

面向高效率细胞注射的实例分割视觉检测方法

刘国志, 李 涛, 康升征, 周 杰

(南京信息工程大学自动化学院, 南京 210044)

摘 要: 针对细胞显微注射过程中细胞与注射器针尖尺度差异大、传统方法难以同时实现高精度检测与高效推理的问题, 本文提出一种面向细胞与注射器实例分割的多分支高效轻量化细节增强模型(Diverse efficient lightweight detail-enhanced YOLO, DELD-YOLO)。该模型基于 YOLO11 架构, 在主干网络中引入结构重参数化的多样性分支块以增强特征表达能力; 设计了一种轻量化多尺度特征金字塔结构以提升跨尺度特征融合效果; 同时构建了基于细节增强卷积的轻量化分割头, 在保证分割精度的同时显著降低计算开销。实验结果表明, DELD-YOLO 在细胞与注射器分割任务中均优于原始 YOLO11 模型, 在交并比(Intersection over Union, IoU)阈值为 0.5:0.95 条件下, 细胞与注射器的平均精度分别达到 98.7%和 75.2%。此外, 模型参数量仅为 2.08M, 计算量为 10.0GFLOPs, 具有良好的精度和效率平衡, 适用于自动化细胞注射及嵌入式视觉应用场景。

关键词: 微注射; 实例分割; DELD-YOLO; 计算效率; 细胞与针尖检测

中图分类号: TP242; TP391 文献标志码: A

Instance Segmentation for Vision Detection in High-Efficiency Cell Injection

LIU Guozhi, LI Tao, KANG Shengzheng, ZHOU Jie

(School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: To address the challenge of accurately and efficiently detecting cells and injection needle tips with large scale discrepancies in automated microinjection, this paper proposes a high-efficiency instance segmentation model named DELD-YOLO. Built upon the YOLO11 architecture, the proposed model incorporates structurally re-parameterized diverse branch blocks into the backbone to enhance feature representation. An efficient lightweight multi-scale feature pyramid network is introduced to improve cross-scale feature fusion, and a detail-enhanced lightweight segmentation head is designed to preserve fine structural information while reducing computational cost. Experimental results demonstrate that DELD-YOLO outperforms the baseline YOLO11 in both cell and injector segmentation tasks, achieving mAP values of 98.7% for cells and 75.2% for injectors at IoU thresholds from 0.5 to 0.95. With only 2.08M parameters and 10.0 GFLOPs, the proposed model exhibits a favorable balance between accuracy and efficiency, making it well suited for automated cell injection and edge-device deployment.

Key words: microinjection; instance segmentation; DELD-YOLO; computational efficiency; cell and needle tip detection

引 言

显微注射属于细胞微操作的一种, 是生物学研究中一种将外源物质导入生物样品的普遍操作, 被广泛应用于疫苗研发、药物筛选、基因研究等方面^[1]。传统的细胞注射任务都是由专业人士进行手动完成, 难以保持稳定的注射效率和可重复性。随着

生命科学方面需求变高以及工业自动化程度的提高, 要求细胞显微注射渐渐朝着自动化、高质量发展。于是人们在显微注射的系统设计^[2]、控制^{[3][4]}、视觉^{[5][6]}等方面做出了众多努力。但是在视觉方面, 传统的图像处理的精度已经不能满足如今高质量的细胞注射需求, 对于细胞的检测也提出更快速, 更精确的要求。

基金项目: 国家自然科学基金(62373195; 52405117); 江苏省研究生科研与实践创新项目(SJ CX25_0495)。

对于细胞的检测已经有阈值化、基于区域的方法、模板匹配、Hough 变换等传统的图像处理算法^[7-8]。如文献^[5]采用灰度阈值变换实现图像二值化,将斑马鱼从其图像背景中区分出来。再通过滤波、二值化、图像分割、角度矫正得到清晰的标准斑马鱼幼虫图像,最终通过边缘特征来确定注射目标位置。Guo 等^[6]介绍了一种新颖的基于自动阈值和过度膨胀的视觉识别算法,以准确识别斑马鱼胚胎和幼鱼卵黄的中心。尽管这些方法在能够大致识别细胞,但是在细胞显微注射的动态系统中容易受到气泡、杂质等因素的影响,进而导致实验失败^[9]。而且传统图像处理算法在检测效果和速率上仍存在不足,于是基于深度学习的目标检测的发展为此提供了新的研究思路。Mori 等^[10]提出一种基于力和视觉呈现的辅助系统,利用高斯混合模型(GMM)和长短期记忆(LSTM)等人工智能技术简化了细胞旋转操作。Qian 等^[11]提出一种基于卷积神经网络的生物样品操控和显微注射系统的视觉处理策略,在保证较高注入精度的情况下实现高通量注入。Pan 等^[12]使用深度卷积神经网络 Mask R-CNN 对非染色标记贴壁细胞进行检测和分割,显著提高了注射系统运算效率和后续吞吐量。Li 等^[12]使用基于深度学习的 Faster R-CNN 目标检测和传统图像处理算法,得到注射目标和蛋黄中心的精确坐标。尽管基于深度学习的视觉检测提高了细胞注射的准度和效率,但是大部分基于 R-CNN 的视觉检测速度较低,难以满足实时性目标检测的需求。

随着深度学习技术的不断进步和现实环境的复杂度不断提升,目标检测领域出现以 YOLO^[13]为代表的一阶段算法。一阶段算法具有更快的检测速度和较高的准确度。如今 YOLO 因其高效的性能已被使用在自动驾驶车辆检测^[14]、无人机图像中小目标检测^[15-20]和农作物检测^[21]、缺陷检测和分割^[22-23]、动物图像分割^[24-25]、火情检测分割^[26]等多个领域。在细胞注射视觉检测方面,YOLO 算法也逐渐被人们所使。Sadak 等^[29]使用 YOLOv2_Resnet-50 识别斑马鱼细胞,使得检测细胞具有较高的实时性和准确性。但是检测针尖却使用有约束的自适应直方图均衡化、高斯滤波器、自适应阈值法、Canny 边缘检测等一系列算法,具有较低的效率和实时性。为此,Xu 等人^[27]提出了一种基于深度学习视觉定位的自动显微注射辅助系统。其采用基于 YOLOv8 的视觉定位算法进行细胞和微注射器针尖的实时检测。然而,由于微注射器针尖尺寸($\sim 10\ \mu\text{m}$)与细胞尺寸($\sim 100\ \mu\text{m}$)之间存在显著的尺度差异,该系统的检测算法无法自适应调节感受野,导致微注

射器针尖检测精度明显下降。而且该算法采用边界框进行粗略定位,难以捕捉细胞的精准轮廓和准确的注射器针尖位置,这为实现高精度的细胞注射和高细胞存活率带来了极大的挑战。

