.740	0.717	0.728
C	—	—	√	0.734	0.701	0.717
D	√	√	—	0.773	**0.760**	0.766
E	√	—	√	0.752	0.733	0.742
F	—	√	√	0.781	0.717	0.748
G	√	√	√	**0.790**	0.756	**0.773**

注："√"表示启用该模块，"—"表示未启用；加粗字体表示本表各指标的最优值。

上下文特征为MFCR细化CAM提供了更为可靠的基础，两者的结合在特征表征与定位精度上产生了显著的协同增强效应。

模型E获得了0.742的F1分数。得益于HCA提取的细粒度特征，DACL能够更有效地挖掘困难负样本，从而进一步增强了对假阳性的抑制能力。

模型F的F1分数达到0.748，其中精确率（Precision）攀升至0.781，但召回率（Recall）回落至0.717。这表明，尽管MFCR为DACL提供了更精准的正负样本区域，但由于缺乏HCA提供的全局特征支持，模型依然会出现对低对比度细胞核的漏检。

当集成所有三个模块（模型G）时，模型在所有配置中取得了最佳的综合性能。

总结而言，表5中的消融实验清晰地证明了这三个核心模块的必要性与协同价值。HCA奠定了坚实的特征提取基石，而MFCR和DACL则分别从精准定位与细粒度判别两个关键维度提升了网络性能。这三个模块相辅相成，共同构成了所提模型的最佳网络架构

3.3.1　HCA消融

为了进一步验证HCA模块的有效性，开展内部消融实验，以深入探究其各项设计选择的合理性。具体而言，评估了混合通道划分策略以及局部与全局信息的联合聚合机制是否对模型性能产生了正向增益，并系统分析了多头机制与池化选择策略对模型表征能力的具体影响。

为了评估动态划分局部与全局通道的实际贡献，将本方法与固定划分策略（前一半通道分配给局部特征分支，剩余半数分配给全局特征分支）进行了对比。该组实验以表5中的模型A为基准进行。实验结果表明，动态通道划分机制展现出更优的性能表现，从而充分证实了该设计的有效性。

此外，在模型A的基础上进一步分析了多头注意力模块中注意力头数对网络性能的具体影响。

本方法采用了不同的池化策略来获取$\boldsymbol{Q}$（查询）与$\boldsymbol{K}$（键）的特征表征。最大池化Max侧重于强化最显著的局部特征响应，而平均池化Avg则能提供更为稳定且平滑的全局特征聚合。为了评估所选池化

© 中国图象图形学报版权所有

表6　通道选择策略消融

Table 6　Ablation of channel selection strategies

方法	精确率	召回率	F1值
不使用通道选择策略	0.713	0.737	0.725
使用通道选择策略	**0.715**	**0.740**	**0.727**

注：最优结果以粗体标注。

表7　注意力头数消融

Table 7　Ablation of the number of attention heads

数量	精确率	召回率	F1值
4	**0.715**	0.740	**0.727**
8	0.707	0.744	0.725
16	0.702	**0.747**	0.724

注：最优结果以粗体标注。

组合策略的有效性，本文在模型G的基础上开展了对比实验。实验结果证实了当前池化设计的可行性与有效性。

表8　池化选择策略消融

Table 8　Ablation of pooling strategies

方法		精确率	召回率	F1值
Q	*K*			
Avg	Avg	0.782	0.750	0.766
Avg	Max	**0.790**	**0.756**	**0.773**
Max	Avg	0.787	0.749	0.768
Max	Max	0.776	**0.756**	0.766

注：最优结果以粗体标注。

3.3.2　MFCR消融

为了评估多特征聚类细化（MFCR）模块的实际贡献及其设计的合理性，在模型B的基础上开展了消融实验，本节重点探讨MFCR中特征来源的选择。

MFCR既可以提取骨干网络单一层的特征，也可以聚合多层特征。浅层特征能够提供细粒度的空间细节，有助于定位细胞核边界和局部纹理；而深层特征则蕴含高阶语义信息。有助于区分有丝分裂核与相似背景结构。为了评估多层特征融合带来的增益，本文对比了三种变体配置：仅使用浅层特征（shallow only）、仅使用深层特征（deep only）以及本文提出的多层融合（multi-level fusion）策略。

如表9所示，多层级融合取得最高F1值。仅使用浅层特征时，模型能够保留较好的空间细节，但语义判别能力不足；仅使用深层特征时，模型具备更强语义表达，但定位响应相对粗糙。多层级融合同时利用浅层空间细节和深层语义信息，能够为CAM细化提供更稳定的软区域分配，从而提升整体检测性能。

