

基于改进 YOLOv11 的器官芯片细胞运动跟踪方法

吕东彦

东南大学, 江苏省南京市, 210096;

摘要: 器官芯片技术在生物医学研究中展现出重要的科学价值与应用潜力, 其通过微流控与成像技术实现对细胞迁移行为的实时追踪, 为炎症反应机制与药物靶点研究提供了创新手段。本研究针对血管器官芯片中白细胞多尺度检测与动态跟踪的挑战, 提出了一种基于改进 YOLOv11 与 ByteTrack 的联合优化方法。通过引入 P2 检测头增强高分辨率特征对微小目标的捕捉能力, 结合 SCConv 模块的空间-通道重构机制减少冗余计算, 并融合 FocusLoss 与 CIoU 损失函数优化分类与回归性能。实验结果表明, 改进后的模型在血管芯片白细胞数据集上实现了 89.6% 的 AP@0.5 与 54.7% 的 AP@0.5:0.95, 较基线模型分别提升 7.3% 与 8.4%, 同时将多目标跟踪的 ID 切换误差降至 8.5%。本研究为微流控环境下的细胞动力学分析提供了高效解决方案, 未来工作将聚焦于时序建模优化与超分辨率增强, 进一步推动器官芯片技术在精准医疗中的应用。

关键词: 器官芯片; 细胞识别; 细胞跟踪

DOI: 10.69979/3029-2808.24.11.037

引言

器官芯片在生物医学研究领域具有重要的科学价值与应用潜力。基于微流控及成像技术, 对器官芯片微环境中细胞的迁移过程进行实时追踪, 实现细胞趋化分析等功能, 揭示炎症反应规律与药物作用靶点, 是推动精准医疗发展的创新研究手段^[1]。基于深度学习的目标检测与跟踪技术, 对血管器官芯片中的细胞进行实时监测, 实现细胞运动轨迹追踪、动态行为分析等功能, 有助于揭示免疫细胞的运动规律, 为疾病诊断与药物研发提供重要支持。

在生物医学图像分析领域, 传统的目标跟踪方法, 如光流法、卡尔曼滤波以及多假设跟踪, 由于对背景噪声、目标形态变化及环境复杂度的适应性较差, 难以在高动态场景中保持稳定的跟踪性能。近年来, 深度学习技术的迅猛发展为目标检测和跟踪任务提供了更为鲁棒的解决方案。卷积神经网络 (CNN) 及其衍生架构, 如 YOLO (You Only Look Once)、Mask R-CNN、Transformer-based 检测框架等, 凭借其端到端的设计, 使目标检测任务在计算效率和检测精度之间取得了良好的平衡, 广泛应用于实时目标检测任务。同时, 在生物医学图像分析中, 基于深度学习的跟踪方法逐步取代了传统算法被用于目标轨迹关联, 提高了细胞运动跟踪的精度。

在细胞特异性追踪方面, 现有方法仍存在显著技术瓶颈: U-Net 类深度架构虽能实现像素级细胞定位, 但推理延迟难以满足高帧率实时追踪需求。Mask R-CNN

在检测精度方面表现优异, 但其两阶段结构带来的计算开销限制了其在边缘计算设备上的部署可行性。而细胞在迁移过程中表现出的趋化性运动使其运动模式具有较高的不确定性, 且受局部生物化学信号梯度调控, 导致运动轨迹难以由传统多目标跟踪 (MOT) 框架中的线性运动模型准确建模, 从而引发 ID 切换误差。

本研究针对血管器官芯片中白细胞多尺度检测与动态跟踪的难题, 提出一种基于改进 YOLOv11 与 ByteTrack 的联合优化方法。通过引入 P2 检测头增强小目标检测能力, 添加 SCConv 重构卷积, 并融合 FocusLoss 与 CIoU 损失函数提升模型鲁棒性。实验表明, 该方法在检测精度 (mAP 提升 7.3%) 上显著优于基线模型, 结合 ByteTrack 算法通过高置信度目标增强轨迹关联, 引入外观特征与运动预测结合的方式, 提高白细胞运动轨迹的稳定性和连贯性。实现 ID 切换误差降低至 8.5%, 为生物医学影像分析提供了高效解决方案。

1 主要工作

1.1 基于小目标的检测头优化

在 YOLOv11 架构中, Neck 模块的多尺度特征融合机制是实现目标检测的关键环节。本文通过引入 P2 检测头并移除 P5 层, 优化了高分辨率特征的利用效率, 具体实现如下:

