

基于 YOLOv11 的重叠细胞识别研究

辛俊蕙 1221001047

摘要

在数字病理图像分析中，细胞核的准确检测与定位是肿瘤分级、组织结构分析及辅助诊断等关键任务的基础。然而，实际病理图像中细胞通常呈现高度重叠、密集排列、边界模糊、形态异质性强等特征，传统的目标检测算法难以在复杂背景中实现对细胞核的精细识别与分离。这对检测模型的特征提取能力、局部细节表达以及对重叠区域的解析能力提出了更高要求。针对上述问题，本文提出了一种基于改进 YOLOv11 的重叠细胞识别方法。在 YOLOv11 主干网络中引入倒置残差模块（Inverted Residual Block）以提升非线性建模与高效特征传递能力，同时融合通道注意力机制（Channel Attention, CA）以动态增强关键信息通道的表达能力，从而实现对复杂结构细胞区域的精准识别。在自行标注的重叠细胞图像数据集上进行训练和评估，实验结果表明，改进后的模型在 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 等多个评价指标上均优于原始 YOLOv11、YOLOv8 及 YOLOv12 等主流模型。其中，mAP@0.5 提升至 0.922，mAP@0.5:0.95 提升至 0.488，显示出较强的目标检测性能。消融实验也验证了倒置残差结构与通道注意力机制在模型精度提升方面的有效性与协同作用。本研究为病理图像中重叠细胞的高精度检测提供了一种高效、可扩展的解决方案，对智能病理诊断系统的构建具有重要的参考价值与实际意义。

关键词：YOLOv11、重叠细胞识别、倒置残差模块、通道注意力

目录

一、引言	3
二、相关工作	3
2.1 通道注意力机制	4
2.2 倒置残差模块	4
三、方法论	6
3.1 改进的 YOLOv11 模型	6
3.2 损失函数	7
四、实验	8
4.1 数据集	8
4.2 实验环境	8
4.3 评价指标	8
4.4 消融实验	9
4.5 对比实验	9
五、结论	10

一、引言

细胞识别与计数在病理诊断、肿瘤分析、药物筛选等生物医学领域中扮演着极为重要的角色。随着数字病理与高分辨率显微成像技术的发展，基于图像的自动细胞识别方法成为研究热点。然而，在实际显微镜图像中，细胞之间常常出现不同程度的重叠与粘连，导致边界模糊、结构遮挡，这种重叠细胞现象严重影响了传统图像分割与目标检测方法的识别准确率。

重叠细胞的识别不仅对算法的精度提出了更高要求，也对其空间感知能力、特征提取能力、模型鲁棒性构成挑战。尤其是在多细胞密集区域，细胞核之间轮廓重合、颜色接近，常规方法往往难以准确分辨个体细胞，导致漏检、误检率显著上升。因此，如何高效、准确地识别重叠细胞，成为细胞图像分析领域的关键问题之一，对于推动智能辅助诊断系统的落地具有重要理论价值与现实意义。近年来，深度学习技术，尤其是 YOLO (You Only Look Once) 系列目标检测算法，在自然图像领域取得了卓越成果。YOLOv11 作为该系列的最新版本，在模型结构、特征融合、注意力机制与小目标检测能力方面进行了多项增强，表现出更强的泛化能力与定位精度，为解决重叠细胞识别这一复杂任务提供了可能。

然而，现有医学图像公开数据集中针对重叠细胞的标注仍较为有限，无法直接满足特定检测需求。为此，本研究采用手工逐帧标注的方式，构建了一套高质量的重叠细胞目标检测数据集，涵盖不同形态、密度与重叠程度的细胞图像样本。基于该数据集，本文提出一种改进的 YOLOv11 模型结构，针对显微图像中细胞尺寸小、密集度高、遮挡严重等特点，引入小目标增强模块与特征融合机制，提升模型在重叠区域的检测能力，本研究的主要贡献包括：

