基于YOLOv11的重叠细胞识别研究

摘要

重叠细胞的精确识别是细胞图像分析中的核心挑战，对于疾病诊断与病理分析等医学应用具有重要意义。传统图像处理方法在面对细胞密集、边界模糊与结构重叠等复杂情况时，往往难以实现鲁棒性好且高效的检测。为应对这一问题，本文提出一种基于改进 YOLOv11 网络的重叠细胞识别方法。该方法在 YOLOv11 主干网络中引入倒置残差模块（Inverted Residual Block），以增强深层语义特征的建模能力，并融合通道注意力机制（Channel Attention, CA），通过动态调节通道权重，提升模型对关键结构的感知能力。此外，围绕任务特点构建了人工标注的专用细胞图像数据集，为网络训练与优化提供了数据支撑。本研究在不依赖额外结构复杂度的前提下，提升了网络对重叠细胞的表达能力与识别潜力，为医学图像中复杂小目标的检测提供了有效的技术路径。

关键词：YOLOv11、重叠细胞识别、倒置残差模块、通道注意力

1. 引言

细胞识别是生物医学图像处理中的一项基础而关键的任务，在细胞计数、癌症检测、药物筛选等多种下游应用中扮演着核心角色。然而，在实际显微图像中，由于细胞间的自然堆叠、密集分布以及形态多样性，细胞之间常常呈现出不同程度的重叠，导致目标边界模糊不清，严重影响了传统图像处理算法的准确性与鲁棒性。针对这一问题，开发一种具备高效分离与精确检测能力的识别方法，具有重要的研究价值与实际意义。

传统的细胞识别方法通常依赖于阈值分割、边缘检测、形态学操作等手工设计的特征提取策略，在面对细胞重叠、尺寸差异大、对比度低等复杂图像特征时往往具有应用瓶颈。近年来，随着深度学习在计算机视觉领域的广泛应用，基于卷积神经网络（CNN）的目标检测算法在医学图像领域逐步展现出强大的特征学习与目标定位能力。其中，YOLO（You Only Look Once）系列算法以其端到端的检测机制和较高的推理效率，受到研究者的广泛关注。特别是最新的YOLOv11架构在模型精度与速度之间取得了更优平衡，并引入多尺度特征融合机制，使其在处理小目标检测和密集场景下更具优势。这为解决重叠细胞识别中的尺寸不均、目标邻近、遮挡严重等难题提供了新的可能。

基于上述背景，本文以重叠细胞识别为研究核心，探索YOLOv11在该任务中的应用潜力与性能优势，围绕数据构建、结构优化等方面展开系统研究。

1. YOLO模型发展背景

目标检测作为计算机视觉中的基础任务之一，长期以来受到广泛关注。自 YOLO（You Only Look Once）系列模型提出以来，端到端的实时检测框架逐渐成为目标检测领域的主流方向。最初的 YOLOv1 将目标检测任务统一为回归问题，具备极高的推理速度，但在定位精度与小目标检测方面存在明显不足。后续版本如 YOLOv2 和 YOLOv3 引入多尺度预测机制与残差结构，有效提升了模型的检测精度与鲁棒性。

YOLOv4 开始系统性集成优化策略，包括 CSPNet 结构、Mish 激活函数与 PANet 特征融合路径，使其在速度与精度之间达到更优权衡。而 YOLOv5 及之后的改进版本（如 YOLOv7、YOLOv8、YOLOv11）则在轻量化、模块设计、多任务融合等方面持续演进，引入诸如自适应锚框、深度可分离卷积、注意力机制、NAS（神经架构搜索）等先进技术，使其更适应资源受限场景与复杂图像结构的处理需求。

在医学图像分析任务中，尤其是细胞图像中，目标往往具备尺寸小、密度高、边界模糊、重叠复杂等特点，对检测模型的特征建模能力与定位精度提出了更高要求。YOLO 系列模型因其结构紧凑、检测速度快、可灵活扩展，在此类任务中展现出良好的适用性。特别是随着 YOLOv11 引入更精细的结构设计，其在特征提取深度与计算效率之间取得更佳平衡，为细粒度细胞检测提供了坚实的基础。

因此，本文基于 YOLOv11 框架，融合结构级与通道级增强机制，旨在进一步挖掘该系列模型在复杂医学图像场景中的应用潜力，解决重叠细胞识别中精度不足与特征弱化等问题。

1. 相关工作

近年来，深度神经网络在图像目标检测任务中展现出强大性能，而细粒度目标识别对模型的特征表达能力提出了更高要求。为增强模型对关键特征的关注能力并提升特征传播效率，研究者们提出了多种结构性改进方法，其中通道注意力机制与倒置残差模块作为代表性结构，被广泛集成于目标检测框架中，并取得了良好效果。

2.1 通道注意力机制

通道注意力机制作为一种深度特征重加权策略，其设计目的在于展现建模通道维度上的特征重要性，强化关键语义表达，削弱无效或干扰信息的影响。该机制通过学习通道间的全局依赖关系，实现对输入特征图的动态加权，从而提升网络对复杂结构、微小差异的识别能力，尤其适用于细粒度识别场景如重叠细胞的检测。

