

基于 CMCD-YOLO 的宫颈癌细胞检测算法

郑馨^{1,2*}, 赵路线^{1,2}, 张亮亮¹, 陈义仁¹, 刘德阳¹¹安庆师范大学计算机与信息学院, 安徽 安庆 246133;²安庆师范大学高性能计算平台, 安徽 安庆 246133

摘要 宫颈细胞检测对宫颈癌早期筛查具有重要意义,但由于宫颈细胞图像中小目标聚集、形态多样等问题,现有宫颈细胞图像检测算法难以满足临床高精度筛查需求。为此,提出一种基于 YOLOv11 目标检测算法的 CMCD-YOLO (Complex Small Cervical Cell Detection-YOLO)宫颈细胞检测模型。首先,针对异常细胞大小占比小且密集的特点,新增 P2 小目标检测层构建四尺度特征金字塔,并结合扩展的 P2 检测层,重构自适应特征融合检测头为多尺度自适应融合头 (MSAFHead),通过多尺度特征的自适应加权融合机制,增强模型对不同尺寸细胞的检测性能;然后,使用双卷积核重构 C3k2 模块为 DuCC3k2,强化对细胞局部特征与全局特征的协同提取;另外,在骨干和颈部引入多分支卷积模块,通过多样化分支结构增强空间感知能力,增强对细胞细节特征的提取能力。实验结果表明在公共数据集上,CMCD-YOLO 在 mAP50 (交并比阈值为 0.50 时的平均精度均值)和 mAP75 (交并比阈值为 0.75 时的平均精度均值)上分别实现了 2.7 百分点、3.0 百分点的提升,优于现有检测方法。

关键词 宫颈细胞; 目标检测; YOLOv11; 自适应特征融合

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP252386

1 引言

宫颈癌是全球女性常见的恶性肿瘤之一,早期筛查与诊断对降低发病率和死亡率具有重要意义^[1]。液基薄层细胞学(TCT)作为宫颈癌筛查的主要手段,能够有效发现宫颈上皮内瘤变及早期癌变^[2]。随着计算机视觉技术的不断进步,医学影像分析领域正经历深刻变革^[3]。基于深度学习的目标检测算法可自动检测和识别异常细胞,为医生提供决策支持,有效减轻其负担,提升诊断的准确性与效率。

YOLO 算法以其出色的速度和较高的准确率著称,它使用深度神经网络同时进行目标分类与位置检测。随着 YOLO 算法的持续更新,YOLOv7 至 YOLOv13 等版本相继推出。近年来,YOLO 算法已广泛应用于细胞图像检测任务。为解决宫颈细胞图像中密集小目标漏检率高的问题,一种融入难例分类器、标签平滑正则化与图像金字塔策略的级联改进 YOLOv3 网络被应用于宫颈异常细胞检测^[4]。郑雯等^[5]提出一种基于 YOLOv3 模型的改进识别算法 mo-YOLOv3,采用数据增强^[6]策略扩充数据集,利用多尺度特征融合算法优化小目标检测,并改进损失函数以

计算相对位置误差。针对血细胞图像中小目标和重叠目标的检测难点,周煜庭等^[7]提出一种改进的 YOLOv7 算法,融合加权双向特征金字塔网络与递归门控卷积的特征金字塔 HorNet-BiFPN 结构,增强了网络的特征融合能力,有效解决了重叠红细胞的检测问题。高仪等^[8]提出轻量化检测模型,通过添加微小目标检测头、去除大目标检测头以专注小目标特征提取,设计部分多尺度特征融合模块丰富灰度影像特征多样性,并构造上下文特征金字塔池化模块强化局部与全局信息表达,显著提升了灰度影像小目标的检测精度与轻量化适配性。岳剑峰等^[9]提出采用分布移位卷积替代传统卷积以降低模型参数量与复杂度,引入高效多尺度注意力机制增强对不同尺度尤其是小特征的识别能力。张泽毅等^[10]提出高效且轻量化的检测网络,通过设计动态特征提取模块动态调整卷积权重以增强多尺度缺陷特征提取与融合能力,开发小波自适应下采样结构减少空间信息损失并抵抗复杂背景噪声,有效提升了微小缺陷的检测效率。涂俊博等^[11]通过 SiLU 激活函数优化模型收敛速度与泛化性能,采用卷积块注意力模块优化主干特征采样结构,引入内容感知特征重组提升检测精度与速率,结合动态非单

收稿日期: 2025-11-10; 修回日期: 2025-12-15; 录用日期: 2025-12-19; 网络首发日期: 2026-01-12

基金项目: 国家自然科学基金(62171002, 62302014)、安徽省自然科学基金青年项目(B类)(2508085Y039)、安徽省高校杰出青年科研项目(2024AH020009)

通信作者: *zxaoyou@aqnu.edu.cn

调损失函数增强对小目标数据集和噪声的鲁棒性,有效提升表面小目标缺陷的检测准确性。吕彦朋等^[12]提出改进型微生物细胞识别算法,通过引入优化的路径聚合网络捕获不同尺度细胞的特征信息,并采用 DIOU_Loss 函数优化边界框定位精度,有效提升了多种类微生物细胞的识别效率与准确性。Xu 等^[13]提出多尺度特征融合网络 MSFF-Net,该网络是近年来宫颈细胞多尺度检测领域的典型方案,虽通过跨尺度池化模块捕获细胞局部细节与全局上下文信息,尝试优化重叠细胞检测效果,但未充分结合宫颈细胞的场景特性设计专属结构:一方面,未构建专门的小目标检测层,对微小病变细胞仅依赖通用跨尺度池化提取特征,易因池化过程丢失关键细节;另一方面,未对特征提取模块进行重构,难以高效协同捕捉细胞局部纹理与全局形态,面对背景干扰时的区分能力有限。

然而,尽管现有算法为宫颈细胞的自动检测提供了诸多可借鉴的实现路径,但仍存在以下问题:其一,现有研究多聚焦于有限尺度特征的适应性优化,对多尺度特征的适配能力不足,难以兼顾微小异常细胞与较大细胞的精准识别;其二,算法对空间信息的挖掘与利用程度有限,在尺度选择上存在固定化局限;其三,面对宫颈细胞受背景噪声干扰的问题,现有特征提取机制对细胞关键特征的提取能力较弱,难以精准捕捉细微病理特征。

针对上述问题,在综合评估 YOLOv11^[14]在宫颈细胞检测场景中的适配性后,本研究选择其作为基准模型。主要原因如下:其一,YOLOv11 是 Ultralytics 于 2024 年 9 月 30 日在年度活动 YOLO Vision 2024 上正式发布的官方版本,已提供完整的开源支持、官方权重文件及成熟的社区技术反馈,具备良好的稳定性与可复现性;该模型的部署门槛低于 YOLOv9^[15]与 YOLOv10^[16],为后续针对宫颈细胞检测的模块优化提供了更灵活的基础架构。其二,YOLOv11 将 C2F 模块升级为计算量更低、特征聚合能力更强的 C3k2 模块,更适配宫颈细胞图像中小目标密集、需频繁进行多尺度特征计算的需求。综合评估表明,YOLOv11 优于 YOLOv9 与 YOLOv10,因此本研究选定 YOLOv11 作为基准模型,并提出一种改进的宫颈细胞检测算法 CMCD-YOLO (Complex Small Cervical Cell Detection-YOLO):