表9　聚类特征组合策略消融

Table 9　Ablation of clustering feature combinations

方法	精确率	召回率	F1值
只用浅层	0.723	0.713	0.718
只用深层	0.711	**0.732**	0.721
深浅融合	**0.740**	0.717	**0.728**

注：最优结果以粗体标注。

此外，图5（c）、（d）分别给出了以完整模型G为参照的位置编码与聚类数对照。其中，位置编码通过同维度全零替换进行比较；聚类数设置为1、2、4和8，K=2对应原配置，K=1作为关闭聚类空间重加权的退化对照。

3.3.3　DACL消融

针对DACL模块，本文以模型G为基线，对预热轮数（warm-up epochs）的超参数选择开展了消融实验。实验结果表明，在第15个训练轮次引入DACL能够取得最佳的综合性能。过早引入该模块往往会导致模型表现次优，其主要原因在于网络在训练初期生成的类激活图定位尚欠准确，极易引入显著的噪声干扰，从而阻碍对比学习的有效优化。反之，若过晚引入该机制，则因网络特征已趋于定型，仅能获得极为有限的性能增益。

表10　预热轮数选择消融

Table 10　Ablation of warm-up epochs

轮数	精确率	召回率	F1值
5	0.740	0.740	0.740
10	0.786	0.752	0.769
15	**0.790**	0.756	**0.773**
20	0.778	0.757	0.767
25	0.752	**0.760**	0.756

注：最优结果以粗体标注。

针对本方法中的对比学习组件，该方法还对比了两种经典的对比学习损失函数，以进一步验证所提方法的有效性。

© 中国图象图形学报版权所有

本文进一步以完整模型G为参照，分析负样本保留比例与采样策略的影响。图5(a)给出了不同比例下最高相似度选择与随机选择的对照，两种策略在保留比例为1.00时共享全部候选负样本。图5(b)固定保留比例为0.50，比较两种策略的精确率、召回率和F1值，用于区分负样本数量变化与困难负样本选择机制的影响。

3.3.4　骨干网络消融实验与T-SNE可视化

此外，为了评估不同骨干网络(backbone networks)对模型性能的具体影响，本文对该网络的骨干网络选择开展相关实验。实验结果表明，当采用ResNet-18作为特征提取骨干时，模型取得了最佳的综合性能表现。为了进一步探究特征表征的有效性，本文利用t-SNE技术对高维特征空间进行了可视化分析。如图4所示，基线模型(左图)提取的有丝分裂核特征(橙色)与背景混淆样本特征(灰色)之间存在严重的混叠现象。相比之下，本文提出的模型(右图)通过DACL模块对困难负样本施加显式的排斥力，从而构建了清晰的决策边界。

表11　对比学习损失函数对比

Table 11　Comparison of contrastive learning losses

损失	精确率	召回率	F1值
SupConLoss(Khosla等，2020)	0.784	**0.758**	0.771
InfoNCELoss(Oord等，2018)	0.784	0.756	0.770
OURS	**0.790**	0.756	**0.773**

注：最优结果以粗体标注。



图4　t-SNE可视化

Fig. 4　t-SNE visualization of feature distributions

表12　基线模型采用不同骨干网络的检测性能对比

Table 12　Detection performance of the baseline with different backbone networks

骨干网络	精确率	召回率	F1值
ResNet-18(He等，2016)	0.685	**0.710**	**0.697**
ResNet-50(He等，2016)	0.682	0.702	0.692
ResNet-101(He等，2016)	**0.687**	0.704	0.695
Swin-transformer(Liu等，2021)	0.678	0.695	0.686
VGG16(Simonyan和Zisserman，2014)	0.663	0.688	0.675

注：最优结果以粗体标注。

3.4　所提方法在临床应用中的一致性验证

该方法进一步在真实的临床场景中进行了评估，以验证其实际应用效能。临床评估邀请四位具有1～9年经验的病理医师，对来自GZMH-V2的27张高倍视野图像分别进行无辅助和模型辅助阅片，记录有丝分裂核计数及用时。首先进行无辅助阅片，间隔2周后进行模型辅助阅片，图像呈现顺序随机，记录各图像的有丝分裂核计数及阅片时间。模型辅助图像由原始图像叠加模型检测标记生成，评估流程如图6所示。