具体实现中，该模块对输入特征图采用全局平均池化与最大池化两种策略整合空间域内的统计信息，分别生成两个一维的通道描述符，消除位置信息的干扰，使网络更专注于通道级别的全局特征建模。平均池化计算如下：

其中，表示第个通道在空间位置处的特征值，和分别为特征图的高度与宽度。

在获取通道统计描述符后，将其输入到包含两个全连接层的子网络中，进行非线性映射与维度压缩操作。第一个全连接层将通道数降至 C/r（其中r为压缩因子，常设为16），以减少参数量，第二个全连接层再将其恢复至原始通道维度。通过 ReLU 和 Sigmoid 激活函数的联合作用，最终输出归一化的通道权重向量，过程如下：

其中，和是可学习权重，在得到通道权重向量之后，将其与原始输入特征进行通道维度上的逐元素乘法，实现对各通道特征的选择性增强或抑制：

通过这一机制，模型可自适应调整不同通道的响应强度，显著提升对重叠细胞边缘、结构及纹理等微小差异的感知能力，为后续检测任务提供更具判别性的特征表示。

2.2 倒置残差模块

现代神经网络的发展推动了模型表达能力的不断提升，但计算量也随之显著上升。在医学图像分析这种有大量数据基础的应用场景中，模型不仅需要高精度，还必须兼顾计算效率。倒置残差模块作为轻量化网络设计的典范，在保持较低计算开销的同时，实现了对复杂语义信息的有效建模。

YOLOv11将倒置残差结构引入其主干与特征融合路径中，有效解决了深层网络中信息流动受阻与梯度衰减的问题。其先升维再深度卷积最后又降维的结构策略，既提升了模型的非线性建模能力，又确保了信息在网络中的高效传递，为实现快速而精确的细胞检测提供了基础。

设输入特征图为，输出特征图为，计算过程可分为三个步骤：首先，使用1×1卷积对通道数进行升维，从原始扩展到，引入更高维度的表达空间：

其次，对升维后的特征图进行深度可分离卷积（depthwise convolution），仅在空间维度内施加卷积，减少参数数量同时捕获局部空间特征：

其中，表示深度卷积操。第三步，采用另一个 1×1 卷积将通道数压缩回，得到最终的输出特征图：

当输入与输出的特征尺寸一致时，模块引入跳跃连接（skip connection）将输入与输出相加，形成残差路径，从而增强特征传递效率和模型稳定性。

倒置残差模块在不显著增加计算负担的前提下，提升了网络的非线性建模能力，尤其适用于重叠细胞的检测任务中对小尺度、密集结构的建模需求。其轻量化与高效性特征，使其能够在YOLOv11架构中有效提升检测精度的同时，维持较高的推理速度，适应医学图像场景对实时性与准确性的双重要求。

1. 方法设计

为增强 YOLOv11 在重叠细胞识别任务中的特征提取能力，本研究对其主干网络进行了结构性优化，在保持模型整体轻量化特点的基础上，引入了 倒置残差模块（Inverted Residual Block）与通道注意力机制（Channel Attention, CA），以提升模型对复杂结构特征的建模能力与关键语义的聚焦能力。

3.1 YOLOv11模型改进结构设计

在 YOLOv11 主干网络（Backbone）中，原始的部分常规卷积结构被替换为倒置残差模块。该结构可以在特征变换过程中构建非线性高维表示，与此同时，残差连接机制保留了浅层特征信息，有效缓解了深层网络中梯度消失及特征退化等问题。该结构在提升深层语义建模能力的同时，也更好地保留了浅层结构信息，使得模型在处理重叠区域的细胞核时能够获得更准确的空间定位。

为进一步加强通道维度上的特征选择能力，模型在关键特征提取阶段引入了通道注意力机制（Channel Attention）该模块通过提取通道级别的全局统计信息，动态调整各通道的重要性权重，使网络更加聚焦于与细胞轮廓、核形态等关键结构相关的特征表达，从而有效抑制背景噪声或冗余干扰。

上述两类模块的结合充分发挥了其在不同维度上的特征建模优势：倒置残差模块提升了模型在空间域中的结构提取能力，通道注意力机制则强化了语义通道的表达深度。两者在模型中的协同嵌入，构建出一种既具表达能力又具计算效率的细粒度识别结构，显著提高了网络在重叠、密集、小尺寸细胞等复杂场景下的检测性能。

1. 结论

本研究聚焦于细胞图像分析中的关键难题：重叠细胞的自动识别与检测。该类任务在实际医学诊断、病理分析等场景中具有重要应用价值，但由于目标密集、边界模糊、结构多变，传统方法往往难以取得理想的检测效果。为解决上述问题，本文在 YOLOv11 网络框架的基础上，提出了一种面向重叠细胞检测的结构优化方案。具体包括：一方面，通过引入倒置残差模块，增强网络对高维语义信息的建模能力，同时保留浅层关键特征，有效改善细胞重叠区域的边界表达；另一方面，结合通道注意力机制动态调整特征通道权重，提升模型对显著区域的关注度，从而提高检测的鲁棒性与准确性。此外，围绕任务特性构建了人工标注数据集，为该类目标检测提供了基础支撑。未来的工作可进一步探索多尺度融合结构、跨模态图像特征融合等策略，以进一步提升模型在复杂医疗图像场景中的适应性与推广能力。