1) 新增 P2 检测层,构建 P2-P5 多尺度金字塔,配合自适应空间特征融合,动态优化权重分配,有效减少微小病变与密集重叠细胞的漏检和误检。

2) 构建 DuCC3k2 模块,以 3×3 分组卷积提取局部纹理, 1×1 逐点卷积整合全局上下文,双路径并行协同,增强宫颈细胞局部与全局特征的融合表达与判别能力。

3) 在主干与颈部引入多分支卷积模块 (Diverse Branches Block, DBB),利用多路径结构扩大感受野差异,强化对细胞边缘、核质细节等空间信息的捕捉,提升细微形态识别精度。

2 CMCD-YOLO 网络及实现

提出的 CMCD-YOLO 网络是基于 YOLOv11 的改进网络模型。其主干与颈部仍沿用 CSPNet (Cross Stage Partial Network) 框架,但将 C2F 模块升级为 C3k2 模块,该模块具有较低计算量与更强特征聚合能力,在保持与 YOLOv8 同量级参数的同时显著降低了 FLOPs (Floating Point Operations),并提升了对多尺度特征的捕获能力。检测头部分引入 head^[17]思想,通过多尺度特征图并行预测,实现大目标与小目标的同步精准定位。相较于 YOLOv8^[18]的 PGI 机制, YOLOv11 以结构级轻量化的方式达到了同等甚至更高的精度,且推理延迟更低、模型尺寸更小,更有利于实验复现与边缘部署。

图 1 为 CMCD-YOLO 网络结构图,由主干 (Backbone)、颈部 (Neck) 和头部 (Head) 三部分组成。在主干与颈部部分,采用 DBB^[19]。在主干网络中,使用双卷积核 DualConv^[20]重构 C3k2 模块。此外,在网络中增加了一个 P2 小目标检测层,并将其添加到检测头部分,同时将自适应空间特征融合^[21-22]应用于 YOLO 检测头,重构为四头结构 MSAFHead。这些改进有效提升了模型在宫颈细胞图像中的检测性能。

2.1 MSAFHead 模块

宫颈细胞图像中的目标细胞属于小目标范畴。Chen 等^[23]将小目标定义为边界框宽高与图像宽高比例均小于 0.1 的特定值。宫颈细胞数据集存在尺度差异大、染色质分布异常等特点,而这些细微特征又易在卷积过程中被削弱或淹没。模型输入图像大小为 640×640 ,原 P3 检测头特征层大小为 80×80 。输入特征图在空间维度进行下采样,在此过程中易丢失空间细节信息,导致小目标表征能力不足。加之复杂背景干扰,进一步加剧了不同层次特征在尺度上不匹配的问题。传统检测头难以精准区分目标与背景,这也凸显了新增 P2 层的必要性。以 P2 层为例,首先将其他三个层次 (P3、P4、P5) 的特征通过尺度调整至相同分辨率,即对 P3 层特征采用二倍上采样,对 P4 层特征先压缩通道再四倍上采样,对 P5 层特征则连续两次二倍上采样并补充边缘信息,确保调整后的特征在空间维度上精准对齐。随后,网络通过训练学习自适应空间融合权重,最终将不同层次的特征按权重融合。在融合过程中,携带矛盾信息的特征 (如黏液的高亮区域与细胞的染色特征) 会被分配较低权重甚至过滤,而更具判别性的病变细胞核的纹理、细胞间的粘连痕迹特征将获得更高权重,从而主导预测过程。MSAFHead 的结构如图 2 所示。

MSAFHead 检测头在进行自适应特征融合时针对第 l 层, l 分别对应 P2、P3、P4、P5,融合过程如下:

$$y_j^l = \alpha_j^l \cdot x_j^{0 \rightarrow l} + \beta_j^l \cdot x_j^{1 \rightarrow l} + \gamma_j^l \cdot x_j^{2 \rightarrow l} + \delta_j^l \cdot x_j^{3 \rightarrow l} \quad (1)$$