随着人工智能、物联网及边缘计算技术的飞速发展,对计算设备的形态与能力提出了全新的要求。传统上,复杂的数据处理和智能算法严重依赖体积庞大、功耗高昂的中央服务器或台式计算机,这在实时性要求高、空间受限或需部署于现场的应用场景中面临显著瓶颈。近年来,以 NVIDIA Jetson、树莓派等为代表的微型计算机的崛起,正彻底改变这一格局。这些嵌入式设备虽仅手掌大小,却集成了强大的处理器、专用 AI 加速器及丰富的 I/O 接口,足以支撑复杂的机器视觉、深度学习和实时分析任务。为在这些微型计算机有限的资源下实现强劲的算力和高效的实时性,这对细胞检测算法的效率也提出了很大的挑战。

为解决显微注射场景中细胞与注射器针尖尺度差异大、边缘设备计算资源受限以及传统检测方法难以兼顾精度与效率的问题,本文围绕高效实例分割模型设计开展研究,主要创新点如下:

(1) 在 YOLO11 主干网络中引入结构重参数化的多样性分支块,通过训练阶段的多分支特征增强与推理阶段的等效单分支转换,在不增加推理开销的前提下显著提升模型对显微图像中小尺度注射器目标的特征表达能力。

(2) 设计了一种高效的轻量化多尺度特征金字塔结构,通过多条跨尺度连接与自适应特征融合机制,加强不同尺度特征之间的信息交互,在降低参数量的同时有效提升多尺度目标的分割精度。

(3) 提出了一种基于细节增强卷积与共享卷积机制的轻量化分割头,强化模型对细胞边缘及注射器针尖等高频细节特征的建模能力,在保证分割精度的同时显著减少模型参数量与计算复杂度。

1 DELD-YOLO算法

本文提出了一种基于 YOLO11 的多分支高效轻量化细节增强模型 DELD-YOLO,其整体架构如图1所示。该模型在骨干网络中引入结构重参数化技术,通过多样性分支块(Diverse branch block, DBB)^[28]增强多尺度特征提取能力。为进一步提升特征融合效果,在骨干网络后接入一种高效的轻量化多尺度卷积特征金字塔结构(Efficient lightweight multi-scale feature pyramid network, ELMSFPN),该模块融合高效卷积与异构核选择机制,增强跨尺度信息交互。同时,设计了一种轻量化的细节增强分割头(Lightweight detail-enhanced segmentation head,

LSDESH)，结合细节保持卷积与参数共享机制，在保持分割精度的同时显著降低计算复杂度。所提出的模型实现了细胞与注射器的高效精准实例分割。

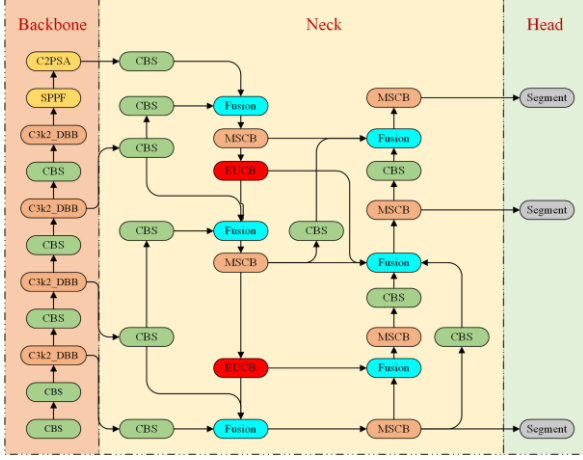


图 1 DELD-YOLO 网络结构

Fig.1 DELD-YOLO network structure

1.1 多样性分支块

YOLO11 的主干网络采用单分支结构，难以在保持轻量化的同时实现精度的最大化，尤其限制了在边缘设备上对小目标注射器进行精准检测的性能。为应对这一挑战，本研究在 YOLO11 主干的 C3k2 模块中引入了再参数化技术，并结合多样性分支块模型对原始主干结构进行增强，以提升显微图像中注射器目标的检测能力。

在训练阶段，DBB 结构借鉴了 Inception 网络的多分支设计理念，支持结构重新参数化。该模块融合了多尺度卷积、序列卷积、平均池化以及分支相加等多种操作，通过组合不同尺度和复杂度的卷积分支，有效扩展了模型的感受野范围。这种多分支架构所带来的多样性与泛化性，显著增强了单个卷积层的表征能力，进而生成更加丰富和多维的特征表示。训练完成后，DBB 可等价转换为单一卷积层用于推理，从而在不引入额外计算开销的前提下，提升模型的表达能力。在实际应用中，可将现有网络中的常规卷积层替换为 DBB 模块，构建更为复杂的混合结构进行训练，并在推理阶段恢复为原始架构，实现训练复杂性与推理效率的兼顾。DBB 通过在保持宏观结构不变的前提下增强训练阶段的微观复杂性，使其成为一种通用的、即插即用的卷积替代方案，有助于显著提升模型性能。

通过整合多尺度操作，DBB 有效增强了模型对复杂模式与细微特征的捕捉能力，构建出判别性更强的特征空间，从而显著提高了模型的泛化性能与鲁棒性，使其更适用于真实场景下的多样化数据及

未知样本的处理。

在 C3k2 模块中，本文将原 Bottleneck 结构中的常规卷积替换为 DBB 模块，具体结构如图 2 (c) 所示。该 Bottleneck 由两个 DBB 模块及跨层短路连接构成，增强了梯度流动并促进了更深层次特征融合。改进后的 C3k2 模块具备以下优势：(1) 在训练阶段可通过更复杂的结构提升模型表征能力与检测性能；

(2) 在推理时能够等价转换为轻量级模块，保持高性能的同时不影响推理速度。如图 2(a) 所示，C3k2-DBB 架构将 C3k2 与 DBB 有机结合，使用 DBB 替换 Bottleneck 中的传统卷积操作。该设计在维持模型轻量化特性的基础上，显著提升了对低分辨率显微图像中小尺寸目标（如注射器）的检测能力。通过融合多尺度特征与增强特征多样性，C3k2-DBB 能够生成更具判别性的特征表示，从而在微纳操作与机器人视觉任务中实现优越的小目标检测性能。此外，该结构还强化了模型的适应性与鲁棒性，有利于在计算资源有限的边缘设备中进行高效部署，满足实时检测的应用需求。

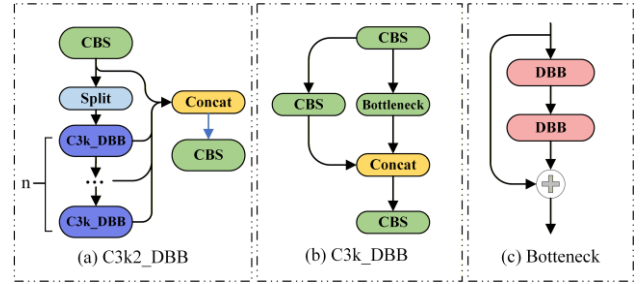


图 2 详细的 C3k2-DBB 结构图

Fig.2 Detailed C3k2-DBB structure diagram

1.2 高效的轻量化多尺度卷积特征金字塔结构

高层次特征蕴含丰富的语义信息，有助于大目标的识别；而低层次特征则保留了丰富的空间细节信息，对于小目标的检测至关重要。特征金字塔网络（FPN）通过一条自上而下的路径，将高分辨率的浅层特征与强语义的深层特征进行融合，从而实现不同尺度目标的准确预测^[30]。然而，在原始 FPN 中，低层特征到顶层特征的路径较长，且融合方式可能不够直接，增加了准确定位信息传递的难度。