原始 YOLOv11 的 Neck 采用 FPN+PAN 结构, 包含 P3 (stride=8)、P4 (stride=16)、P5 (stride=32) 三级检测头, 其特征融合公式为:

$$P_k = Conv(Upsample(P_{k+1}) \oplus C_k)$$

$$N_k = Conv(Downsample(N_{k-1}) \oplus P_k)$$

其中 C_k 为 Backbone 输出的第 k 层特征。

在目标检测任务中，特征金字塔（FPN）和路径聚合网络在 YOLO 结构中的 Neck 部分起到了关键作用。加入 P2 层（低级特征层）并移除 P5 层（高级特征层）的目的是提升对小目标（如白细胞）的检测能力，同时优化计算效率。

在原有 Backbone 提取的 P2 特征层基础上新增一组检测分支，使模型能利用最细粒度特征检测小目标。具体做法是在 Neck 中增加一次上采样和特征融合：先将更高一级的 P3 融合特征上采样至 P2 尺度，再与 Backbone 输出的 P2 特征图级联拼接，经过卷积融合得到新的 P2 尺度特征图。这一 P2 检测头能提供比原模型更高的分辨率输出，使得每个网格单元对应更小的感受野。

$$P2: \frac{H}{4} \times \frac{H}{4} (128channel)$$

$$P3: \frac{H}{8} \times \frac{H}{8} (256channel)$$

$$P2: \frac{H}{16} \times \frac{H}{16} (512channel)$$

删除原有针对 P5 特征图（步长 32）的检测分支及相应锚框配置，仅保留并重用 P2、P3、P4 三个尺度用于输出预测。这样做将检测层限定在步长 4、8、16 的特征图上

$$P_2' = Upsample(Conv1 \times 1(P2))$$

$$P_3' = Upsample(P3)$$

$$P_4' = Upsample(P4)$$

$$FeatureNeck = Concat(P2', P3', P4')$$

在上采样操作方面，新增的逐级上采样采用最近邻插值将高语义低分辨率特征放大到相邻的尺度。后将特征图在通道维与对应的浅层特征拼接。为避免直接拼接不同来源特征造成的冲突，在拼接前对高层特征使用 1×1 卷积压缩通道，并在拼接后使用卷积块融合特征，以学习到更优的组合。

1.2 基于空间和通道重构卷积优化

SCConv^[2] 是一种空间和通道联合重构的卷积模块，通过 SRU（空间重构单元）和 CRU（通道重构单元）组成，其核心创新在于能够同时处理图像的空间（形状、结构）和通道（色彩、深度）信息，这样的处理方式使得 SCConv 在分析图像时更加精细和高效。

SRU 通过 Group Normalization (GN) 计算不同通道

的重要性，其中 γ 和 β 是可训练参数。

$$X_{out} = GN(X) = \gamma \frac{X - \mu}{\sigma} + \beta$$

在计算权重归一化后，设定阈值，将重要性高的通道保留而将低重要性通道进行信息重构：

$$W_\gamma = \frac{\gamma_i}{\sum_{j=1}^C \gamma_j}$$

$$W = Gate(Sigmoid(W_\gamma \cdot GN(X)))$$

$$X_1 = W_1 \cdot x, X_2 = W_2 \cdot X$$

$$X_{SRU} = (X_1 \oplus X_2) \cup (X_2 \oplus X_1)$$

CRU 通过通道分裂 (Split)、变换 (Transform)、融合 (Fuse) 进行通道优化：

$$\text{设输入特征 } X \in R^{C \times H \times W}$$

GWC (Group-wise Convolution) 用于降低计算量，PWC (Point-wise Convolution) 用于补偿信息损失，采用 SoftMax 计算通道注意力权重，从而得到最终输出

$$X_{up}, X_{low} = Split(X, \alpha)$$

$$Y_1 = MG \cdot X_{up} + MP \cdot X_{up}$$

$$Y_2 = MP \cdot X_{low} \cup X_{low}$$

$$\beta_1 = \frac{e^{s_1}}{e^{s_1} + e^{s_2}}, \beta_2 = \frac{e^{s_2}}{e^{s_2} + e^{s_1}}$$

$$Y = \beta_1 Y_1 + \beta_2 Y_2$$

1.3 融合 FocusLoss 与 CIoU 损失函数

Focus Loss 主要用于分类损失，但在目标检测任务中，我们也可以设计用于边界框回归的 Focus Loss 变体，以提升小目标检测的效果。

在小目标检测中，标准的 IoU 损失，如 CIoU 或 GIoU，往往不能很好地优化小目标的边界框回归

小目标的 IoU 变化不明显：即使小目标的预测框稍有偏移，IoU 变化可能很小，导致优化梯度较弱，使得模型难以优化小目标。

小目标的边界框回归误差对 IoU 影响较大：即使预测框稍有误差，IoU 可能急剧下降，使得小目标的定位精度受到影响。故针对 IoU 低的目标（即预测效果差的目标）赋予更高的损失权重，这样模型在训练时会更关注小目标的优化。