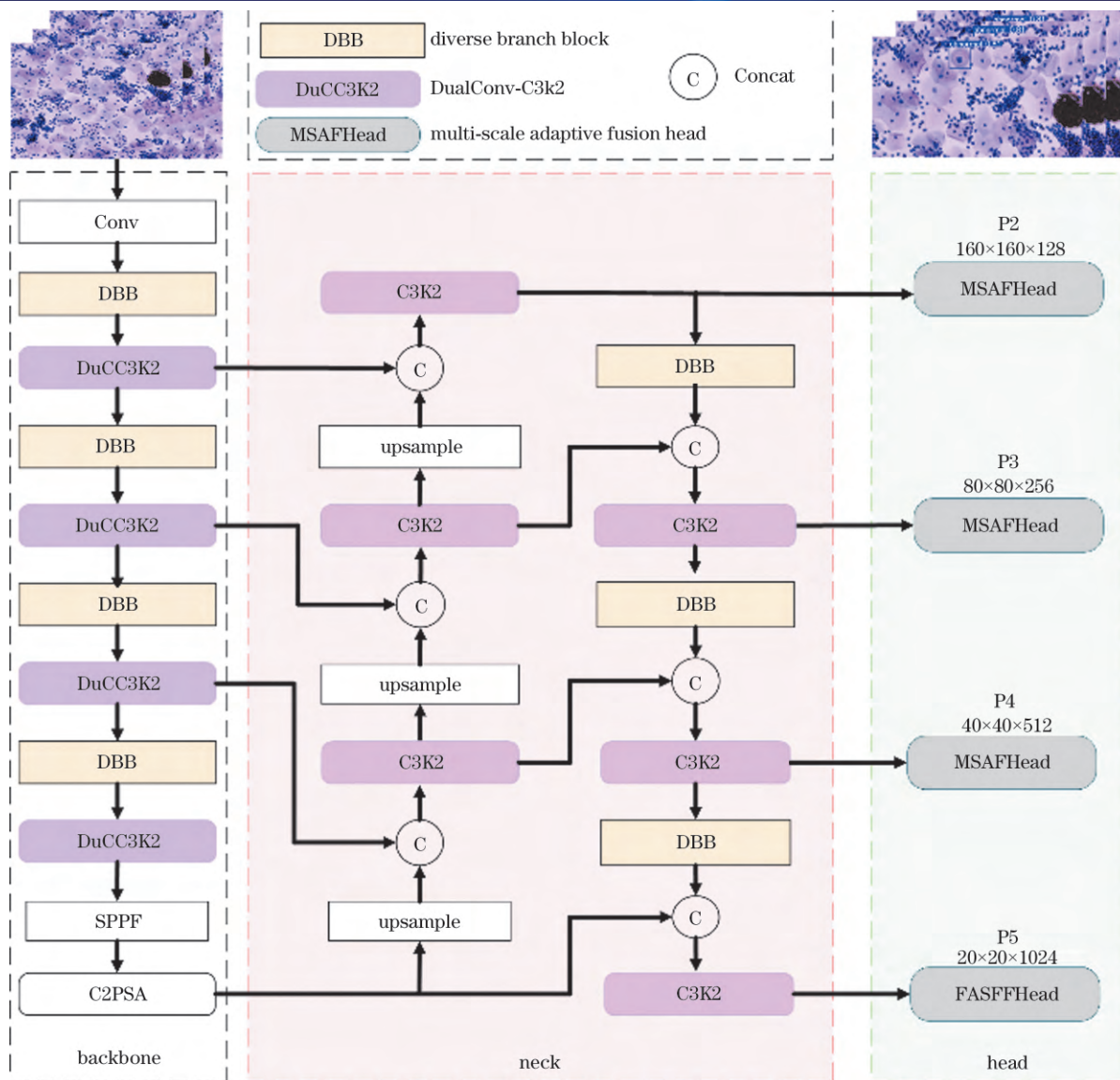


图1 CMCD-YOLO网络结构

Fig. 1 Structure of CMCD-YOLO network

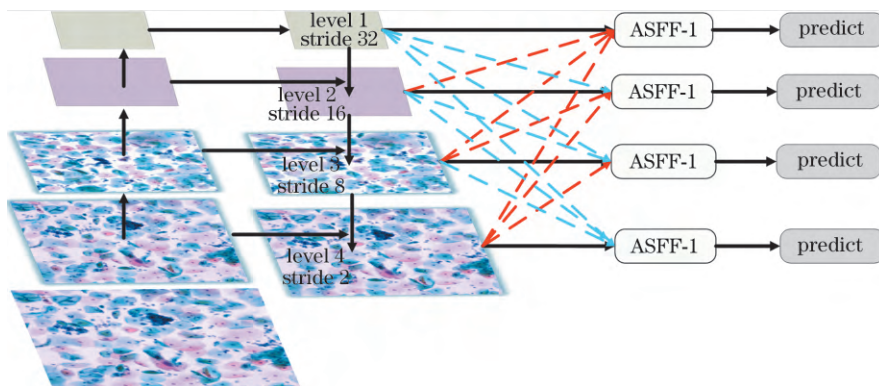


图2 MSAFHead模块结构

Fig. 2 Structure of MSAFHead module

式中: y_{ij}^l 为在第 l 层特征图上位置 (i, j) 处融合后的输出特征向量; $x_{ij}^{n \rightarrow l}$ 表示特征图上位置 (i, j) 处的 x 特征向量从第 n 层调整到第 l 层后的结果, 其调整过程结合

了卷积压缩与插值采样, 既保留原始特征信息又确保尺度匹配; $\alpha_{ij}^l, \beta_{ij}^l, \gamma_{ij}^l, \delta_{ij}^l$ 为网络自适应学习的空间重要性权重, 满足 $\alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l + \delta_{ij}^l = 1$ 且均在 $[0, 1]$ 范围

内,通过 Softmax 函数实现,以 α_{ij}^l 为例:

$$\alpha_{ij}^l = \frac{\exp(\lambda_{a_{ij}}^l)}{\exp(\lambda_{a_{ij}}^l) + \exp(\lambda_{\beta_{ij}}^l) + \exp(\lambda_{\gamma_{ij}}^l) + \exp(\lambda_{\delta_{ij}}^l)} \quad (2)$$

式中: $\lambda_{a_{ij}}^l, \lambda_{\beta_{ij}}^l, \lambda_{\gamma_{ij}}^l, \lambda_{\delta_{ij}}^l$ 为通过 1×1 卷积从调整后的特征中学习的控制参数。

MSAFHead 检测头能够在每个空间位置动态过滤冲突信息,有效改善不同尺度特征间的不一致性,增强特征的尺度不变性。在检测微小的细胞时,P2层的高分辨率特征与P3层的细胞形态特征会被赋予更高权重,而P4、P5层的全局特征则作为辅助;在检测细胞集群时,P4、P5层的上下文特征权重上升,确保模型同时捕获全局分布与局部细节。这种自适应融合机制显

著提升了网络对密集小目标细胞和复杂背景下细微病变特征的检测能力。

2.2 DuCC3k2模块

在宫颈细胞检测场景中,低级特征图携带丰富的局部细节信息,而高级特征图则富含更多细胞整体形态等全局信息。低级与高级特征的精准融合是实现病变细胞高效识别的关键。YOLOv11中的C3k2模块能通过深层网络结构提取多尺度特征,但在处理宫颈细胞图像时,其原始卷积操作难以有针对性地捕捉病变细胞与正常细胞的细微差异。本文使用双重卷积核 DualConv 重构 C3k2 中的卷积操作,以增强对宫颈细胞细微特征的提取能力。DualConv 结构如图 3 所示。

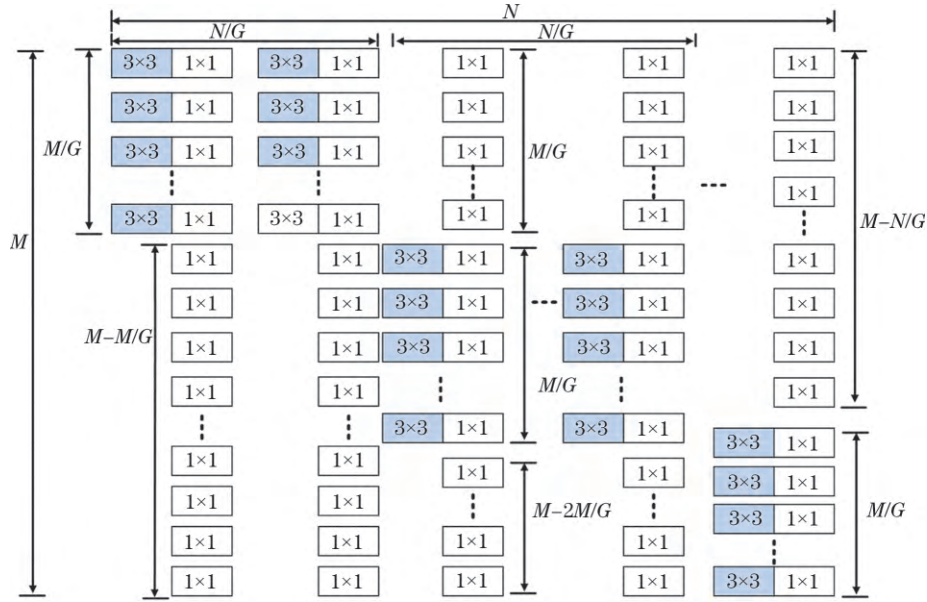


图3 DualConv模块结构

Fig. 3 Structure of DualConv module

其核心设计是将 3×3 分组卷积与 1×1 逐点卷积并行结合,对输入特征图的通道同时进行处理。对C3k2模块的改进集中在Bottleneck中的卷积层,将原卷积核替换为DualConv,实现了更高效的卷积计算:

$$\text{DuCC3k2}(\mathbf{x}) = \text{Conv}(\text{Concat}(\text{Seq}(\text{DuCBottleneck}(\text{Conv}(\mathbf{x})), \text{Conv}(\mathbf{x}))) \quad (3)$$

式中: $\mathbf{x}$ 为输入特征图;Seq(DuCBottleneck)为由多个DuCBottleneck模块组成的串行序列,对特征进行深度提取。在Bottleneck结构中,先通过 1×1 卷积将输入通道数降维,再将原本的 3×3 标准卷积替换为DualConv。该设计使 3×3 分组卷积聚焦于细胞核与细胞质的边界特征,如核膜的模糊程度, 1×1 卷积则强化染色梯度,病变区域的染色深浅差异等跨通道信息,整体仍遵循CSP风格以保证特征融合的有效性。重构后的DuCC3k2结构如图4所示。