为此，本文设计了一种高效的轻量化多尺度特征金字塔结构（ELMSFPN）。该结构融入了高效多尺度卷积模块、高效上采样操作、全局异构核选择机制以及自适应融合策略，以实现高效的多尺度特征融合。同时，为提升模型效率，本文借鉴 BiFPN^[31] 的跨尺度连接思想，引入了额外的自下而上路径，加强了不同层级节点间的信息交互与融合能力。

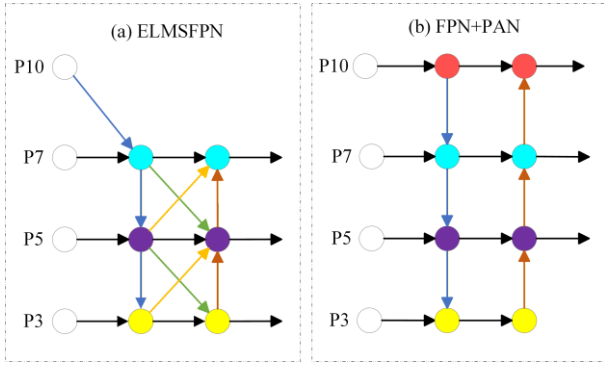


图 3 两种不同的 Neck 结构

Fig.3 Two different Neck structures

图像经过 YOLO11 主干的提取后, 得到 8 个语义信息逐渐变高的主要特征层(P3 ~ P10), 选择经过三个 C3k2_DBB 和 C2PSA 处理过的特征层(P3、P5、P7、P10)交给 ELMSFPN 网络做进一步处理。ELMSFPN 结合跨尺度连接和自适应特征融合, 不

$$P_5^{out} = \text{Conv} \left(\frac{\omega_1' \cdot \text{Resize}(P_3^{td}) + \omega_2' \cdot P_5^{td} + \omega_3' \cdot \text{Resize}(P_3^{out}) + \omega_4' \cdot \text{Resize}(P_7^{td})}{\omega_1' + \omega_2' + \omega_3' + \omega_4' + \theta} \right) \quad (5)$$

式中: P_5^{in} 为第 5 层的输入特征, P_5^{td} 为自上而下路径中第 5 层的中间特征, P_5^{out} 为自下而上路径中第 5 层的输出特征。 ω_i 和 ω_i' 表示不同路径输入特征对应的可学习权重参数。该权重并非人工设定, 而是在模型训练过程中作为网络参数通过反向传播进行端到端学习。为保证数值稳定性并避免不同路径权重尺度不一致的问题, 本文对路径权重进行归一化约束, 使其在融合过程中能够根据不同尺度特征对目标检测与分割任务的贡献自适应分配权重。该机制避免了固定融合比例带来的不确定性, 有助于提升多尺度特征融合的鲁棒性与表达能力。 $\theta = 0.0001$ 是一个很小的值, 以避免数值不稳定。

“Conv” 通常是卷积操作, “Resize” 是用于分辨率匹配的上采样或下采样操作^[45]。

1.2.1 有效的上卷积块

在 YOLO 的特征融合网络中, 默认采用双线性插值^[32]作为上采样方法。该方法通过线性加权平均周围像素值实现上采样, 具有计算简单、结果平滑的优点。然而, 在分割任务中, 双线性插值可能导致边界模糊, 对掩码精度产生一定影响, 并在某些情况下限制模型对细微结构的捕捉能力, 从而降低对小目标的识别与分割精度。

为提升上采样过程中的细节还原能力, 本文引入了一种高效上卷积块 (Efficient Up-convolution Block, EUCB), 以逐步对当前阶段的特征映射进行上采样, 使其尺寸和分辨率与下一跳跃连接所对应的特征映射相匹配^[33]。EUCB 首先使用缩放因子

断进行特征融合, 进一步加强特征提取。以第 5 层 (P5) 特征融合为例, 原来的 FPN+PAN 结构如图 2(a) 所示, 合成形式如下

$$P_7^{td} = \text{Conv}(P_7^{in}) \quad (1)$$

$$P_5^{td} = \text{Conv}(P_5^{in} + \text{Resize}(P_7^{td})) \quad (2)$$

$$P_5^{out} = \text{Conv}(P_5^{td} + \text{Resize}(P_3^{out})) \quad (3)$$

改进后的 ELMSFPN 结构合成形式如下

$$P_5^{td} = \text{Conv} \left(\frac{\omega_1 \cdot P_5^{in} + \omega_2 \cdot \text{Resize}(P_7^{td})}{\omega_1 + \omega_2 + \theta} \right) \quad (4)$$

为 2 的上采样操作扩大特征图尺寸; 随后采用 3×3 深度可分离卷积、批量归一化以及 SiLU 激活函数对上采样后的特征进行增强; 最后通过 1×1 卷积调整通道数, 使其与下一阶段的特征维度一致。EUCB 的结构如图 4 所示。由于使用深度可分离卷积替代标准 3×3 卷积, EUCB 在保持较强特征表达能力的同时显著降低了计算开销, 具备较高的效率。

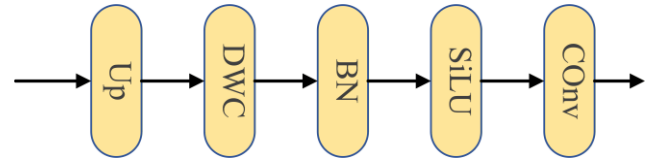


图 4 EUCB 结构

Fig.4 EUCB structure diagram

1.2.2 多尺度卷积模块

本文引入了一种高效的多尺度卷积块 (Multi-scale convolution block, MSCB), 以增强级联扩展路径所生成的特征表征能力。MSCB 借鉴了 MobileNetV2 中反向残差块的设计思想^[34], 但在结构上有所区别。具体而言, 该模块首先使用逐点卷积扩展通道维度, 并经过批量归一化与 SiLU 激活函数处理; 随后通过多尺度深度卷积 (Multi-scale depth-wise convolution, MSDC, 结构如图 5 所示) 提取多尺度与多分辨率的上下文特征; 最后再利用逐点卷积与批量归一化将通道数恢复至原始维度, 该过程同时编码了通道间的依赖关系。MSCB 的大致结构如图 6 所示。