- 1、手工标注数据集：**聚焦重叠细胞识别这一细胞图像处理中的核心难题，分析其成因与挑战，并构建了手工标注的专用数据集。
- 2、提出改进模型：**基于 YOLOv11 提出优化检测结构，提升模型对重叠、密集、小尺寸细胞的检测效果。
- 3、实验验证结果：**通过系统实验验证所提方法的有效性，展示其在识别准确率、召回率、泛化能力方面的优势。

二、相关工作

YOLOv11 是 YOLO (You Only Look Once) 系列最新一代的目标检测算法，在保持高速检测能力的基础上，对特征提取网络、特征融合结构和注意力机制等模块进行了进一步优化。与前代模型相比，YOLOv11 引入了更高效的骨干网络、更细致的小目标检测机制以及动态标签分配策略，显著提升了在复杂场景下的检测精度和鲁棒性，YOLOv11 具备更强的定位能力和特征区分能力。

2.1 通道注意力机制

通道注意力机制是一种基于深度学习的特征优化方法，其核心目标是通过显式建模特征通道之间的相互依赖关系，动态调整各通道的权重分布，从而增强模型对关键特征的敏感性，同时抑制噪声或冗余通道的干扰。该模块通过自适应地为各通道分配差异化权重，动态优化特征表征，从而增强网络检测细微结构特征的能力，并表征复杂的影像模式。通道注意力结构图如图 1 所示。

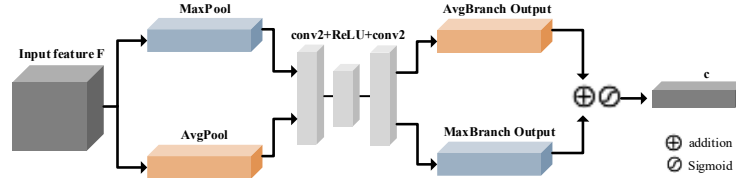


图 1 通道注意力结构图

从图 3 中可以看出，它通过同时使用平均池化和最大池化操作来聚合特征图 \mathbf{F} 的空间信息，生成描述符 $\mathbf{z} \in \mathbb{R}^C$ ，从而消除空间位置的影响，使网络专注与通道间全局关系而非局部细节，数学表达如下：

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \mathbf{F}_c(i, j) \quad (1)$$

其中， $\mathbf{F}_c(i, j)$ 表示第 c 个通道在空间位置 (i, j) 处的特征值， H 为特征图高度， W 为特征图宽度。在获取通道统计描述符 \mathbf{z} 后，通道注意力机制通过两层全连接层，学习通道间的非线性依赖关系。全连接层首先对 \mathbf{z} 进行降维，减少计算开销，再通过非线性激活函数增强表达能力，最后恢复原始通道数，并利用激活函数输出归一化的权重向量 $\mathbf{s} \in [0, 1]^C$ 。数学表达如下：

$$\mathbf{s} = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{z} + \mathbf{b}_1) + \mathbf{b}_2) \quad (2)$$

其中， $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r} \times C}$ 和 $\mathbf{W}_2 \in \mathbb{R}^{C \times \frac{C}{r}}$ 是可学习参数， r 是压缩比（通常设为 16），用于控制模型的复杂度。在得到通道权重向量 \mathbf{s} 之后，将其与原始输入特征 \mathbf{F} 进行逐个通道乘法运算，生成优化后的特征图 \mathbf{M}_c 如下：

$$\mathbf{M}_c = \mathbf{s}_c \cdot \mathbf{F}_c \quad (3)$$

经过注意力机制操作，每个通道进行动态缩放，使重要的特征通道得到增强，不重要的特征通道被抑制，从而提高模型的精度。

2.2 倒置残差模块

在医学图像中，特别是细胞识别任务中，网络需要具备在高分辨率图像中有

效提取细粒度特征的能力。YOLOv11 引入了轻量化设计思想以提升网络速度与准确性，其中倒置残差模块（Inverted Residual Block）是关键组成部分之一，旨在减少计算量的同时提升特征表达能力。倒置残差结构图如图 2 所示。