从特征响应角度分析,DualConv对宫颈细胞特征的提取能力可通过输出特征图的响应值分布体现。设某一空间位置 (h, w) 的输出特征值为 $\mathbf{O}_{j,h,w}$,其由 3×3

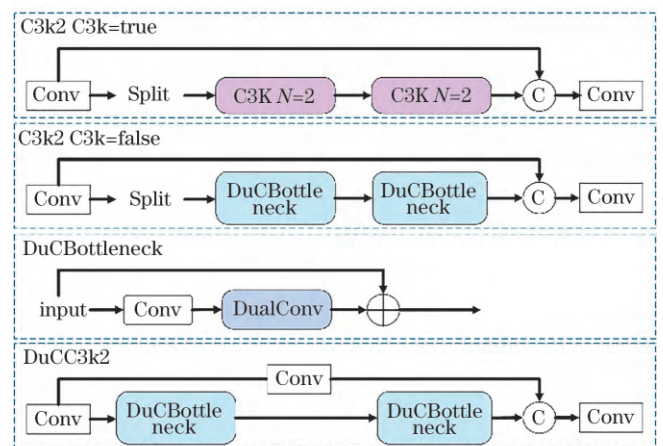


图4 DuCC3k2模块结构

Fig. 4 Structure of DuCC3k2 module

分组卷积响应 $\mathbf{O}_{j,h,w}^{\text{GC}}$ 与 1×1 逐点卷积响应 $\mathbf{O}_{j,h,w}^{\text{PC}}$ 叠加而成:

$$\mathbf{O}_{j,h,w} = \mathbf{O}_{j,h,w}^{\text{GC}} + \mathbf{O}_{j,h,w}^{\text{PC}} \quad (4)$$

$O_{j,h,w}^{GC}$ 对局部纹理变化更敏感(如细胞核内的染色颗粒), $O_{j,h,w}^{PC}$ 则更侧重全局特征(如细胞的整体轮廓), 二者协同提升对病变细胞的辨识度。

针对宫颈细胞形态多样、病变特征细微、细胞核微小形变、染色强度差异及背景干扰复杂等特点, DuCC3k2 模块通过构建双重卷积模块 DualConv, 实现了复杂宫颈细胞低级特征与高级特征的有效融合。

2.3 DBB 模块

针对宫颈细胞尺寸微小、形态多样、特征难以捕捉的问题, 标准卷积与普通空洞卷积因固定卷积核与步长的限制, 难以精准提取细胞核边缘、纹理等关键特征。DBB 模块通过并行多分支结构, 包含原始卷积、 1×1 卷积、平均池化及 $1 \times 1-k \times k$ 组合卷积等丰富特征空间, 可从多维度聚合细胞复杂特征, 增强对细微形态与局部颜色变化的感知能力, 更适配宫颈细胞检测场景。无需改变网络整体架构, 即可通过多分支特征融合弥补普通卷积跨层信息挖掘不足的缺陷。细胞图像中, 不同类型、状态的细胞形态差异显著, 早期病变细胞局部颜色变化的细微特征难以捕捉。训练阶段通过多分支结构, 包括原始卷积分支 dbb_origin、平均池化分支 dbb_avg、 1×1 卷积分支 dbb_1x1 及 $1 \times 1-k \times k$ 组合卷积分支 dbb_1x1_kxk, 充分学习细胞的多样特征, 增强模型对细胞细微特征和多样形态的感知能力。训练阶段 DBB 模块的数学表达式如下:

$$O = F_{\text{origin}}(F_{\text{origin}}(I) + F_{1 \times 1}(I) + F_{\text{avg}}(I) + F_{1 \times 1-k \times k}(I)) \quad (5)$$

式中: I 为输入特征图; F_{origin} 为原始卷积分支; $F_{1 \times 1}$ 为 1×1 卷积分支; F_{avg} 为平均池化分支; $F_{1 \times 1-k \times k}$ 为 $1 \times 1-k \times k$ 组合卷积分支; O 为输出特征图。

本研究将 YOLOv11 主干网络和颈部的卷积替换为 DBB 卷积, 这种替换无需对网络整体架构进行大规模修改, 即可增强模型对细胞多样特征的提取能力。

3 数据集

CDetector^[24] 数据集包含 7410 张宫颈显微图像, 其中训练集 6666 张, 测试集 744 张, 涵盖 11 个类别: 意义不明的非典型鳞状细胞(ASCUS)、低级别鳞状上皮内病变(LSIL)、非典型鳞状细胞不排除高级别鳞状上皮内病变(ASCH)、高级别鳞状上皮内病变(HSIL)、鳞状细胞癌(SCC)、毛滴虫(TRICH)、念珠菌(CAND)、非典型腺细胞(AGC)、菌群(FLORA)、疱疹(HERPS)和放线菌(ACTIN)。各类别样本分布详见表 1。尽管不同类别间样本数量存在一定差异, 但整体未见显著不平衡, 各类别样本量均处于合理范围内。因此, 本研究未采用额外的类别不平衡处理, 以保持数据及训练过程的原始性与一致性。

HMCHH-TCT-CellDet^[25] 数据集包含 8037 张宫颈细胞学图像, 其中训练集 5577 张、验证集 1348 张与测试集 1111 张, 仅涵盖异常细胞 1 个类别, 由于其中

表 1 数据集中的病变分类及对应的标注数量

Table 1 Lesion categories in the dataset and the corresponding number of annotation boxes

Lesion category	Train	Test	Total
ASCUS	1835	222	2057
LSIL	1466	173	1639
ASCH	3891	410	4301
HSIL	26305	2823	29128
SCC	1991	229	2290
TRICH	4977	481	5453
CAND	336	27	363
AGC	4989	668	5657
FLORA	127	24	151
HERPS	272	37	309
ACTIN	144	18	162
Total	46333	5112	51445

1 张图像无有效标注, 本文实验实际使用 8036 张有效图像。异常宫颈细胞具有显著的形态特征: 细胞形态畸变, 呈不规则形态; 细胞核异常, 表现为核增大、形态不规则及染色深浅不均。示例图像如图 5 所示, 其细胞核表现出增大、形态不规则及染色不均的典型异常特征。

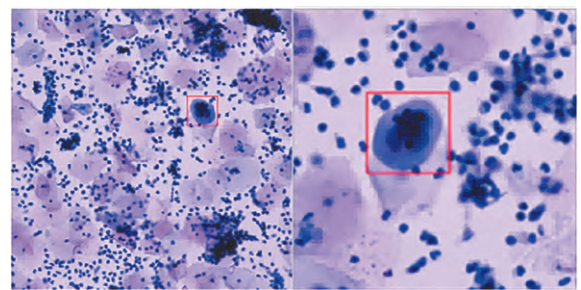


图 5 数据集样本示例

Fig. 5 Sample example of dataset

两个数据集均按照原始文献提供的官方划分方式使用, 以确保结果的可比性与可复现性, 同时保持类别分布与数据原貌一致。为增强实验设计的针对性与结果分析的合理性, 结合图像分辨率与宫颈细胞实际物理尺寸, 统一界定小目标范围并补充微小细胞的专属界定标准: 本研究所有实验图像统一采用 $640 \text{ pixel} \times 640 \text{ pixel}$ 分辨率, 经标定, 该分辨率下图像像素与涂片实际视野的对应关系为“ $640 \text{ pixel} \times 640 \text{ pixel}$ 对应实际视野直径 2 mm ”。据此, 将两数据集中像素面积不超过 $64 \text{ pixel} \times 64 \text{ pixel}$ 、对应物理直径 $15 \mu\text{m}$ 及以下的宫颈细胞统一界定为小目标。其中, CDetector 数据集内的小目标主要为 ASCUS、LSIL 等病变细胞, HMCHH-TCT-CellDet 数据集内的小目标为小型异常细胞。