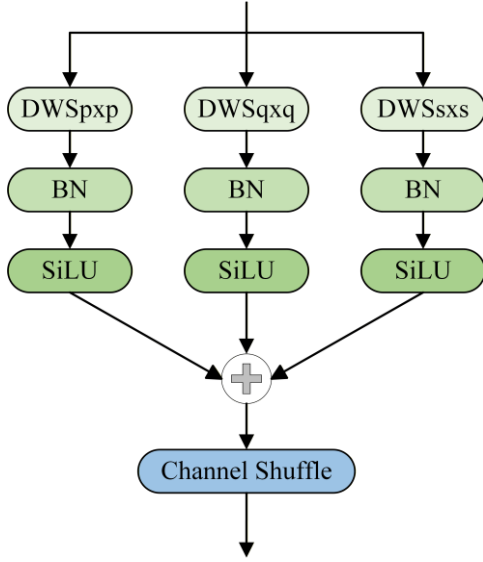


图 5 MSDC 结构

Fig.5 MSDC structure diagram

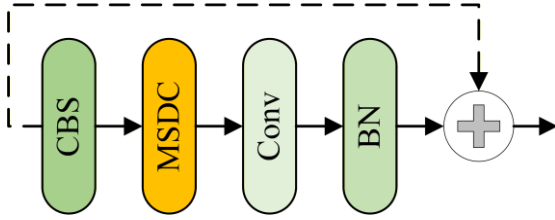


图 6 MSCB 结构

Fig.6 MSCB structure diagram

大卷积核有助于捕获更大范围的局部与全局特征，适度增大卷积核尺寸以扩展有效感受野的做法已在多项研究中被证明是有效的。通常而言，具备较大感受野的网络更适用于检测大尺寸目标，而较小感受野则更有助于捕捉小尺度目标。YOLO-MS^[42]提出了异构核选择协议，在主干网络中采用逐渐增大的卷积核（如 3、5、7、9）以平衡性能与推理速度。受该工作的启发，我们进一步提出全局异构核选择机制^[36]，将异构大卷积核结构集成至 MSCB 模块中。具体地，我们在 MSDC 子模块中引入多组卷积核组合（包括 1/3/5、3/5/7、5/7/9），以适应不同分辨率的特征图需求，从而逐步获取多尺度感受野信息。该模块通过将大卷积核与多个小卷积核并行组合，在不引入额外推理开销的前提下扩展了模型的感知范围，同时保留了对小目标细节的敏感度。这一设计有效兼顾了显微图像中较大目标（如细胞）与极小目标（如注射器）的同步识别需求。此外，由于深度卷积操作忽略了通道间的相关性，本文在模块最后引入通道洗牌操作^[35]，以增强跨通道的信息交互与整合。

其中，不同内核大小（K）的并行 MSDC（如图 6 所示）可由式 6 表示

$$MSDC(x) = \sum_{k \in K} DWCB_k(x) \quad (6)$$

$$DWCB_k(x) = SiLU(BN(DWC_k(x))) \quad (7)$$

式中： $DWC_k(\cdot)$ 是一个深度可分离卷积，其核大小为 k ； $BN(\cdot)$ 和 $SiLU(\cdot)$ 分别是批量归一化和 SiLU 激活。此外，本文的序列 $MSDC(\cdot)$ 使用递归更新的输入 x ，其中输入 x 与之前的 $DWCB_k(\cdot)$ 残差连接，以便更好地正则化，如在式 7 中所示

$$x = x + DWCB_k(x) \quad (8)$$

1.2.3 多尺度特征加权融合

在特征融合网络中，融合不同分辨率的特征通常需要先将它们调整至相同分辨率，再进行聚合操作（如相加或拼接）。常见的融合方式包括加权融合与级联融合等。级联融合^[37]是一种顺序或分层式的融合策略，适用于存在顺序依赖或层次结构关系的模型，但其结构灵活性较差，难以适应数据或任务需求的快速变化。加权融合^[31]则通过对不同来源的特征分配固定权重，再进行加权求和，实现线性融合。该方法适用于输出可直接相加的场景，然而在不同分辨率的输入特征对输出贡献通常不相等的情况下，固定权重难以有效捕捉模型输出之间复杂的非线性关系。

为解决上述问题，本文在 ELMSFPN 中引入了一种自适应特征融合策略^[38]。该策略以形状为（BS, C, H, W）的输入特征为基础，通过卷积、拼接和 Softmax 操作生成对应尺寸为（BS, 3, H, W）的空间自适应权重矩阵。其中 3 个权重通道分别与 3 个输入特征一一对应，通过加权求和实现上下文信息的自适应聚合，从而更有效地融合多尺度特征。该模块结构如图 7 所示。

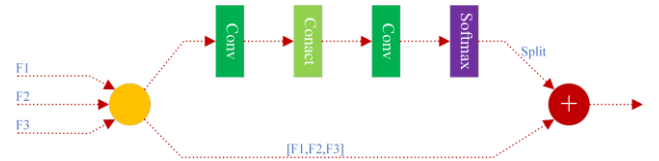


图 7 自适应特征融合

Fig.7 Adaptive feature fusion

1.3 轻量化的细节增强分割头

YOLO 系列模型中的分割头设计借鉴了 YOLACT 算法^[39]的思想，将实例分割任务解耦为两个并行分支：检测分支负责生成边界框及类别信息，分割分支则生成一组原型掩码；最终通过线性组合掩码系数与原型掩码，得到每个实例的分割结果。尽管该范式在通用场景中表现良好，但在复杂显微图像中对高频细节（如细胞精细边界与注射器针尖）

建模仍显不足。分析其根因在于图像中的光照与颜色变化通常属于低频信息，对背景建模较为重要；而物体边缘与轮廓等高频率细节，则对精确分割微小目标（如注射器针尖）至关重要。传统卷积操作往往更倾向于编码低频成分，容易忽略对分割任务关键的高频细节，导致对细胞与注射器等目标的精细边界捕捉不足。

为改善上述问题，本文对 YOLO11 的分割头进行了重新设计，提出了一种基于细节增强卷积（Detail-enhanced convolution, DEConv）^[40]的分割结构。该模块能够更有效地提取和保持高频特征，从而增强对微小边缘和复杂轮廓的敏感性。此外，我们还在分割头中引入共享卷积机制，在保持特征表达能力的同时显著减少参数量，进一步实现了模型在精度与轻量化之间的平衡。

DEConv 模块由 5 个并行部署的卷积层构成，包括 4 个差分卷积层和 1 个普通卷积层，用于协同增强特征提取能力。具体而言，该模块融合了 4 种结构不同的差分卷积操作：中心差分卷积（Central difference convolution, CDC）、角度差分卷积（Angular difference convolution, ADC）、水平差分卷积（Horizontal difference convolution, HDC）和垂直差分卷积（Vertical difference convolution, VDC），从而将传统局部描述符的思想引入卷积过程中，以提升模型的表示能力与泛化性能。

此外，为进一步优化模型效率，本文对并行卷积结构学习到的核权重进行了重新参数化处理。通

过施加适当的约束条件，并利用卷积运算的线性特性，将 5 个并行卷积层等效合并为 1 个单一的标准卷积层。该方法显著减少了参数量，同时加速了模型的训练与推理过程。原理如图 8 所示。因此，DEConv 可以提取丰富的特征以提高分割性能，同时保持与标准卷积相同的参数数量和计算成本。