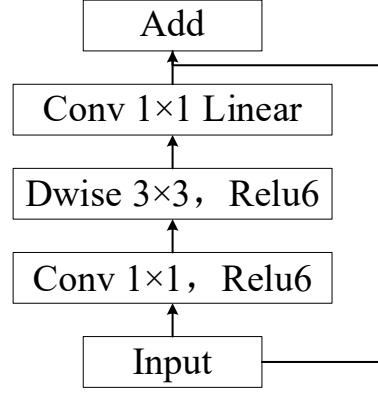


图 2 倒置残差结构图

倒置残差模块最早由 MobileNetV2 提出，其核心思想是通过升维—深度卷积—降维的方式对信息进行非线性建模，并结合残差连接（Residual Connection）增强特征保留与信息流动能力。与传统残差模块“先压缩再扩展”相反，倒置残差模块采用“先扩展再压缩”的策略，从而被称为“倒置”，

具体而言，设输入特征图为 $\mathbf{X} \in R^{H \times W \times C_{in}}$ ，输出特征图为 $\mathbf{Y} \in R^{H' \times W' \times C_{out}}$ ，倒置残差模块的计算过程包含三个步骤：首先，使用 1×1 卷积将输入通道数由 C_{in} 扩展到 C_{out} ，其数学表达式如下：

$$\mathbf{X}_1 = \text{ReLU6}(\text{BN}(\mathbf{X} * \mathbf{W}_1)), \mathbf{W}_1 \in R^{1 \times 1 \times C_{in} \times tC_{in}} \quad (4)$$

其中， $*$ 表示卷积操作，BN 为批量归一化，ReLU6 是激活函数。然后，深度可分离卷积，对升维后的特征图 \mathbf{X}_1 进行逐通道卷积，提取空间特征：

$$\mathbf{X}_2 = \text{ReLU6}(\text{BN}(\mathbf{X}_1 \odot \mathbf{W}_2)), \mathbf{W}_2 \in R^{3 \times 3 \times tC_{in} \times 1} \quad (5)$$

其中， \odot 表示深度卷积操作，该步骤减少参数量并捕获空间信息。通过 1×1 卷积将通道数恢复到 C_{out} ，数学表达式如下：

$$\mathbf{X}_3 = \text{BN}(\mathbf{X}_2 * \mathbf{W}_3), \mathbf{W}_3 \in R^{1 \times 1 \times tC_{in} \times C_{out}} \quad (6)$$

最后通过残差连接当输入输出特征图尺寸和通道数相同时，采用跳跃连接进行融合。通过以上操作，倒置残差模块不仅在高维特征空间增强了非线性表达能力，同时利用残差结构有效避免梯度消失问题，提升了模型的稳定性和准确性。其轻量化设计使其非常适合用于 YOLOv11 中复杂的重叠细胞检测任务，兼顾了

模型性能和推理速度。

三、方法论

本文提出了一种改进的 YOLOv11 网络结构，旨在提升对重叠细胞的识别能力。在主干网络（Backbone）中引入倒置残差模块（Inverted Residual Block）与通道注意力机制（Channel Attention, CA），实现对细胞图像特征的更深层次挖掘与高效表达。

3.1 改进的 YOLOv11 模型

为了提升医学图像中重叠细胞的检测精度，本文提出了一种改进的目标检测模型，其在传统 YOLOv11 框架的主干网络中融合了倒置残差模块（Inverted Residual Block）与通道注意力机制（Channel Attention, CA）。如图 3 所示，在不增加过多计算开销的前提下，有效增强了特征提取能力。