4 实验结果与分析

实验所用计算机处理器为 Intel Core i7 (2.50 GHz), 搭配 64 GB 内存; 显卡型号为 NVIDIA RTX 3090

GPU,显存 24 GB;运行系统为 Windows 10,深度学习框架为 PyTorch 1.12.1,开发环境为 Python 3.9.7,CUDA 版本为 11.3。当 epoch 设置为 200 时,模型在验证集上的损失曲线趋于平稳,故将训练轮数设为 200。每批次输入图像数量为 4,输入图像分辨率为 640 pixel×640 pixel,初始学习率为 0.01,采用 SGD 优化器进行模型训练。

4.1 评价标准

为全面评估模型性能并与基线模型进行对比,本文采用召回率(R)、mAP50、mAP75、mAP50-95 和 mAR50-95 作为评价指标。首先计算 11 个类别的精确率和召回率。 P 代表精确率, R 代表正确检测数量与需检测总数量的比值,具体计算公式如下:

$$P = N_{TP} / (N_{FP} + N_{TP}) \quad (6)$$

$$R = N_{TP} / (N_{FN} + N_{TP}) \quad (7)$$

式中: N_{TP} 为正确检测到的病变细胞数量; N_{FP} 为错误检测到的病变细胞数量; N_{FN} 为漏检的病变细胞数量。每个类别的平均精确率(AP)和平均精确率均值(mAP)计算方式如下:

$$P_{AP} = \int_0^1 P(R) dR \quad (8)$$

$$P_{mAP} = \frac{1}{N} \sum_{i=1}^N P_{AP,i} \quad (9)$$

对所有类别的 AP 求和取平均即得到 mAP。mAP50 表示预测框与真实框交并比为 0.50 时所有类别的平均检测精度均值;mAP75 表示交并比为 0.75 时的平均检测精度均值;mAP50-95 表示交并比在 0.50~0.95 之间、以 0.05 为间隔递增时所有类别的平均检测精度均值;mAR50-95 表示预测框与真实框交并比在 0.50~0.95 之间、以 0.05 为间隔递增时所有类别的平均召回率均值。

表 3 CMCD-YOLO 模型的消融实验

Table 3 Ablation experiments of CMCD-YOLO model

Group	DBB	DuCC3k2	MSAFHead	mAP50 / %	mAP75 / %	mAP50-95 / %	mAR50-95 / %
1				68.1	35.3	36.5	62.6
2	✓			68.9	36.2	36.8	61.5
3		✓		68.2	36.3	36.3	63.3
4			✓	69.3	34.2	36.9	66.0
5	✓	✓		69.5	37.0	37.1	64.5
6	✓		✓	70.1	37.3	37.5	66.6
7		✓	✓	69.7	36.0	37.0	66.3
8	✓	✓	✓	70.8	38.3	38.0	66.9

4.3 可视化结果

为直观展示改进算法对宫颈细胞检测性能的提升效果,使用原模型、YOLOv12、YOLOv13 和本文改进算法对测试集图像进行检测,在 HMCHH-TCT-CellDet 数据集上的可视化结果如图 6 所示。在细胞密集分布区域(第一张图像),细胞间距离近、特征易混淆,

4.2 消融实验

为验证新增 P2 层的有效性,在 CDetector 数据集上开展了添加不同检测头的对比实验,相关结果如表 2 所示。可见,仅采用 P3、P4、P5 层或 P2、P3、P4 层检测时,检测性能存在一定局限。而当同时纳入 P2、P3、P4、P5 层进行检测时,各项指标均得到提升,表明新增 P2 层能够更充分地利用不同尺度的特征信息,有效增强对不同大小细胞目标的检测能力,展现出明显优势。

表 2 不同检测头对比

Table 2 Comparison of different detection heads

Detector	Detection Head	mAP50 / %	mAP50-95 / %	Recall / %
P3 P4 P5	Detect	61.5	35.4	60.9
P2 P3 P4 P5	Detect	62.0	35.7	61.5
P3 P4 P5	MASFF	62.3	35.9	62.9
P2 P3 P4 P5	MASFF	62.5	36.1	63.4

为验证各改进模块对原模型性能提升的有效性,在 HMCHH-TCT-CellDet 数据集上设计了消融实验,结果如表 3 所示。基准模型(组别 1)在 mAP50、mAP75、mAP50-95 上分别为 68.1%、35.3%、36.5%。依次添加 DBB(组别 2)、DuCC3k2(组别 3)与 MSAFHead(组别 4)后,除组别 3 的 mAP50-95 较基准模型小幅下降 0.2 百分点外,其余主要指标整体有所提升;当依次组合 DBB 和 DuCC3k2(组别 5)、DBB 和 MSAFHead(组别 6)、DuCC3k2 和 MSAFHead(组别 7)时,性能持续优化;而当三项模块完全集成(组别 8)时,CMCD-YOLO 的 mAP50 提升至 70.8%,mAP75 升至 38.3%,mAP50-95 升至 38.0%。消融结果从定量角度证实了各模块的互补性与整体架构的有效性。

CMCD-YOLO 模型检测边界框置信度达 0.69,优于对比模型。相比之下,YOLOv11 置信度为 0.57,YOLOv12 为 0.58,YOLOv13 为 0.51,充分体现了模型在细节特征提取方面的有效性,能够精准捕捉密集环境下每个异常细胞的特征。再看背景复杂的样本(第三张图像),除目标细胞外存在诸多干扰元素,CMCD-

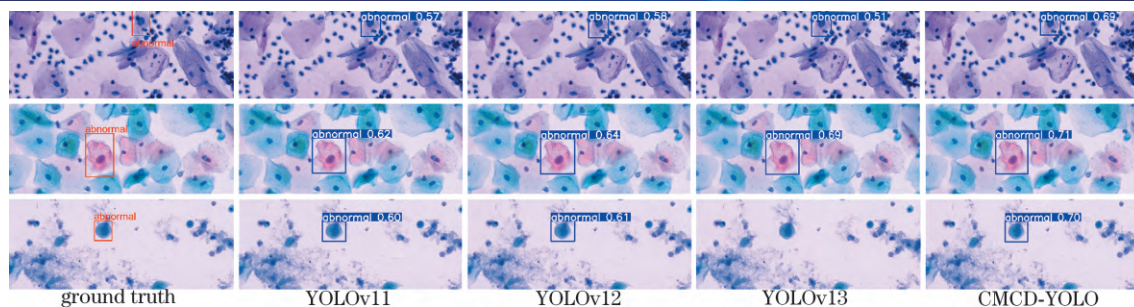


图 6 不同模型在 HMCHH-TCT-CellDet 数据集上的可视化结果

Fig. 6 Visualization results of different models on HMCHH-TCT-CellDet dataset

YOLO 不仅避免了 YOLOv13 出现的漏检,其边界框置信度 0.70 也显著高于 YOLOv11 的 0.60 和 YOLOv12 的 0.61。第二张图像中 CMCD-YOLO 置信度为 0.71, 相比 YOLOv11 的 0.62、YOLOv12 的 0.64、YOLOv13 的 0.69, 同样处于领先。这充分验证了 CMCD-YOLO 在融合低级纹理特征和高级语义特征方面的优势,能从复杂背景和多样细胞形态中准确识别出异常细胞。