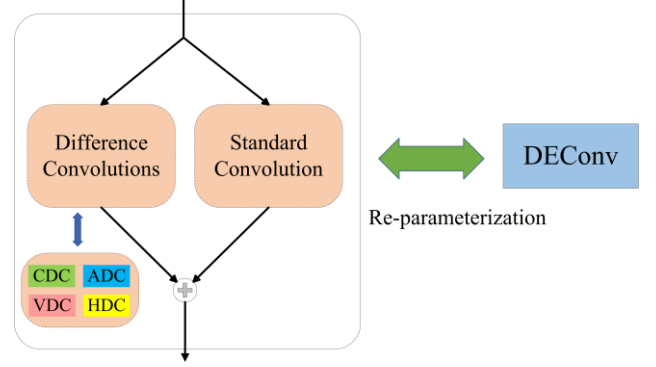


图 8 细节增强卷积

Fig.8 Detail-enhanced convolution

在分割头中，本文在每个普通卷积操作后均引入了组归一化（Group normalization, GN）。该设计的动机源于 FCOS^[41]的研究结论，其指出在分类与回归头中移除 GN 会导致模型性能显著下降。尽管分类与回归分支采用共享卷积结构以提升参数效率，但由于不同检测头所需处理的目标尺度存在差异，本文额外引入了一个 Scale 层，专门用于对回归分支的特征进行尺度缩放，以适应多尺度目标的检测需求。最终构建的 LSDESH 分割头结构如图 9 所示。

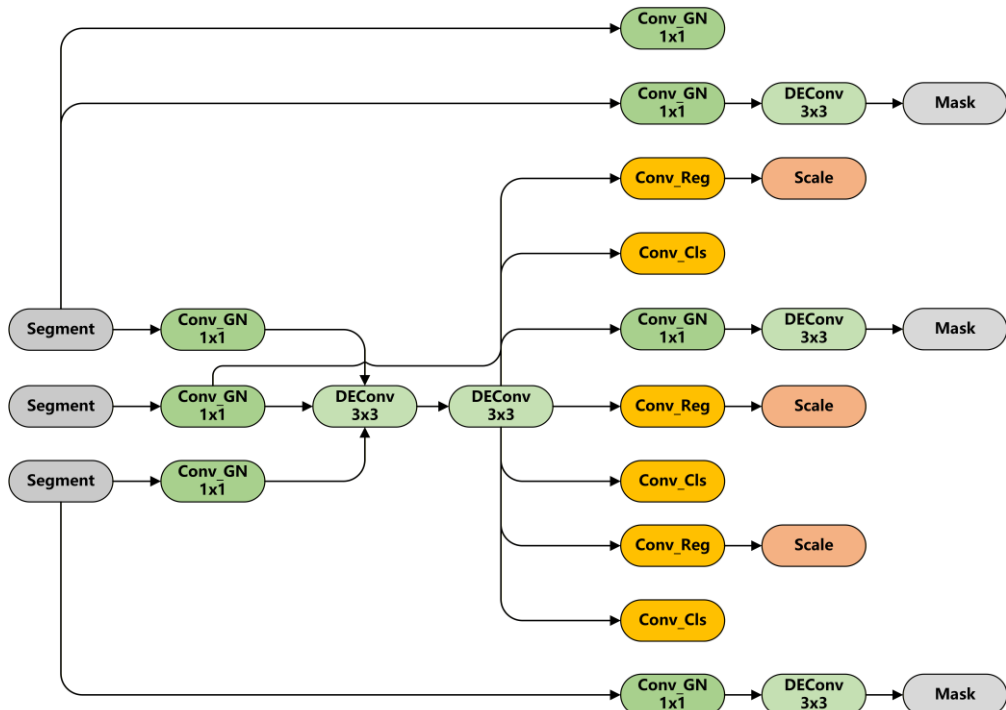


图 9 轻量化的细节增强分割头

Fig.9 LSDESH

2 实验结果与讨论

2.1 实验平台

整个系统的硬件配置如图 10 所示,主要由显微图像采集模块、运动控制模块与主控处理模块三部分构成。显微图像采集模块包括一台光学显微镜(型号:ZY-HDMI2800)和一台高分辨率相机,用于实时获取显微视野下的图像数据。相机的原始分辨率为 1920×1080 ,该分辨率对应相机传感器的物理采样能力。为满足网络输入尺寸及计算效率要求,所有原始图像在预处理阶段被统一缩放至 640×640 ,该尺寸作为后续模型训练与推理阶段的图像尺寸。细胞固定平台采用培养皿与载玻片结构,用于稳定放置并固定待操作的细胞样本。

运动控制模块采用宏-微复合驱动策略:宏驱动平台负责执行大范围位移定位,而末端执行器则用于实现精确的细胞穿刺操作。末端执行器由压电执行器与柔顺式位移放大机构组成,显微注射器的粗端固定于一个力传感器上,并整体嵌入柔顺机构的末端,从而实现对注射过程中力学信号的实时监测。

主控处理模块以工控机为核心,搭配运动控制卡(NI PCI-6229)完成对整个系统运动指令的生成与发送,并协调图像采集与运动执行流程,保证操作的同步性与精确性。

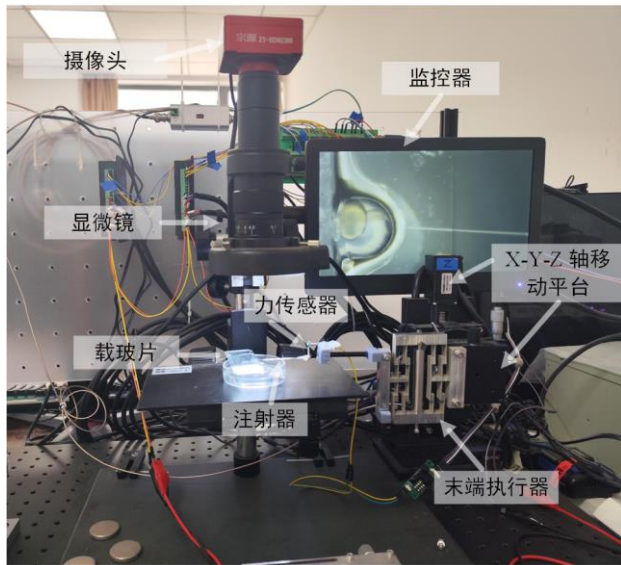


图 10 系统硬件平台

Fig.10 System hardware platform

2.2 数据集准备

斑马鱼胚胎因其胚胎透明、体积较大等特性,在生物医学研究中被广泛应用^[9]。本研究亦将其作为实验模型,以验证所提出方法的有效性。通过安装在微操作平台上的相机,采集了不同状态下的斑

马鱼胚胎细胞及注射器图像,并采用 Labelme 图形图像标注工具对图像进行标注,构建了一个专用的斑马鱼胚胎与注射器数据集。该数据集共包含 1 123 张图像,按 70% : 20% : 10% 的比例划分为训练集、验证集和测试集。划分过程采用随机划分策略,以避免样本顺序或采集条件带来的偏置。具体而言,在划分前对全部样本进行随机打乱,并设置固定的随机种子以保证实验结果的可复现性,随后依据预设比例依次划分得到训练集、验证集和测试集。验证集用于监控训练过程并评估模型性能,测试集则用于最终检验模型的泛化能力。