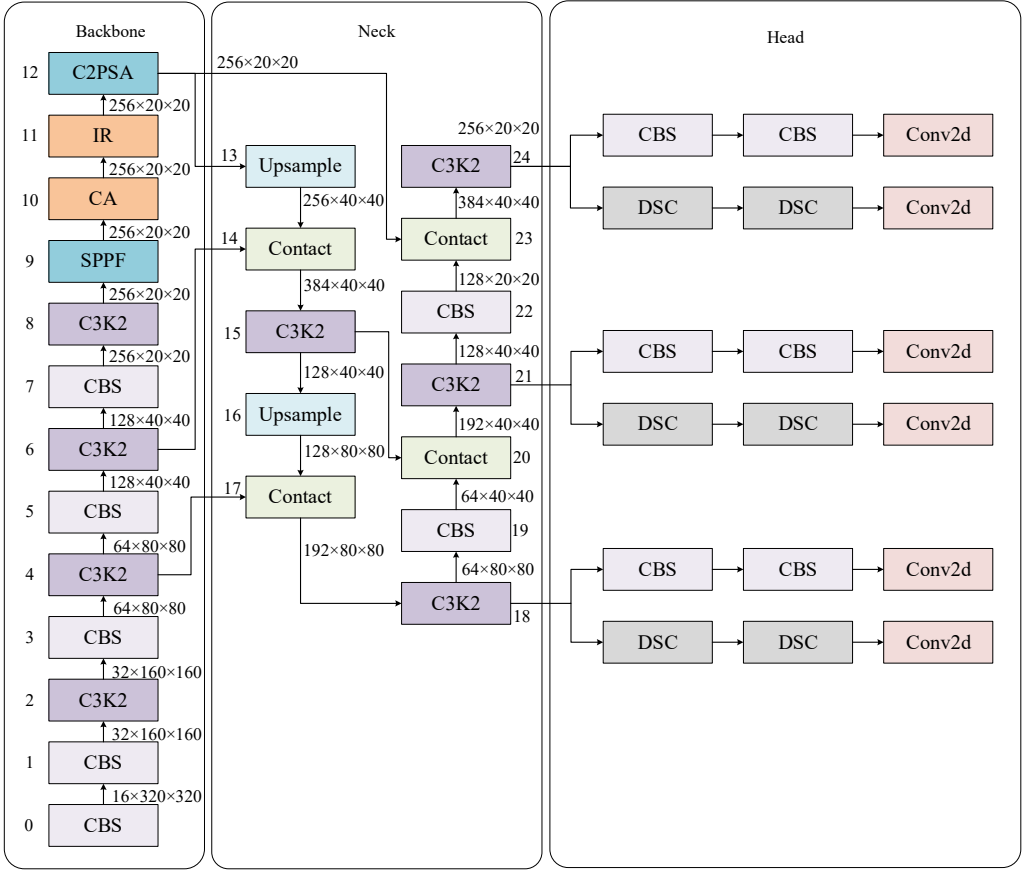


图 3 改进的 YOLOv11 结构图

在该结构中，倒置残差模块被嵌入至 Backbone 的多个阶段，用于替代常规卷积模块。该模块通过先升维、再深度卷积、再降维的方式对输入特征进行高效建模，不仅提升了模型对高维语义特征的表达能力，还通过残差连接机制保留了浅层关键信息，从而有助于准确定位重叠区域内的细胞核。为了进一步提升通道特征的表达效果，我们加入了 CA 模块。该模块通过全局平均池化提取通道维度

上的全局信息，并通过两层全连接结构动态生成每个通道的重要性权重，从而增强关键信息，抑制冗余特征。这种模块级联设计充分结合了空间建模与通道重标定的互补优势：倒置残差模块提升了空间特征的建模能力，CA 模块强化了通道语义表达，使网络能够在不同层次全面提取重叠细胞的显著特征，从而显著提高目标检测性能。

3.2 损失函数

本研究采用 YOLOv11 官方提供的多任务损失函数进行模型训练，目标是同时优化目标定位精度与分类准确率，以满足重叠细胞识别任务对小目标、高密度区域检测的精度要求。YOLOv11 的损失函数主要由以下三部分组成：