在 CDetector 数据集上的可视化结果如图 7 所示。CMCD-YOLO 模型针对 11 类病变及病原体混合场景展现出卓越的细粒度检测能力。图 7 第二行 (HSIL-LSIL 重叠切片) 显示,该模型对两类重叠病变细胞的定位更完整、类别误检较少,说明跨尺度自适应加权机

制有助于降低复杂重叠场景下的误检概率。在图 7 第三行 (图像左上角细胞重叠区域) 中, YOLOv11、YOLOv12、YOLOv13 均存在不同程度的漏检,而 CMCD-YOLO 模型不仅对该重叠区域的细胞实现了准确识别,且具备较高的检测置信度,未出现漏检,充分体现其在细胞重叠复杂场景下的更优检测性能。综上,CMCD-YOLO 模型在多类别、多尺度及复杂背景的宫颈细胞检测中,凭借更高定位精度与更低误报率,可为宫颈病变临床筛查提供客观可靠的量化依据,具有重要临床应用价值。其稳定的高置信度检测结果,可支撑病理诊断精准度,为病变早发现、早干预提供技术支持。

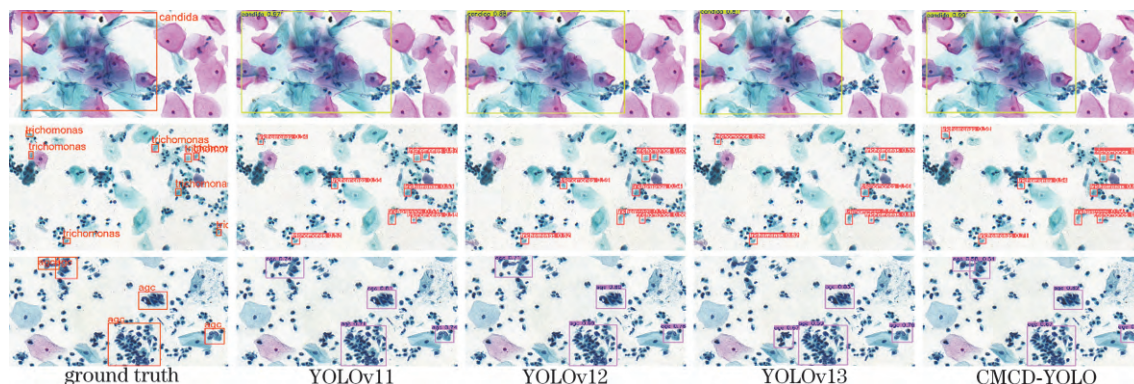


图 7 不同模型在 CDetector 数据集上的可视化结果

Fig. 7 Visualization results of different models on CDetector dataset

4.4 不同模型对比实验

为验证本文改进算法在宫颈细胞图像上的有效性,确保验证实验的严格性,在相同数据集和训练参数下,对比 Faster R-CNN^[26]、DETR、MSFF-Net^[13]、YOLOv8、YOLOv12、YOLOv13 等检测网络与本文提出的 CMCD-YOLO 模型的实验效果,全方位评估各算法在不同类别规模宫颈细胞图像检测任务中的性能。

表 4 详细展示了在相同实验条件下,上述算法在 HMCHH-TCT-CellDet 数据集上的对比结果。在小目标密集的 HMCHH-TCT-CellDet 数据集上,CMCD-YOLO 的平均检测精度更高,mAP50 达 70.8%,mAP75 达 38.3%,mAP50-95 达 38.0%。与

YOLOv7 相比,mAP50 高 33.5 百分点,mAP50-95 高 24.7 百分点;与 YOLOv8 相比,mAP50 高 3.4 百分点,mAP50-95 高 2.7 百分点;与 YOLOv12 相比,mAP50 高 2.9 百分点,mAP50-95 高 1.9 百分点;与 YOLOv13 相比,mAP50 高 2.5 百分点,mAP50-95 高 1.4 百分点;与 D-FINE 模型相比,mAP50 提升 0.9 百分点,mAP75 提升 2.7 百分点,mAP50-95 提升 1.4 百分点,mAR50-95 提升 4.3 百分点。因此,CMCD-YOLO 检测精度高于其他模型,验证了其在细胞密集检测场景中的优势。

表 5 展示了在相同实验条件下,上述算法在 CDetector 数据集上的对比结果。该数据集包含 11 种类别,涵盖大、中、小等不同尺度的细胞目标,检测难度

表 4 不同算法在 HMCHH-TCT-CellDet 数据集上的目标检测结果对比

Table 4 Comparison of object detection results of different algorithms on HMCHH-TCT-CellDet dataset

Method	mAP50 / %	mAP75 / %	mAP50-95 / %	mAR50-95 / %
Retina Net	54.8	20.3	25.7	34.4
FCOS	61.6	21.2	27.6	37.7
Faster R-CNN ^[26]	67.4	25.0	31.5	45.2
Sparse R-CNN	50.1	19.0	23.2	32.0
DETR	47.1	12.2	19.4	36.5
D-FINE	69.9	35.6	36.6	62.6
YOLOv7	37.3	5.4	13.3	22.6
GD-YOLO	67.2	34.6	35.3	62.3
YOLOv8 ^[18]	67.4	34.0	35.3	62.6
YOLOv9 ^[15]	68.5	36.0	36.9	63.6
YOLOv10 ^[16]	65.3	32.3	34.0	61.8
YOLOv11 ^[14]	68.1	35.3	36.5	62.6
YOLOv12	67.9	35.8	36.1	65.2
YOLOv13	68.3	36.1	36.6	65.0
CMCD-YOLO (ours)	70.8	38.3	38.0	66.9

表 5 不同算法在 CDetector 数据集上的目标检测结果对比

Table 5 Comparison of object detection results of different algorithms on the CDetector dataset

Method	mAP50 / %	mAP75 / %	mAP50-95 / %	Recall / %
Faster R-CNN ^[26]	57.8	30.5	27.8	63.8
DETR	47.1	10.8	19.4	36.5
YOLOv7	60.6	15.8	33.0	61.5
GD-YOLO	61.0	27.4	34.2	61.5
YOLOv8 ^[18]	60.6	30.6	34.2	61.2
YOLOv9 ^[15]	61.4	34.9	35.0	62.9
YOLOv10 ^[16]	60.9	32.2	34.5	62.6
YOLOv11 ^[14]	61.5	32.8	35.4	60.9
YOLOv12	61.8	32.5	35.6	63.7
YOLOv13	60.6	33.7	34.7	59.1
MSFF-Net ^[13]	63.2	—	36.4	—
CMCD-YOLO (ours)	63.6	35.3	37.0	64.3