2.3 模型训练设置与实现细节

实验硬件环境为一台搭载 Windows 11 操作系统的计算机,配备 Intel i7-13700H CPU 和 NVIDIA GeForce RTX 4070 GPU。软件环境为 Python 3.10, CUDA 11.7。模型训练基于 PyTorch 2.0.0 深度学习框架实现,采用 SGD 优化器对网络参数进行更新,数据加载线程数为 4。初始学习率设为 0.01,动量参数为 0.937,权重衰减系数为 0.000 5,批大小设置为 16。训练总轮数设 300 轮。在训练初期引入学习率预热策略,对学习率进行线性预热,以提升模型收敛稳定性;初始预热学习率设为 0.1,预热动量为 0.8,偏置项学习率为 0.1。为减小随机因素对实验结果的影响,本文在训练过程中固定随机种子,并启用确定性训练。训练过程中启用自动混合精度以提升训练效率。

在模型训练阶段,对训练集采用 Ultralytics 框架提供的在线数据增广策略。该策略在每个训练迭代中对输入样本随机施加多种几何与光照变换,主要包括马赛克拼接增强、样本混合增强、随机擦除、随机缩放、平移、翻转以及颜色扰动等增强策略。由于数据增广过程为在线随机生成,训练集中样本数量在数据层面保持不变,但样本在不同训练轮次中呈现出多样化的增强形态,从而有效提升模型的泛化能力。

整个 DELD-YOLO 网络基于 PyTorch 框架构建,实现在 PyCharm 平台中进行训练与测试,最终完成对斑马鱼胚胎、注射器及针尖位置的深度学习检测。

2.4 指标

采用精度 (Precision)、召回率 (Recall)、平均精度 (mAP)、在交并比 (Intersection over Union, IoU) 阈值为 50% 的平均精度 (mAP@0.5)、在 IoU 阈值从 50% 到 95% 范围内的平均精度

(mAP@0.5:0.95)对检测器进行评价。此外，引入模型大小(Parameters)和每秒千兆浮点运算(GFLOPs)来评估模型的复杂性。

$$P = \frac{TP}{TP + FP} \quad (9)$$

$$R = \frac{TP}{TP + FN} \quad (10)$$

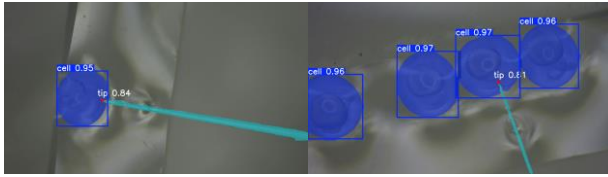
$$AP = \int_0^1 P(R) dR \quad (11)$$

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (12)$$

式中： P 代表精度, R 代表召回率, TP 代表真阳性样本的数量, FP 代表假阳性样本的数量, FN 代表假阴性样本的数量, $P(R)$ 代表曲线绘制的精度和召回。

2.5 检测效果展示

传统的细胞检测方法通常依赖边界框进行粗略定位，难以准确捕捉细胞的精细形态结构及注射器针尖的精确位置，在显微注射过程中易造成细胞损伤，从而降低存活率。相比之下，基于分割的方法能够精确提取细胞与注射器的轮廓信息，更易于识别注射器针尖的准确位置，其效果如图 11 所示。



(a) 单细胞注射 (b) 多细胞注射
(a)Single-cell injection (b)Multi-cell injection

图 11 细胞与注射针分割效果

Fig.11 Cell and needle segmentation effect

由于细胞本身较为脆弱，注射过程中需避开极

点、细胞核等关键区域，因此选择恰当的注射点至关重要。传统图像处理方法难以精准获取细胞轮廓上的理想注射点，而 DELD-YOLO 算法通过完整分割细胞轮廓，后期可快速有效识别并标注多个候选注射点，为自动化注射操作提供可靠支持。相关效果如图 12 和表 1 所示。

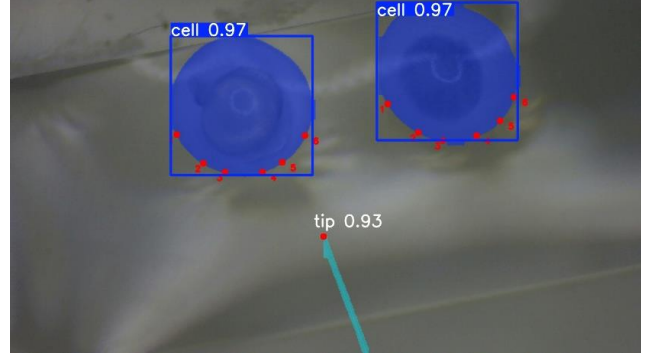


图 12 多注射点选择效果图

Fig.12 Multi-injection point selection rendering

表 1 注射点坐标

Table 1 Injection point coordinates

坐标点序号	细胞（左）	细胞（右）
1	(507, 417)	(1 149, 324)
2	(588, 504)	(1 242, 411)
3	(654, 531)	(1 317, 435)
4	(768, 531)	(1 419, 420)
5	(828, 501)	(1 491, 375)
6	(897, 420)	(1 533, 303)

2.6 消融实验

分割效果表明，DELD-YOLO 模型对细胞和注射器有很好的分割效果。但是改进模块对模型的分割性能的具体提升效果还尚待探究。因此，依次在基线模型的基础上对改进模块进行消融实验。实验结果如表 2 所示。

表 2 消融实验结果

Table 2 Ablation experimental results

Model			cell		injector		Parameters/M	FLOPs/G
DBB	ELMSFPN	LSDESH	mAP@0.5	mAP@0.5:0.95	mAP@0.5	mAP@0.5:0.95		
--	--	--	99.5%	97.3%	99.2%	68.1%	2.83	10.2
✓			99.5%	97.7%	99.2%	72.0%	2.82	10.1
	✓		99.5%	98.3%	96.6%	70.4%	2.30	10.3
		✓	99.5%	97.7%	99.2%	74.6%	2.40	9.5
✓	✓		99.5%	98.0%	99.1%	74.7%	2.30	10.3
✓	✓	✓	99.5%	98.7%	99.4%	75.2%	2.08	10.0

由表2可知，DBB、ELMSFPN和LSDESH等模块均对模型精度产生了积极影响。在显微图像中，细胞属于大尺度目标，因此改进后模型在细胞分割上的mAP提升相对有限；而注射器作为小目标，其

分割精度的显著提高更能够体现出各模块的优势。

单独引入DBB模块后，细胞与注射器的分割精度均有所提升，同时计算量与参数量还略有下降。这表明经训练后，DBB可等价转换为单一卷积层进

行推理,不会引入额外推理开销。ELMSFPN模块在仅增加0.1 GFLOPs计算量的前提下,不仅提升了分割精度,还降低了参数量,进一步验证了本研究设计的特征融合网络的高效性。LSDESH模块使注射器分割的 $mAP@0.5:0.95$ 提升了约6.5%,参数量减少43万,计算量也显著下降。该模块所采用的DEConv与共享卷积结构对参数进行了重新组织与精简,提高了参数利用率;通过重参数化技术,更加高效地利用了参数空间,从而减少了冗余参数与计算负担。

最终,本研究提出的DELD-YOLO模型相较于原始模型,在细胞分割的 $mAP@0.5:0.95$ 上提升了1.4%,注射器分割的 $mAP@0.5:0.95$ 提升了7.1%,参数量降低75万,计算量也进一步减少。以上结果表明,DELD-YOLO凭借其创新的结构设计与优化策