以HPF为评价对象，将每张HPF上4位医师计数的算术均值与参考计数作为两列数据，采用双向混合效应、一致性型、单次测量的组内相关系数(intraclass correlation coefficient，ICC)进行分析，记为ICC(C，1)。该指标评价医师平均计数与参考计数之间的一致性。

无辅助和模型辅助条件下，均以每张HPF上4位医师计数的算术均值与参考计数计算ICC(C，1)，评估结果详见表13和表14。在符号表示中，C代表一致性，1代表单测量ICC，K代表平均测量ICC。

无辅助和模型辅助条件下的ICC分别为0.224(95%置信区间：−0.164～0.552)和0.323(95%置信区间：−0.057～0.622)，绝对增加0.099。这一增益表明，该模型不仅提高了观察者间的一致性，还能帮助病理学家更可靠地逼近参考标准。

换言之，模型的预测结果可作为一种有效的决

© 中国图象图形学报版权所有

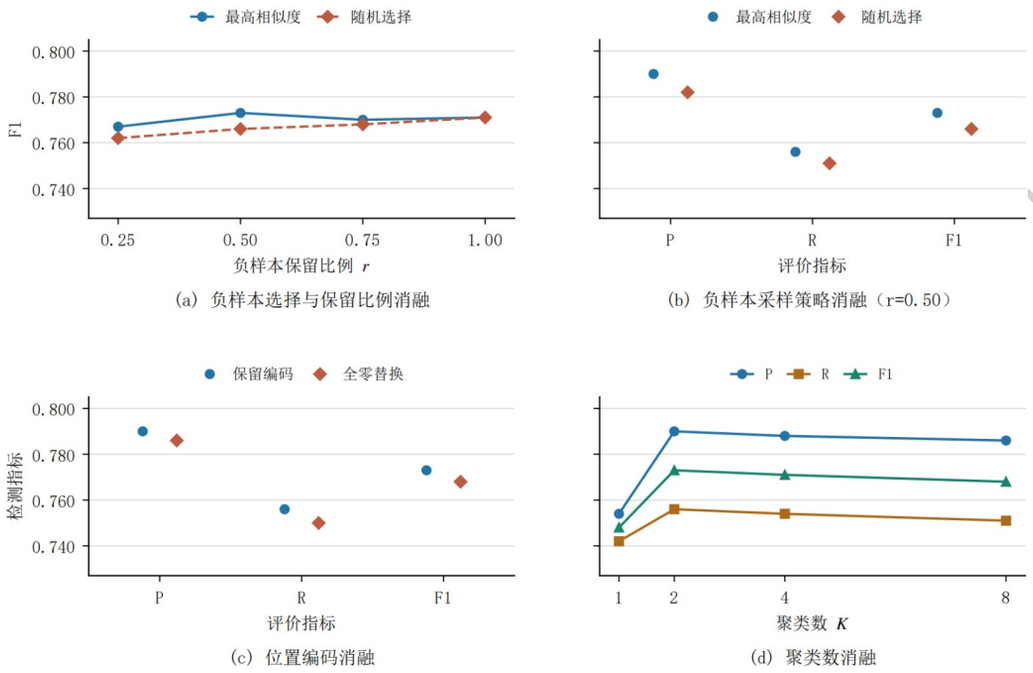

图5　TUPAC16数据集上的关键设置消融

Fig. 5　Ablation study of key settings on TUPAC16 ((a) negative-sample selection and retention ratio; (b) negative sampling strategy at r=0.50; (c) positional encoding; (d) number of clusters)