较高。与基线模型 YOLOv11 (mAP50 为 61.5%, Recall 为 60.9%) 相比, CMCD-YOLO 的 mAP50 提升 2.1 百分点, Recall 提升 3.4 百分点; 与 MSFF-Net 相比, mAP50 提升 0.4 百分点, mAP50-95 提升 0.6 百分点, Recall 也具有明显优势。这充分表明模型在多类别、多尺度细胞目标检测任务中具备良好的性能。

5 结 论

针对复杂宫颈细胞图像的自动检测问题, 提出了一种改进 YOLOv11 的 CMCD-YOLO 算法。为增强模型对细胞细微特征的提取能力, 在模型结构中引入多分支卷积; 通过重构 C3k2 模块为 DuCC3k2, 采用双卷积核与分组卷积策略提升特征融合效果; 新增 P2 小目标检

测层, 并将自适应空间特征融合检测头重构为多尺度结构 MSAFHead, 扩展多尺度感受野以适配不同尺度细胞, 从而降低异常细胞尤其是小目标细胞的漏检概率。实验结果表明, 所提算法明显优于基准模型及其他主流算法, 为宫颈细胞的自动检测提供了更精准的解决方案, 对宫颈癌早期筛查的自动化发展具有积极意义。

本研究算法虽在宫颈癌细胞检测精度上优于主流模型, 但在落地应用中仍面临精度优化瓶颈与效率平衡难题, 且在极端密集细胞场景下仍存在一定漏检风险。当面对细胞高度重叠、簇状分布的极端密集场景时, 模型的特征解耦能力与细胞表征能力有限, 易出现部分细胞特征被“淹没”的漏检情况; 此外, 模型结构复杂度上升导致计算开销增加, 在基层边缘设备上的推理速度难以满足临床实时性需求。后续工作将聚焦精度与效率协同优化的框架构建, 探索医疗场景专属的轻量化路径, 同步优化极端密集场景下的特征解耦与边界框分配机制。同时, 期待目标检测领域的技术取得突破, 为解决精度-效率平衡难题与漏检问题提供思路, 推动算法更贴近临床应用。

参 考 文 献

- [1] Arbyn M, Weiderpass E, Bruni L, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis[J]. The Lancet Global Health, 2020, 8(2): e191-e203.
- [2] Hashmi A A, Naz S, Ahmed O, et al. Comparison of liquid-based cytology and conventional Papanicolaou smear for cervical cancer screening: an experience from Pakistan[J]. Cureus, 2020, 12(12): e12293.
- [3] 王嘉旭, 薛鹏, 江宇. 人工智能在宫颈癌筛查中的应用研究进展[J]. 中国肿瘤临床, 2021, 48(9): 468-471.
Wang J X, Xue P, Jiang Y. Research progress in the application of artificial intelligence in cervical cancer screening[J]. Chinese Journal of Clinical Oncology, 2021, 48(9): 468-471.
- [4] Xiang Y, Sun W X, Pan C L, et al. A novel automation-assisted cervical cancer reading method based on convolutional neural network[J]. Biocybernetics and Biomedical Engineering, 2020, 40(2): 611-623.
- [5] 郑雯, 张标标, 吴俊宏, 等. 适于多尺度宫颈癌细胞检测的改进算法[J]. 光电子·激光, 2022, 33(9): 948-958.
Zheng W, Zhang B B, Wu J H, et al. Improved algorithm of multi-scale cervical cancer cells detection[J]. Journal of Optoelectronics·Laser, 2022, 33(9): 948-958.
- [6] Buslaev A, Iglovikov V I, Khvedchenya E, et al. Albumentations: fast and flexible image augmentations[J]. Information, 2020, 11(2): 125.
- [7] 周煜庭, 余华平, 肖粮钧, 等. 基于改进 YOLOv7 模型的血细胞检测算法研究[J]. 国外电子测量技术, 2024, 43(4): 1-9.
Zhou Y T, Yu H P, Xiao L J, et al. Research on blood cell detection algorithm based on YOLOv7 improved model[J]. Foreign Electronic Measurement Technology,

- 2024, 43(4): 1-9.
- [8] 高仪, 陈鹏, 刘子奇, 等. GUAV-YOLO: 面向灰度影像无人机小目标的轻量化检测模型[J]. 光学学报, 2025, 45(23): 2310002.
- Gao Y, Chen P, Liu Z Q, et al. GUAV-YOLO: lightweight detection model for small UAV targets in grayscale images[J]. Acta Optica Sinica, 2025, 45(23): 2310002.
- [9] 岳剑峰, 李伟明, 宁黎华, 等. 基于 YOLO-DEFW 的焊缝缺陷实时检测算法研究[J]. 中国激光, 2025, 52(8): 0802105.
- Yue J F, Li W M, Ning L H, et al. Real-time weld defect detection algorithm based on YOLO-DEFW[J]. Chinese Journal of Lasers, 2025, 52(8): 0802105.
- [10] 张泽毅, 陶志勇, 林森, 等. DFE-YOLO: 高效且轻量化的钢材缺陷检测网络[J]. 激光与光电子学进展, 2025, 62(22): 2237007.
- Zhang Z Y, Tao Z Y, Lin S, et al. DFE-YOLO: efficient and lightweight steel defect detection network[J]. Laser & Optoelectronics Progress, 2025, 62(22): 2237007.
- [11] 涂俊博, 曾佳林, 唐越新, 等. 基于深度学习的太阳能电池表面缺陷检测及分类[J]. 激光与光电子学进展, 2025, 62(2): 0237013.
- Tu J B, Zeng J L, Tang Y X, et al. Detection and classification of surface defects on solar cell panels based on deep learning[J]. Laser & Optoelectronics Progress, 2025, 62(2): 0237013.
- [12] 吕彦朋, 赵颖彤, 苏晓明, 等. 基于改进 YOLOv7 的微生物细胞识别算法[J]. 现代电子技术, 2025, 48(1): 47-54.
- Lü Y P, Zhao Y T, Su X M, et al. Microbial cell recognition algorithm based on improved YOLOv7[J]. Modern Electronics Technique, 2025, 48(1): 47-54.
- [13] Xu C Y, Hu D, Zhang Y, et al. Multi-scale feature fusion network for accurate detection of cervical abnormal cells[J]. Computers, Materials & Continua, 2025, 83(1): 559-574.
- [14] Khanam R, Hussain M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. (2024-10-23) [2025-11-12]. <https://arxiv.org/abs/2410.17725>.
- [15] Wang C Y, Yeh I H, Mark Liao H Y. YOLOv9: learning what you want to learn using programmable gradient information[M]//Leonardis A, Ricci E, Roth S, et al. Computer vision-ECCV 2024. Lecture notes in computer science. Cham: Springer Nature Switzerland, 2024, 15089: 1-21.
- [16] Wang A, Chen H, Liu L H, et al. YOLOv10: real-time end-to-end object detection[EB/OL]. (2024-05-23) [2025-11-12]. <https://arxiv.org/abs/2405.14458>.
- [17] Liang F T, Liang F T, Zhou Y, et al. Review of target detection technology based on deep learning[C]//Proceedings of the 5th International Conference on Control Engineering and Artificial Intelligence, January 14-16, 2021, Sanya, China. New York: ACM Press, 2021: 132-135.
- [18] Jocher G, Qiu J, Chaurasia A, Ultralytics YOLO[EB/OL]. (2023-01-10) [2025-11-12]. <https://github.com/ultralytics/ultralytics>.
- [19] Ding X H, Zhang X Y, Han J G, et al. Diverse branch block: building a convolution as an inception-like unit [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 10881-10890.
- [20] Zhong J C, Chen J Y, Mian A. DualConv: dual convolutional kernels for lightweight deep neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(11): 9528-9535.
- [21] Li X M, Li W Y, Ren D W, et al. Enhanced blind face restoration with multi-exemplar images and adaptive spatial feature fusion[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 2703-2712.
- [22] Fu L, Gu W B, Ai Y B, et al. Adaptive spatial pixel-level feature fusion network for multispectral pedestrian detection[J]. Infrared Physics & Technology, 2021, 116: 103770.
- [23] Chen C Y, Liu M Y, Tuzel O, et al. R-CNN for small object detection[M]//Lai S H, Lepetit V, Nishino K, et al. Computer vision-ACCV 2016. Lecture notes in computer science. Cham: Springer, 2017, 10115: 214-230.
- [24] Liang Y X, Tang Z H, Yan M, et al. Comparison detector for cervical cell/clumps detection in the limited data scenario[J]. Neurocomputing, 2021, 437: 195-205.
- [25] Zhang X, Ji J X, Zhang Q, et al. A large annotated cervical cytology images dataset for AI models to aid cervical cancer screening[J]. Scientific Data, 2025, 12: 23.
- [26] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