略,在计算效率与分割精度之间实现了良好平衡,有效缓解了计算资源与模型性能之间的冲突。

2.7 对比实验

为验证本文算法在细胞与注射器分割任务上的有效性,我们将其与若干同类先进算法进行了对比实验,结果如表3所示。实验结果表明,DELD-YOLO模型在分割精度方面表现优异,显著优于其他对比模型。具体而言,该模型在细胞与注射器分割任务上的 $mAP@0.5$ 分别达99.5%和99.4%,接近完全识别;细胞分割的 $mAP@0.5:0.95$ 也达到了98.7%。尽管注射器在该指标下的结果为75.2%,但仍高于所有对比模型。在效率方面,DELD-YOLO的参数量和计算量也优于大多数对比模型,进一步体现了其良好的综合性能与部署潜力。

表 3 DELD-YOLO 与其他实例分割算法的对比实验

Table 3 Comparative experiment of DELD-YOLO and other instance segmentation algorithms

Model	cell		injector		Parameters/M	FLOPs/G
	$mAP@0.5$	$mAP@0.5:0.95$	$mAP@0.5$	$mAP@0.5:0.95$		
Mask-RCNN	98.5%	97.2%	98.0%	63.7%	43.98	259.0
YOLOAT	99.0%	96.1%	70.3%	20.3%	34.73	61.5
ConvNeXt-V2	98.9%	89.3%	62.6%	21.3%	47.68	262.0
Solov2	98.5%	95.9%	100%	64.9%	46.59	246.0
RTMDet	98.9%	97.9%	100%	45.6%	10.18	21.6
YOLOv8n	99.5%	98.0%	99.4%	75.0%	3.26	12.0
YOLOv12n	99.5%	97.6%	99.4%	74.6%	2.81	10.2
DELD-YOLO	99.5%	98.7%	99.4%	75.2%	2.08	10.0

2.8 参数分析与稳定性验证

为评估所提出DELD-YOLO模型对关键训练参数的稳定性,本文进一步分析了不同初始学习率设置对模型分割性能的影响。实验中,在保持网络结构、数据集划分及其余训练参数一致的前提下,仅对初始学习率进行调整,分别设置为0.005、0.01和0.02,并对模型在测试集上的性能进行对比分析。

不同学习率条件下的实验结果如表4所示。可以看出,在3种学习率设置下,模型在细胞与注射器分割任务上的整体性能变化较小。对于细胞这一大尺度目标,其 $mAP@0.5$ 在不同学习率下均保持在99.5%, $mAP@0.5:0.95$ 仅呈现轻微波动,表明模型对学习率变化具有较强的稳定性。对于注射器这一小尺度目标,其分割性能在不同学习率条件下同样保持稳定, $mAP@0.5:0.95$ 的变化幅度不超过0.3%,未出现明显性能退化现象。

表 4 不同学习率下的性能对比

Table 4 Effect of different learning rates on model performance

Learning rate	cell		injector	
	$mAP@0.5$	$mAP@0.5:0.95$	$mAP@0.5$	$mAP@0.5:0.95$
0.005	99.5%	98.5%	99.4%	75.0%
0.01	99.5%	98.7%	99.4%	75.2%
0.02	99.5%	98.4%	99.3%	75.3%

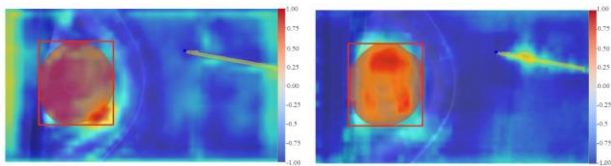
综合分析表明,在合理的学习率取值范围内,DELD-YOLO模型能够稳定收敛并保持较高的分割精度,说明所提出方法对初始学习率参数不敏感,具有良好的参数稳定性与鲁棒性。同时,实验结果也验证了本文所采用的学习率设置在性能与稳定性之间达到了较优平衡。

2.9 热力图:

类别激活图(Class activation mapping, CAM)是一种广泛使用的可视化工具,能够直观显示模型

在输入图像中的关注区域及其重要性分布,有助于理解模型的决策机制与识别潜在问题区域^[43-44]。该类图通常以颜色编码表示激活强度(如红色代表高激活,蓝色代表低激活),以突出关键特征区域。

从图 13 的类别激活图对比中可以看出,与 YOLO11 相比,DELD-YOLO 对细胞内部结构表现出更显著的关注,其热力响应峰值更紧密地集中于细胞中心区域,表明模型能够提取更具判别性的特征。同时,注射器区域呈现出更大范围的黄色至红色激活,说明 DELD-YOLO 对微小目标的感知与聚焦能力也更强,这与其在分割精度上的提升是一致的。



(a) YOLO11 算法 (b) DELD-YOLO 算法
(a) YOLO11 algorithm (b) DELD-YOLO algorithm

图 13 YOLO11 与 DELD-YOLO 算法热力图对比

Fig.13 Heat map comparison between YOLO11 and DELD-YOLO algorithms

3 结束语

本文面向自动细胞注射系统中的视觉检测任务,提出了一种基于 DELD-YOLO 的实例分割方法,以取代传统的目标检测方法。该方法在 YOLO11 的主干网络中引入多样性分支块,增强模型的多尺度特征捕获能力;设计了一种高效的轻量化多尺度特征金字塔结构,显著提升了不同层级特征的融合效率;并构建了基于细节增强卷积的分割头,强化对细胞及注射器细微结构的感知与分割精度。最终,DELD-YOLO 模型在参数量与计算量较低的情况下,实现了对细胞和注射器的高效率分割,并能够准确识别注射点及注射器针尖位置。

该系统与注射控制系统结合,可有效减轻实验人员繁重重复的操作负担,并为缺乏经验的用户提供可靠、高效的斑马鱼胚胎注射辅助。在未来工作中,本方法可与运动控制系统进一步集成,适用于大批量细胞注射任务,具有重要的应用前景。