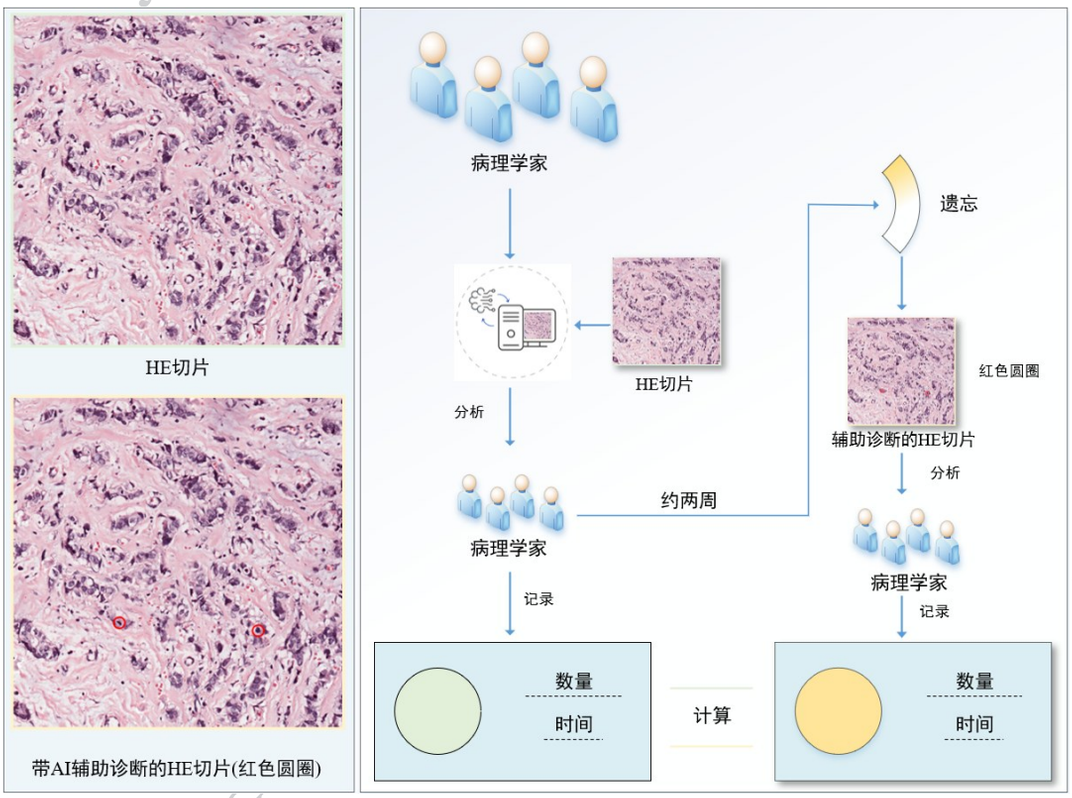

图6　所提方法在临床应用中的一致性验证流程

Fig. 6　Workflow for consistency validation in clinical application

策支持工具，减少了临界病例判读时的模棱两可，提升了真实临床环境中有丝分裂核评估的鲁棒性与可靠性。

所有参与评估的医师均具有1至9年的临床经

© 中国图象图形学报版权所有

表13　医师无辅助均值与真实标注对比

Table 13　Agreement between unaided pathologist measurements and reference annotations

一致性类型	ICC	95%CI
单测量ICC(C,1)	0.224	−0.164 ~ 0.552
平均测量ICC(C,k)	0.366	−0.391 ~ 0.711

表14　医师有辅助均值与真实标注对比

Table 14　Agreement between assisted pathologist measurements and reference annotations

一致性类型	ICC	95%CI
单测量ICC(C,1)	0.323	−0.057 ~ 0.622
平均测量ICC(C,k)	0.489	−0.122 ~ 0.767

验。为了进一步剖析本模型在辅助临床诊断方面的优越性，本文根据经验年限对医师进行了分层，以展开更为细致的分析。具体而言，具有6年及以下（≤6年）临床经验的医师被归为低年资医师，而具有6年以上（>6年）经验的医师被归为高年资医师。各年资组均先计算每张HPF上组内医师计数的算术均值，再与参考计数计算ICC（C，1），比较模型辅助前后的结果。

评估结果汇总于表15、表16、表17和表18中。低年资组ICC由0.279（95%置信区间：−0.106～0.591）增至0.450（95%置信区间：0.092～0.705）。

表15　低年资医师无辅助均值与真实标注对比

Table 15　Agreement between unaided less-experienced pathologists and reference annotations

一致性类型	ICC	95%CI
单测量ICC(C,1)	0.279	−0.106 ~ 0.591
平均测量ICC(C,k)	0.437	−0.236 ~ 0.743

表16　低年资医师有辅助均值与真实标注对比

Table 16　Agreement between assisted less-experienced pathologists and reference annotations

一致性类型	ICC	95%CI
单测量ICC(C,1)	0.450	0.092 ~ 0.705
平均测量ICC(C,k)	0.621	0.168 ~ 0.827

高年资组由0.122（95%置信区间：−0.264～0.474）增至0.167（95%置信区间：−0.220～0.509）。对比低年资与高年资医师的结果，观察到低年资医师的ICC提升了61.3%，而高年资医师则提升了36.9%。这表明，该模型对低年资医师的辅助效果更为显著，其临床评估过程更容易从模型支持中获益，这一结果也与临床实际预期高度一致。