Cervical Cancer Cell Detection Algorithm Based on CMCD-YOLO

Zheng Xin^{1,2*}, Zhao Luxian^{1,2}, Zhang Liangliang¹, Chen Yiren¹, Liu Deyang¹

¹School of Computing and Information, Anqing Normal University, Anqing 246133, Anhui, China;

²High Performance Computing Platform, Anqing Normal University, Anqing 246133, Anhui, China

Abstract

Objective Cervical cell detection is critical for early cervical cancer screening. However, existing algorithms fail to meet clinical

precision demands due to the small size, clustered distribution, and high morphological variability of abnormal cells in cervical cytology images. To address these issues, we propose CMCD-YOLO (complex small cervical cell detection-YOLO), an improved model based on YOLOv11. First, an additional P2 detection layer is added to build a four-scale feature pyramid (P2–P5), and the original detection head is redeveloped into a multi-scale adaptive fusion head (MSAFHead) for adaptive weighted fusion of multi-scale features. Second, the conventional C3k2 module is redesigned as DuCC3k2, integrating 3×3 group convolutions (for local textures) and 1×1 pointwise convolutions (for global context) to realize collaborative feature extraction. Moreover, diverse branches blocks (DBB) are inserted into the backbone and neck to enhance spatial information capture. Experiments on two public datasets (CDetector and HMCHH-TCT-CellDet) show that CMCD-YOLO outperforms state-of-the-art methods, achieving 2.7 and 3.0 percentage points improvements in mAP50 (mean average precision at an intersection over union threshold of 0.50) and mAP75 (mean average precision at an intersection over union threshold of 0.75) compared with the baseline YOLOv11, and providing a more accurate solution for automated cervical cancer cell detection.

Methods CMCD-YOLO is built on the YOLOv11 framework, retaining the CSPNet-based backbone and neck while introducing three synergistic improvements to enhance small-target detection and feature discrimination. First, MSAFHead is developed by adding a P2 small-target detection layer to complement the existing P3–P5 layers, forming a four-level feature pyramid (P2–P5) that preserves fine-grained spatial information of tiny cells easily lost during downsampling. The MSAFHead aligns multi-scale features via upsampling and convolution, dynamically learns spatial importance weights through softmax, suppresses conflicting or background features, and emphasizes discriminative cues such as lesion cell textures and adhesion traces. Second, the original C3k2 module is replaced with DuCC3k2, integrating dual convolutional paths: 3×3 group convolution captures local details like nuclear chromatin distribution, while 1×1 pointwise convolution aggregates global context such as cell contour morphology, enabling collaborative extraction of local and global features. Third, DBB are embedded in the backbone and neck, incorporating parallel branches (original convolution, 1×1 convolution, average pooling, and $1\times 1-k\times k$ combined convolution) to enrich feature diversity, expand receptive field differences, and enhance sensitivity to subtle morphological changes like cell edges and nucleocytoplasmic ratios. Experiments are conducted on a platform with an Intel Core i7 (2.50 GHz) processor, 64 GB memory, and NVIDIA RTX 3090 GPU (24 GB VRAM), using PyTorch 1.12.1 and Python 3.9.7 under Windows 10 and CUDA 11.3. The model is trained for 200 epochs with a batch size of 4, 640 pixel $\times$ 640 pixel input resolution, 0.01 initial learning rate, and SGD optimizer, evaluated via recall (R), mAP50, mAP75, mAP50-95 (mean average precision averaged over intersection over union thresholds from 0.50 to 0.95), and mAR50-95 (mean average recall averaged over intersection over union thresholds from 0.50 to 0.95).

Results and Discussions CMCD-YOLO addresses the challenges of small, dense, and morphologically variable abnormal cells in cervical cytology images, with its three targeted enhancements validated through quantitative and qualitative analyses. Ablation experiments on the HMCHH-TCT-CellDet dataset confirm each module's effectiveness: the full model achieves mAP50 of 70.8% and mAP75 of 38.3%, representing 2.7 and 3.0 percentage points improvements over the YOLOv11 baseline, with monotonic performance gains as modules are incrementally integrated. On the CDetector dataset (11 cell categories), CMCD-YOLO reaches mAP50 of 63.6% and recall of 64.3%, outperforming MSFF-Net and YOLOv13. Qualitative visualizations show CMCD-YOLO exhibits higher detection confidence up to 0.71 in dense clusters and complex backgrounds, reducing missed detections and false positives compared to YOLOv11, YOLOv12, and YOLOv13. It accurately identifies overlapping or morphologically diverse lesion cells by leveraging the P2–P5 pyramid's fine-grained information, DuCC3k2's local-global feature fusion, and DBB's enhanced spatial perception. These results demonstrate CMCD-YOLO's ability to provide a more accurate and reliable solution for automated cervical cancer cell detection, supporting early screening automation. Limitations include constrained feature decoupling in extremely dense, overlapping cell clusters and higher computational overhead affecting real-time inference on edge devices. Future work will optimize the accuracy-efficiency trade-off and enhance feature processing mechanisms for extreme dense scenarios.

Conclusions CMCD-YOLO effectively tackles the core challenges posed by cervical cells—including their small size, dense distribution, and morphological variability—through targeted structural optimizations. By integrating a dedicated P2 small-target detection layer, the MSAFHead, the dual-convolution DuCC3k2 module, and DBB, the model achieves superior multi-scale feature fusion and precise fine-grained feature extraction, all while maintaining efficient computational performance. Rigorous experiments on two public datasets, CDetector and HMCHH-TCT-CellDet validate that CMCD-YOLO outperforms mainstream detectors, delivering reliable technical support for the automation of early cervical cancer screening. Future research will focus on three key directions: optimizing the balance between detection accuracy and computational efficiency to enable edge device deployment, enhancing feature decoupling and bounding box assignment mechanisms to reduce missed detections in extremely dense cell clusters, and exploring weakly supervised training strategies to further improve the model's generalization across diverse clinical scenarios.

Key words cervical cell; object detection; YOLOv11; adaptive feature fusion