参考文献

- [1] 沈耀辉.基于视觉伺服的辅助细胞操作系统研究[D].南京:南京航空航天大学,2021.
- [2] Li L, Wang C, Gong H, et al. Automatic high-throughput injection system for zebrafish larvae based on precise positioning of injection target[J]. IEEE Transactions on Automation Science and Engineering, 2023, 21(4): 6423-6433.
- [3] Zhan S, Bi H, Wei J, et al. Development of a micromanipulator robot for cell injection with sub-picoliter scale[C]//2023 IEEE International Conference on Manipulation, Manufacturing and Measurement on the Nanoscale (3M-NANO). [S.l.]: IEEE, 2023: 233-237.
- [4] Zhang G, Tong M, Qian C, et al. Visual-based contact detection for automated zebrafish larva heart microinjection[J]. IEEE Transactions on Automation Science and Engineering, 2020, 18(4): 1803-1813.
- [5] Chi Z, Xu Q, Ai N, et al. Design and development of an automatic microinjection system for high-throughput injection of zebrafish larvae[J]. IEEE Transactions on Automation Science and Engineering, 2021, 19(4): 3409-3418.
- [6] Guo Z, Ai N, Ge W, et al. Design of an automated robotic microinjection system for batch injection of zebrafish embryos and larvae[J]. Microsystems & Nanoengineering, 2024, 10(1): 20.
- [7] Ekstrom M P. Digital image processing techniques[M]. [S.l.]: Academic Press, 2012.
- [8] Permana S, Grant E, Walker G M, et al. A review of automated microinjection systems for single cells in the embryogenesis stage[J]. IEEE/ASME Transactions on Mechatronics, 2016, 21(5): 2391-2404.
- [9] Zhao Y, Sun H, Sha X, et al. A review of automated microinjection of zebrafish embryos[J]. Micromachines, 2018, 10(1): 7.
- [10] Mori R, Aoyama T, Kobayashi T, et al. Real-time spatiotemporal assistance for micromanipulation using imitation learning[J]. IEEE Robotics and Automation Letters, 2024, 9(4): 3506-3513.
- [11] Qian C, Tong M, Yu X, et al. CNN-based visual processing approach for biological sample microinjection systems[J]. Neurocomputing, 2021, 459: 70-80.
- [12] Pan F, Jiao Y, Chen S, et al. Deep learning-enhanced dual-module large-throughput microinjection system for adherent cells[J]. IEEE Transactions on Automation Science and Engineering, 2022, 20(4): 2409-2422.
- [13] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 779-788.
- [14] Wang H, Liu C, Cai Y, et al. YOLOv8-QSD: An improved small object detection algorithm for autonomous vehicles based on YOLOv8[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1-16.
- [15] Yi H, Liu B, Zhao B, et al. Small object detection algorithm based on improved YOLOv8 for remote sensing[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 17: 1734-1747.
- [16] 涂育智,王法翔,吴春霖.融合多注意力机制的轻量级无人机航拍小目标检测模型[J].计算机工程与应用, 2025, 61(11): 93-104.
- [17] 许景科, 索祥龙, 周磊.改进 YOLOv8 的无人机航拍小目标检测算法[J].计算机工程与应用, 2025, 61(11):119-131.
- [18] 李彬,李生林.改进 YOLOv11n 的无人机小目标检测算法[J].计算机工程与应用, 2025, 61(7): 96-104.
- [19] 卢彬,王俊.基于 YOLOv11 改进的无人机图像小目标检测模型[J].信息与电脑(理论版), 2024, 36(22): 150-152.
- [20] 周璇,葛琦,邵文泽.高分辨率特征增强的无人机航拍小目标检测[J].数据采集与处理, 2024, 39(4): 908-921.
- [21] Wang X, Gao H, Jia Z, et al. BL-YOLOv8: An improved road defect detection model based on YOLOv8[J]. Sensors, 2023, 23(20): 8361.
- [22] 高家源,张新.DMH-YOLO: 针对屏幕外观缺陷检测与分割的高效算法[J/OL].计算机工程与应用, 1-13[2025-06-09].

- http://kns.cnki.net/kcms/detail/11.2127.TP.20250121.1616.008.html.
- [23] 朱佳慧, 刘艺, 张登银. 基于改进 YOLOv8n 的道路裂缝检测轻量化模型[J]. 数据采集与处理, 2025, 40(5): 1333-1347.
- [24] 王福顺, 王旺, 孙小华, 等. 基于改进 YOLO v8n-seg 的羊只实例分割方法[J]. 农业机械学报, 2024, 55(8): 322-332.
- [25] 王兴家, 郑纪业, 盛清凯, 等. 基于改进 YOLO v8n-seg 的群猪分割方法研究[J/OL]. 中国农业科技导报, 1-11[2025-06-09]. <https://doi.org/10.13304/j.nykjdb.2024.0393>.
- [26] 何阳. 面向森林火情监测的图像检测分割方法研究[D]. 西安: 西安理工大学, 2024.
- [27] Xu Y, Kang S, Li T, et al. Automatic microinjection assistance system based on deep-learning visual positioning[C]//2023 China Automation Congress (CAC). [S.l.]: IEEE, 2023: 1641-1645.
- [28] Ding X, Zhang X, Han J, et al. Diverse branch block: Building a convolution as an inception-like unit[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 10886-10895.
- [29] Sadak F, Saadat M, Hajiyavand A M. A vision-guided methodology for the automation of biological cell injection[C]//2020 2nd International Conference on Electrical, Control and Instrumentation Engineering (ICECIE). [S.l.]: IEEE, 2020: 1-9.
- [30] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 2117-2125.
- [31] Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 10781-10790.
- [32] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2015: 3431-3440.
- [33] Rahman M M, Munir M, Marculescu R. Emcad: Efficient multi-scale convolutional attention decoding for medical image segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2024: 11769-11779.
- [34] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 4510-4520.
- [35] Zhang X, Zhou X, Lin M, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 6848-6856.
- [36] Yang Z, Guan Q, Zhao K, et al. Multi-branch auxiliary fusion YOLO with re-parameterization heterogeneous convolutional for accurate object detection[C]//Chinese Conference on Pattern Recognition and Computer Vision (PRCV). Singapore: Springer Nature Singapore, 2024: 492-505.
- [37] Zhao H, Qi X, Shen X, et al. ICNet for real-time semantic segmentation on high-resolution images[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 405-420.
- [38] Xiao J, Zhao T, Yao Y, et al. Context augmentation and feature refinement network for tiny object detection[J]. 2022.
- [39] Bolya D, Zhou C, Xiao F, et al. Yolact: Real-time instance segmentation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2019: 9157-9166.
- [40] Chen Z, He Z, Lu Z M. DEA-Net: Single image dehazing based on detail-enhanced convolution and content-guided attention[J]. IEEE Transactions on Image Processing, 2024, 33: 1002-1015.
- [41] Tian Z, Shen C, Chen H, et al. FCOS: A simple and strong anchor-free object detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(4): 1922-1933.
- [42] Chen Y, Yuan X, Wang J, et al. YOLO-MS: Rethinking multi-scale representation learning for real-time object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(6): 4240-4252.
- [43] Zhou B, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2016: 2921-2929.
- [44] Selvaraju R R, Cogswell M, Das A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2017: 618-626.
- [45] Chi Z, Xu Q, Ai N, et al. Design and implementation of an automatic batch microinjection system for zebrafish larvae[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 1848-1855.

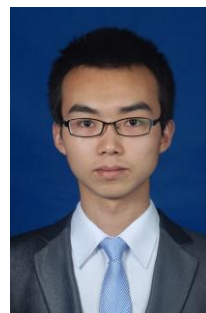
作者简介:



刘国志 (1999-), 男, 硕士研究生, 研究方向: 微纳机器人、深度学习, E-mail: 202312220024@nuist.edu.cn。



李涛 (1979-), 通信作者, 男, 教授, 研究方向: 抗干扰控制、故障诊断, E-mail: litaojia@163.com。



康升征 (1992-), 男, 副教授, 研究方向: 微纳操作机器人、外骨骼机器人。



周杰 (2001-), 男, 硕士研究生, 研究方向: 擦窗机器人。