表17　高年资医师无辅助均值与真实标注对比

Table 17　Agreement between unaided more-experienced pathologists and reference annotations

一致性类型	ICC	95%CI
单测量ICC(C,1)	0.122	−0.264 ~ 0.474
平均测量ICC(C,k)	0.217	−0.717 ~ 0.643

表18　高年资医师有辅助均值与真实标注对比

Table 18　Agreement between assisted more-experienced pathologists and reference annotations

一致性类型	ICC	95%CI
单测量ICC(C,1)	0.167	−0.220 ~ 0.509
平均测量ICC(C,k)	0.287	−0.565 ~ 0.675

不同医师判读细胞核所需的时间各不相同，而缩短这一评估时间对于提升临床诊断效率至关重要。阅片时间同样先在每张HPF内对4位医师取均值，再以27张HPF构成27个配对进行双侧配对t检验。无辅助和模型辅助条件下的平均用时分别为64.407±28.357秒和54.470±15.710秒，平均缩短9.937秒，时间差的95%置信区间为2.136～17.738秒，t值（paired t-test）为2.618，自由度为26，双侧p值约0.015。结果如表19所示。基于实验发现，可以得出以下结论：在本实验条件下，模型辅助阅片的平均用时较无辅助阅片缩短9.937秒，差异具有统计学意义（双侧配对t检验，p值约为0.015）。这表明，该模型在显著缩短临床诊断时间方面做出了重要贡献。

综上所述，该模型在同时提升病理学家的诊断准确率与时间效率方面，均展现出了积极的促进作用。

表19　辅助前后耗时差值数据

Table 19　Differences in assessment time before and after model assistance

均值差/s	标准差/s	95%置信区间/s	t值	p值
9.937	19.719	2.136 ~ 17.738	2.618	0.015

© 中国图象图形学报版权所有

4 结 论

本文提出了一种用于乳腺癌组织病理学图像有丝分裂核检测的难度感知对比学习与特征优化框架。该框架引入了混合上下文聚合(HCA)模块,该模块有效地融合了局部细节与全局语义,显著增强了模型的上下文理解与特征表征能力。该框架进一步设计了多特征聚类细化(MFCR)模块,通过软聚类与位置编码机制细化类激活图(CAM)区域,从而有效提升了定位精度与边界一致性。此外,该框架还开发了难度感知对比学习(DACL)机制,结合预热策略以有效缓解训练初期的噪声干扰,并通过像素级的正负样本选择与困难负样本挖掘,强化了模型区分模棱两可或视觉高度相似细胞核的判别能力。

所提方法在公开数据集TUPAC16、临床数据集GZMH-V2以及MIDOG22数据集上均取得了具有竞争力的检测性能,验证了其在有丝分裂核检测任务中的有效性。此外,临床验证表明,该方法提升了病理学家在实际阅片中的诊断一致性与时间效率。这些详实的实验结果突显了所提方法的临床实用价值,并展现了其作为一种先进、有效的乳腺癌病理诊断AI辅助解决方案的潜力。然而,本文方法仍存在一定局限性。在细胞密集、染色不均或组织结构复杂区域,CAM响应也可能出现偏移,导致定位不准。此外,MFCR依赖初始CAM质量,细化过程可能受噪声干扰。未来将结合跨中心数据和减少噪声干扰,进一步提升模型的泛化性与稳定性。

参考文献(References)

Akram S U, Qaiser T, Graham S, Kannala J, Heikkilä J and Rajpoot N. 2018. Leveraging unlabeled whole-slide-images for mitosis detection//Computational Pathology and Ophthalmic Medical Image Analysis. Cham: Springer: 69-77 [DOI: 10.1007/978-3-030-00949-6_9].

Alom M Z, Aspiras T, Taha T M, Bowen T J and Asari V K. 2020. MitosisNet: end-to-end mitotic cell detection by multi-task learning. IEEE Access, 8: 68695-68710 [DOI: 10.1109/ACCESS.2020.2983995].

Aubreville M, Stathonikos N, Donovan T A, Klopfleisch R, Ammeling J, Ganz J, et al. 2024. Domain generalization across tumor types, laboratories, and species - insights from the 2022 edition of the Mitosis Domain Generalization Challenge. Medical Image Analysis, 94: 103155 [DOI: 10.1016/j.media.2024.103155].

Barlow P W. 1977. Changes in chromatin structure during the mitotic cycle. Protoplasma, 91 (2): 207-211 [DOI: 10.1007/BF01276735].