边界框回归损失（Box Loss），用于回归预测框与真实框之间的位置偏差，YOLOv11 使用了 CIoU（Complete IoU）损失函数，综合考虑了重叠面积、中心距离和长宽比三方面的信息，其数学表达如下：

$$L_{box} = 1 - CIoU(b, \hat{b}) \quad (7)$$

其中， b 表示预测框， \hat{b} 表示真实框，CIoU 能有效引导模型学习到更加准确且收敛更快的目标框。

类别预测损失（Classification Loss），YOLOv11 使用二元交叉熵损失（BCE）对目标类别进行监督，适用于多类目标或背景检测任务。设预测概率为 p ，真实标签为 y ，则分类损失为：

$$L_{cls} = - \sum_{i=1}^C [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (8)$$

其中 C 表示类别数量，当前任务中为细胞类型分类。

分布式回归损失（Distribution Focal Loss, DFL），为了进一步提升边界框精度，YOLOv11 引入了 分布式回归损失（DFL），该方法将边界框预测为一组离散分布，并采用 soft label 与 cross-entropy 进行训练，优化目标如下：

$$L_{DFL} = \sum_{i=1}^n CE(q_i, \hat{q}_i) \quad (9)$$

其中 q_i 为离散位置的预测分布， \hat{q}_i 为真实位置的离散标签。综上，YOLOv11 的总损失函数可以表示为三者加权和：

$$L_{total} = \lambda_1 L_{box} + \lambda_2 L_{cls} + \lambda_3 L_{dfl} \quad (10)$$

其中, $\lambda_1, \lambda_2, \lambda_3$ 各子损失项的权重系数, 本研究遵循 YOLOv11 默认权重设定, 并在医学图像下对其稳定性与收敛性进行验证, 确保模型在面对重叠密集的细胞结构时具有优异的识别性能。

四、实验

4.1 数据集

本研究使用的数据集为自行构建的重叠细胞图像数据集。该数据集共包含 367 张显微镜图像, 图像大小统一为 1040×1360 像素, 分辨率较高, 有利于保留细胞核的结构细节。所有图像均经过精细标注, 标注内容包括每个重叠细胞的边界以及对应的类别信息。为了充分训练模型并合理评估其性能, 数据集按照 7:1:2 的比例划分为训练集、验证集和测试集。



图 4 数据集样例

4.2 实验环境

实验环境如下: 操作系统为 Ubuntu 22.10, 使用的 GPU 为 RTX-3090。所有实验均基于 Python 3.10 和 PyTorch 2.1.0 深度学习框架进行。优化器采用 Adam, 初始学习率设置为 $1e-3$ 。模型使用 L2 正则化, 权重衰减系数为 $1e-4$ 。所有模型的最大训练轮数设置为 150。

4.3 评价指标

为了全面评价改进后的 YOLOv11 模型在重叠细胞检测任务中的性能, 本文采用 YOLOv11 官方推荐的多种评价指标, 主要包括精准率 (Precision)、召回率 (Recall) 以及平均精度均值 (mean Average Precision, mAP)。数学表达如下:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

$$mAP@0.5 = \frac{1}{N_c} \sum_{i=1}^{N_c} AP_i^{\text{IoU}=0.5} \quad (13)$$

$$mAP@0.5:0.95 = \frac{1}{10 \times N_c} \sum_{\tau=0.5}^{0.95} \sum_{i=1}^{N_c} AP_i^{IoU=\tau} \quad (14)$$

其中, TP 表示真正例数量, FP 表示假正例数量, FN 表示假负例数量。平均精度均值 (mAP) 是目标检测任务中综合评价模型性能的重要指标。本文采用两种常见的 mAP 计算方式, 其中, N_c 表示类别总数, $AP_i^{IoU=0.5}$ 为第*i*类在 IoU 阈值 0.5 时的平均精度。 τ 表示不同的 IoU 阈值, 包含 0.5、0.55、0.6、 \cdots 、0.95 共十个阈值。