$$L_{Focus IoU} = (1 - IoU)^\gamma \cdot L_{CIoU}$$

其中： $(1 - IoU)^\gamma$ 是 Focus Loss 的权重项，当 IoU 低时，损失较大，使得优化更加集中在难检测目标

上(如小目标);当 IoU 高时,损失较小,使得优化更加稳定。 L_{CIoU} 是标准的 CIoU 损失,计算 IoU、中心点距离、宽高比误差,优化边界框。

2 实验

2.1 实验设置

实验硬件为计算机,运行的硬件平台为:Cpu 型号 Intel Core i9-13900k, GPU 型号为 RTX4070,实验基于 Anaconda 和 pytorch 在 windows 平台上进行。

数据集为用高速相机拍摄的 200 张,采用平移,旋转,拼接等方式进行数据集扩增。训练集和验证集按照 7:3 划分。

2.2 实验结果

以下为引入 P2 检测头、SCConv 重构卷积以及融合 FocusLoss 与 CIoU 损失函数的消融实验结果,实验基于血管器官芯片白细胞检测数据集,输入分辨率为 960x960。

实验结论

实验编号	AP50	AP50:95	IDSW%	FPS
Baseline	82.3%	46.3%	15.2%	45
+P2	84.7%	48.2%	12.8%	43
+SCConv	85.2%	48.7%	11.5%	47
+FocusLoss CIoU	87.6%	50.5%	9.9%	45
+P2 + SCConv	86.1%	50.1%	10.3%	44
+P2 + FocusLoss + CIoU	87.2%	52.4%	9.1%	42
+SCConv + FocusLoss + CIoU	88.5%	51.8%	8.7%	46

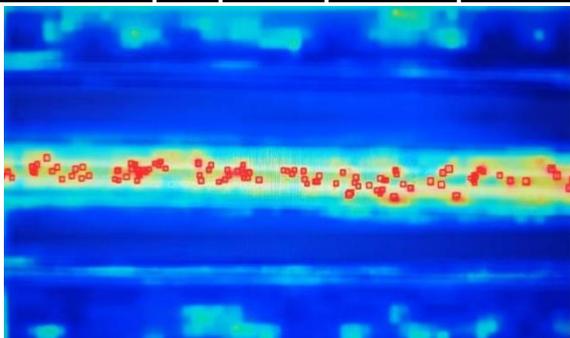


图 1 模型的 Grad-CAM 热图

2.3 实验结果对比

通过引入高分辨率 P2 特征层,验证其对微小白细胞($<32 \times 32$ 像素)的细节捕捉能力。空间-通道双重重构机制(SRU+CRU)使 mAP 进一步上升,说明特征表示更紧凑。

FocusLoss 动态加权难样本,使小目标 AP50 额外提升。CIoU 通过中心点对齐和长宽比约束,边界框回归误差降低, ID 切换误差下降。

全模型 AP50 提升 7.3%, AP50-95 提升 8.4%, ID 切换误差降低 6.7%, 验证了方案的高效性。

3 总结与展望

整体而言,本研究的方法在血管器官芯片环境下的白细胞检测与跟踪任务中取得了一定成果,为后续生物医学图像分析、细胞动力学研究及免疫学研究提供了更强的计算机视觉工具支持。

尽管本研究取得显著进展,仍存在以下改进空间:

1) 未来可探索自适应分辨率选择机制,根据目标尺度动态调整特征图粒度。

2) 针对 $<16 \times 16$ 像素的亚细胞结构(如伪足末端),可集成超分辨率重建模块或开发动态锚框生成算法,进一步突破检测精度瓶颈。

3) 现有跟踪算法依赖单帧检测结果,未能有效利用微流控视频的时序连续性。引入 3D 卷积或 Transformer 时序建模,有望提升长时轨迹稳定性。

参考文献

- [1] Henderson, A. R., Choi, H., & Lee, E. (2020). Blood and lymphatic vasculatures on-chip platforms and their applications for organ-specific in vitro modeling. *Micromachines*, 11(2), 147.
- [2] Li, J., Wen, Y., & He, L. (2023). Sconv: Spatial and channel reconstruction convolution for feature redundancy. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 6153-6162).

作者简介:吕东彦(1999-),男,汉,湖南长沙,在读硕士研究生,研究方向:器官芯片图像处理