Beevi K S, Nair M S and Bindu G R. 2016a. Automatic segmentation of cell nuclei using Krill Herd optimization based multi-thresholding and localized active contour model. Biocybernetics and Biomedical Engineering, 36(4): 584-596 [DOI: 10.1016/j.bbe.2016.06.005].

Beevi K S, Nair M S and Bindu G R. 2016b. Detection of mitotic nuclei in breast histopathology images using localized ACM and Random Kitchen Sink based classifier//2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Orlando, FL, USA: IEEE: 2435-2439 [DOI: 10.1109/EMBC.2016.7591222].

Beevi K S, Nair M S and Bindu G R. 2019. Automatic mitosis detection in breast histopathology images using convolutional neural network based deep transfer learning. Biocybernetics and Biomedical Engineering, 39(1): 214-223 [DOI: 10.1016/j.bbe.2018.10.007].

Cireşan D C, Giusti A, Gambardella L M and Schmidhuber J. 2013. Mitosis detection in breast cancer histology images with deep neural networks//Medical Image Computing and Computer-Assisted Intervention - MICCAI 2013, Part II. Berlin, Heidelberg: Springer: 411-418 [DOI: 10.1007/978-3-642-40763-5_51].

Dodballapur V, Song Y, Huang H, Chen M, Chrzanowski W and Cai W D. 2019. Mask-driven mitosis detection in histopathology images//2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). Venice, Italy: IEEE: 1855-1859 [DOI: 10.1109/ISBI.2019.8759164].

Fernandez-Martín C, Silva-Rodriguez J, Kiraz U, Morales S, Janssen E A M and Naranjo V. 2024. Uninformed Teacher-Student for hard-samples distillation in weakly supervised mitosis localization. Computerized Medical Imaging and Graphics, 112: 102328 [DOI: 10.1016/j.compmedimag.2024.102328].

Gonzalez D G, Carias J, Castilla Y C, Rodrigues J, Adão T, Jesus R, et al. 2023. Evaluating rotation invariant strategies for mitosis detection through YOLO algorithms//Wireless Mobile Communication and Healthcare. Cham: Springer Nature Switzerland: 24-33 [DOI: 10.1007/978-3-031-32029-3_3].

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90].

He K M, Gkioxari G, Dollár P and Girshick R. 2017. Mask R-CNN//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE: 2980-2988 [DOI: 10.1109/ICCV.2017.322].

Irshad H, Jalali S, Roux L, Racoceanu D, Lim J H, Le Naour G, et al. 2013. Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach. Journal of Pathology Infor-

© 中国图象图形学报版权所有

matics, 4(2): 12 [DOI: 10.4103/2153-3539.109870].

Jahanifar M, Shephard A, Zamanitajeddin N, Graham S, Raza S E A, Minhas F, et al. 2024. Mitosis detection, fast and slow: robust and efficient detection of mitotic figures. Medical Image Analysis, 94: 103132 [DOI: 10.1016/j.media.2024.103132].

Khan A M, ElDaly H and Rajpoot N M. 2013. A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images. Journal of Pathology Informatics, 4(1): 11 [DOI: 10.4103/2153-3539.112696].

Khanam R and Hussain M. 2024. YOLOv11: an overview of the key architectural enhancements[EB/OL]. [2026-11-05]. https://arxiv.org/abs/2410.17725v1.

Khosla P, Teterwak P, Wang C, Sarna A, Tian Y L, Isola P, et al. 2020. Supervised contrastive learning//Advances in Neural Information Processing Systems 33. Red Hook, NY, USA: Curran Associates, Inc.: 18661-18673.

Lafarge M W, Pluim J P W, Eppenhof K A J, Moeskops P and Veta M. 2017. Domain-adversarial neural networks to address the appearance variability of histopathology images//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer International Publishing: 83-91 [DOI: 10.1007/978-3-319-67558-9_10].

Lafarge M W and Koelzer V H. 2022. Rotation invariance and extensive data augmentation: a strategy for the MItosis DOmain Generalization (MIDOG) challenge//Biomedical Image Registration, Domain Generalisation and Out-of-Distribution Analysis. Cham: Springer International Publishing: 62-67 [DOI: 10.1007/978-3-030-97281-3_9].

Lafarge M W and Koelzer V H. 2023. Fine-grained hard-negative mining: generalizing mitosis detection with a fifth of the MIDOG 2022 dataset//Mitosis Domain Generalization and Diabetic Retinopathy Analysis. Cham: Springer Nature Switzerland: 226-233 [DOI: 10.1007/978-3-031-33658-4_24].