4.4 消融实验

为了分析倒置残差模块(Inverted Residual Block)与通道注意力机制(Channel Attention)在改进 YOLOv11 模型中的作用, 我们设计了一系列模块配置不同的消融实验, 以评估各个组件对检测性能的影响。实验共包括四种配置: 原始 YOLOv11 基线模型、仅加入倒置残差模块的模型、仅加入通道注意力模块的模型, 以及同时加入两者的完整改进模型。各组实验结果如表 1 所示。

表 1 消融实验表

模型	P	R	mAP50	mAP50-95
YOLOv11	0.761	0.862	0.893	0.462
YOLOv11+CA	0.891	0.785	0.918	0.482
YOLOv11+IR	0.897	0.855	0.915	0.485
YOLOv11+CA+IR	0.905	0.883	0.922	0.488

如表 1 所示, 当仅在 YOLOv11 中引入通道注意力机制(CA)时, Precision、mAP@0.5 和 mAP@0.5:0.95 分别提升了 13.0%、2.5%和 2.0%。若仅引入倒置残差模块(IR), Recall 显著提升了 0.7%, 同时 mAP@0.5 和 mAP@0.5:0.95 分别提高了 2.2%和 2.3%。当同时引入通道注意力与倒置残差模块(YOLOv11+CA+IR)时, 模型在各项指标上均达到最优, Precision 提升至 0.905, Recall 提升至 0.883, mAP@0.5 和 mAP@0.5:0.95 分别达到 0.922 和 0.488, 均优于其他配置。该结果表明, 通道注意力机制与倒置残差结构在改进 YOLOv11 中具有明显的互补性与协同增强效果, 能够更有效地提升模型在重叠细胞检测中的整体性能。

4.5 对比实验

将提出的改进 YOLOv11 与几种先进方法进行了比较, 包括 SSD、YOLOv8、YOLOv11 和 YOLOv12。为了进行公平比较, 所有方法都在相同的数据集上进行训练, 并使用相同的损失函数。

表 2 对比实验结果

模型	P	R	mAP50	mAP50-95
SSD	-	-	0.865	-
YOLOv8	0.826	0.874	0.908	0.485
YOLOv11	0.761	0.862	0.893	0.462
YOLOv12	0.833	0.846	0.908	0.457
YOLOv11+CA+IR	0.905	0.883	0.922	0.488

表 2 展示了改进后的 YOLOv11（即 YOLOv11+CA+IR）在重叠细胞识别任务中的性能提升。与原始 YOLOv11 相比，改进模型的 Precision 提高了 14.4%，Recall 提高了 2.1%，mAP@0.5 提升了 2.9%，mAP@0.5:0.95 提升了 2.6%。在与 YOLOv12 的对比中，YOLOv11+CA+IR 在 Precision 上提升了 8.6%，Recall 提高了 3.7%，mAP@0.5 提升了 1.4%，mAP@0.5:0.95 提升了 3.1%。图 6 展示了各模型在典型图像上的检测可视化效果。YOLOv11+CA+IR 重叠区域识别方面相较其他方法具有更细致的表现和更高的置信度，进一步证明其在图像识别中的优势。

图 5 预测结果可视化

五、结论

本文针对重叠细胞识别中的目标检测难题，提出了一种改进的 YOLOv11 模型，结合倒置残差模块（Inverted Residual Block）与通道注意力机制（Channel Attention），在保持模型轻量化的同时，有效增强了特征提取与信息表达能力。通过在自建的重叠细胞图像数据集上进行大量实验，结果表明本方法在 Precision、Recall、mAP@0.5 和 mAP@0.5:0.95 等评价指标上均优于原始 YOLOv11、YOLOv8、YOLOv12 等主流检测模型。特别是在重叠区域和细粒度细胞边界的识别方面，改进后的模型表现出更高的准确性和鲁棒性，验证了所引入模块的有效性与协同

优势。未来的研究中，可进一步探索多尺度融合结构与空间注意力机制的结合方式，以进一步提升模型在复杂医学图像中的识别能力与泛化性能。