Lei H J, Liu S M, Xie H, Kuo J Y and Lei B Y. 2019. An improved object detection method for mitosis detection//2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Berlin, Germany: IEEE: 130-133 [DOI: 10.1109/EMBC.2019.8857343].

Lei H J, Liu S M, Elazab A, Gong X H and Lei B Y. 2021. Attention-guided multi-branch convolutional neural network for mitosis detection from histopathological images. IEEE Journal of Biomedical and Health Informatics, 25(2): 358-370 [DOI: 10.1109/JBHI.2020.3027566].

Li C, Wang X G, Liu W Y and Latecki L J. 2018. DeepMitosis: mitosis detection via deep detection, verification and segmentation networks. Medical Image Analysis, 45: 121-133 [DOI: 10.1016/j.media.2017.12.002].

Li C, Wang X G, Liu W Y, Latecki L J, Wang B and Huang J Z. 2019. Weakly supervised mitosis detection in breast histopathology images using concentric loss. Medical Image Analysis, 53: 165-178 [DOI: 10.1016/j.media.2019.01.013].

Lin F C, Xie H T, Liu C B and Zhang Y D. 2023. Mitosis detection by appearance and motion pattern perception. Journal of Image and Graphics, 28(9): 2733-2748 (林凡超, 谢洪涛, 刘传彬, 张勇东. 2023. 外观和运动模式感知的有丝分裂细胞检测. 中国图象图形学报, 28(9): 2733-2748) [DOI: 10.11834/jig.220901].

Liu R, Tang Q L, Wang Y, Chen P Z, Shu C, Wang S, et al. 2026. Adaptive context-aware and multi-granularity domain adversarial for mitosis detection in breast cancer histopathology images[J/OL]. Journal of Image and Graphics: 1-14 (刘蓉, 唐奇伶, 王艳, 陈鹏舟, 舒畅, 王帅, 等. 2026. 自适应上下文结合多粒度域对抗的乳腺癌病理图像有丝分裂检测. 中国图象图形学报: 1-14 [DOI: 10.11834/jig.260056].

Liu Z, Lin Y T, Cao Y, Hu H, Wei Y X, Zhang Z, et al. 2021. Swin Transformer: hierarchical vision transformer using shifted windows//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: IEEE: 9992-10002 [DOI: 10.1109/ICCV48922.2021.00986].

Maroof N, Khan A, Qureshi S A, Rehman A U, Khalil R K and Shim S O. 2020. Mitosis detection in breast cancer histopathology images using hybrid feature space. Photodiagnosis and Photodynamic Therapy, 31: 101885 [DOI: 10.1016/j.pdpdt.2020.101885].

Nateghi R, Danyali H and Helfroush M S. 2021. A deep learning approach for mitosis detection: application in tumor proliferation prediction from whole slide images. Artificial Intelligence in Medicine, 114: 102048 [DOI: 10.1016/j.artmed.2021.102048].

Paeng K, Hwang S, Park S and Kim M. 2017. A unified framework for tumor proliferation score prediction in breast histopathology//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer International Publishing: 231-239 [DOI: 10.1007/978-3-319-67558-9_27].

Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015, Part III. Cham: Springer International Publishing: 234-241 [DOI: 10.1007/978-3-319-24574-4_28].

Sebai M, Wang X G and Wang T J. 2020. MaskMitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images. Medical & Biological Engineering & Computing, 58(7): 1603-1623 [DOI: 10.1007/s11517-020-02175-z].

Simonyan K and Zisserman A. 2014. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2026-11-05]. https://arxiv.org/abs/1409.1556v1.

Sohail A, Khan A, Nisar H, Tabassum S and Zameer A. 2021. Mitotic nuclei analysis in breast cancer histopathology images using deep ensemble classifier. Medical Image Analysis, 72: 102121 [DOI: 10.1016/j.media.2021.102121].

© 中国图象图形学报版权所有

Sun Y S, Zhao Z, Yang Z N, Xu F, Lu H J, Zhu Z Y, et al. 2017. Risk factors and preventions of breast cancer. International Journal of Biological Sciences, 13(11): 1387-1397 [DOI: 10.7150/ijbs.21635].

Tang Q L and Cai Y. 2024. Deep radial basis function networks with subcategorization for mitosis detection in breast histopathology images. Medical Image Analysis, 95: 103204 [DOI: 10.1016/j.media.2024.103204].

Tellez D, Balkenhol M, Otte-Höller I, van de Loo R, Vogels R, Bult P, et al. 2018. Whole-slide mitosis detection in H&E breast histology using PHH3 as a reference to train distilled stain-invariant convolutional networks. IEEE Transactions on Medical Imaging, 37(9): 2126-2136 [DOI: 10.1109/TMI.2018.2820199].

van den Oord A, Li Y Z and Vinyals O. 2018. Representation learning with contrastive predictive coding[EB/OL]. [2026-11-05]. https://arxiv.org/abs/1807.03748v1.

Veta M, Heng Y J, Stathonikos N, Bejnordi B E, Beca F, Wollmann T, et al. 2019. Predicting breast tumor proliferation from whole-slide images: the TUPAC16 challenge. Medical Image Analysis, 54: 111-121 [DOI: 10.1016/j.media.2019.02.012].

Wahab N, Khan A and Lee Y S. 2019. Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images. Microscopy, 68(3): 216-233 [DOI: 10.1093/jmicro/dfz002].

Wang H D, Xu H, Li B B, Pan X P, Zeng L Q, Lan R S, et al. 2024a. A novel dataset and a two-stage mitosis nuclei detection method based on hybrid anchor branch. Biomedical Signal Processing and Control, 87, Part A: 105374 [DOI: 10.1016/j.bspc.2023.105374].

Wang H D, Wang X X, Li B B, Liu Z P, Xu H, Pan X P, et al. 2024b. GZMH: a dataset of breast cancer pathological images for mitosis nuclei detection and segmentation. Journal of Image and Graphics, 29(3): 608-619 (汪华登, 王雪馨, 黎兵兵, 刘志鹏, 许浩, 潘细朋, 等. 2024b. GZMH:用于有丝分裂细胞核检测和分割的乳腺癌病理图像数据集. 中国图象图形学报, 29(3): 608-619 [DOI: 10.11834/jig.230332]).

Wang H D, Liu Z P, Pan X P, Yu K, Lan R S, Guan J L, et al. 2025. A novel dataset and a two-stage deep learning method for breast cancer mitosis nuclei identification. Digital Signal Processing, 158: 104978 [DOI: 10.1016/j.dsp.2024.104978].

Wang X Y, Zhang J, Yang S, Xiang J X, Luo F, Wang M H, et al. 2023. A generalizable and robust deep learning algorithm for mitosis detection in multicenter breast histopathological images. Medical Image Analysis, 84: 102703 [DOI: 10.1016/j.media.2022.102703].

Wollmann T and Rohr K. 2021. Deep Consensus Network: aggregating predictions to improve object detection in microscopy images. Medical Image Analysis, 70: 102019 [DOI: 10.1016/j.media.2021.102019].

Zerhouni E, Lanyi D, Viana M and Gabrani M. 2017. Wide residual networks for mitosis detection//2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). Melbourne, Australia: IEEE: 924-928 [DOI: 10.1109/ISBI.2017.7950667].

作者简介

黄振杰，性别男，硕士研究生，研究方向为计算病理学。E-mail：ZhenJieHuang@mails.guet.edu.cn

汪华登，通信作者，性别男，教授，主要研究方向为计算病理学。E-mail：whd@guet.edu.cn

黎兵兵，男，主治医师，主要研究方向为肿瘤病理学与医学图像人工智能研究，E-mail：libingbing19932021@126.com

蓝如师，性别男，教授，研究方向为：医学人工智能。E-mail：rslan2016@163.com

肖秋香，女，主任医师，主要研究方向为消化系统、软组织及乳腺疾病的病理诊断，E-mail：xqx1812@163.com

方伟兰，女，主治医师，主要研究方向为结直肠癌，E-mail：fangweilan715@163.com

何芬，女，住院医师，主要研究方向为肿瘤病理学，E-mail：1224140503@qq.com

刘力源，男，主治医师，主要研究方向为恶性肿瘤性疾病的病理诊断及分子靶向治疗，E-mail：jxxflly@163.com。

郭振强，男，主治医师，主要研究方向为细胞病理和生殖道肿瘤病理，E-mail：827044749@qq.com

陈龙娇，女，主治医师，主要研究方向为乳腺病理，E-mail：978442323@qq.com

梁莉，女，教授，主要研究方向为肿瘤病理学与医学图像人工智能，E-mail：lli@smu.edu.cn

© 中国图象图形学报版权所有