

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2024)09-2806-19

论文引用格式: Jiao R H, Chu J J, Liu J J and Yu J M. 2024. SC-Net: a contextual information skip connection network for overlapping chromosome segmentation. Journal of Image and Graphics, 29(09):2806-2824(焦润海, 褚佳杰, 刘嘉骥, 余济民. 2024. SC-Net:用于重叠染色体分割的上下文信息跳跃连接网络. 中国图象图形学报, 29(09):2806-2824)[DOI:10.11834/jig.230599]

SC-Net: 用于重叠染色体分割的 上下文信息跳跃连接网络

焦润海*, 褚佳杰, 刘嘉骥, 余济民

华北电力大学控制与计算机工程学院, 北京 102206

摘要: 目的 染色体核型分析从细胞分裂中期图像中分离和分类染色体, 是遗传疾病诊断广泛采用的方法, 其中形态多样的重叠染色体簇的分割, 依赖于准确的边界等细节特征。为此, 本文融合目标的上下文信息, 构建了一种两阶段的重叠染色体分割模型 SC-Net(skip connection network)。方法 首先, 在语义分割基线模型 U-Net++ 中增加混合池化模块捕获重叠染色体的局部上下文信息, 在解码器网络中并联上下文融合模块和上下文先验辅助分支, 增强通道和空间上的全局上下文信息。其次, 利用已标注样本的类别先验信息生成真实亲和矩阵, 加入训练过程以有效区分重叠染色体图像中易混淆的空间信息。最后, 通过染色体实例重建算法对重叠与非重叠区域的元素迭代进行配对, 拼接形成单条染色体。结果 在公开的 ChromSeg(chromosome segmentation)数据集上进行实验, 结果表明 SC-Net 分割出的重叠染色体区域交并比值为 83.5%, 与对比算法中的较优算法相比性能提升 2.7%。结论 本文构建的重叠染色体分割模型通过融合上下文信息, 能更有效地解决形态多样的重叠染色体簇的分割问题, 相比对比方法可以得到更精细和准确的结果。

关键词: 重叠染色体分割; 混合池化模块 (HPM); 上下文融合模块 (CFM); 上下文先验辅助分支 (CPAB); 真实亲和矩阵

SC-Net: a contextual information skip connection network for overlapping chromosome segmentation

Jiao Runhai*, Chu Jiajie, Liu Jiaji, Yu Jimin

School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China

Abstract: Objective Chromosome karyotype analysis separates and categorizes chromosomes in midcell division images, and it is widely used for the diagnosis of genetic diseases, in which overlapping chromosome segmentation is one of the key steps. Based on image analysis of overlapping chromosomes, the morphologically diverse chromosome clusters depend on detailed features, such as accurate boundaries during segmentation, in addition to obtaining the basic contour, texture, and semantic information. For this reason, in this paper, a two-stage overlapping chromosome segmentation model SC-Net was constructed through fusion of the contextual information of the target to improve the segmentation performance of the network. **Method** First, the model SC-UNet++ added the hybrid pooling module (HPM) to the baseline model U-Net++ for

收稿日期: 2023-08-25; 修回日期: 2023-12-26; 预印本日期: 2024-01-02

* 通信作者: 焦润海 runhaijiao@ncepu.edu.cn

基金项目: 中央高校基本科研业务费专项资金资助(2022JG004)

Supported by: Fundamental Research Funds for the Central Universities (2022JG004)

semantic segmentation to capture the local context information of overlapping chromosomes and complemented the detailed features of chromosomes, such as color, thickness, and stripes, based on the superposition operation of empty space pyramid pooling and stripe pooling. The context fusion module (CFM) was connected in parallel to a decoder network, i. e., the channel correlation of input features was extracted using the efficient channel attention module, and the features obtained via the multiplication of the output with the input were subsequently fed to the HPM and the spatial attention module (SAM), which explored the correlation of the region around the pixel to obtain the local context and extract the global context through global pooling operation, respectively. In addition, context prior auxiliary branch (CPAB) was introduced after CFM to improve the global context information on channel and space. Second, the category a priori information of labeled training samples, which serves as an additional source of supervisory information during training and effectively distinguishes confusing spatial features in overlapping chromosome images, was used to generate the true affinity matrix. Finally, the elements of overlapping and non-overlapping regions were iteratively paired by the chromosome instance reconstruction algorithm to splice and form a single chromosome. In this paper, experimental analysis were based on ChromSeg dataset, and the hardware resources used included a desktop server with 32 GB RAM, a 3.3 GHz Intel Xeon CPU, and an NVIDIA RTX 3070 GPU. The model was used based on the semantic segmentation toolkit MMSegmentation version 0.30.0 and implemented under the Ubuntu 18.04 operating system, with PyTorch 1.10.0 serving as a deep learning framework. The following sections describe relevant hyperparameter settings and initialization methods for network training and the loss function selection strategy. **Result** SC-Net fully extracted and utilized the contextual and category prior information of overlapping chromosome images and showed good performance in segmentation scenarios with various numbers of overlapping chromosomes. The effect of each improvement on the algorithm performance was investigated through ablation experiments, where various combinations of CFM, HPM, CPAB, and segmentation loss were designed on the baseline model U-Net++. The results proved the better performance of SC-UNet++ compared with models in all evaluation metrics. This condition confirms the effectiveness of the method proposed in this paper, i. e., SC-UNet++ attained better performance in the segmentation of overlapping chromosomes. Through comparative experiments, the SC-Net proposed in this paper caused improvements on the ChromSeg dataset, which outperformed several models in terms of all metrics, and the model achieved an overlapping chromosome region intersection and merger ratio of 83.5%. The overall accuracy obtained after the reconstruction of chromosome instances was 92.3%, which is higher than the best and the same two-stage ChromSeg segmentation methods by 2.7% and 1.8%. The SC-Net outperformed these models mainly due to its capability to extract contextual information and category relevance of the target, and it enables the model to further gain insights into overlapping regions. **Conclusion** The overlapping chromosome segmentation model constructed in this paper can effectively solve the segmentation problem of morphologically diverse overlapping chromosome clusters by fusing contextual information and obtaining finer and more accurate results compared with the existing methods.

Key words: overlapping chromosome segmentation; hybrid pooling module (HPM); contextual fusion module (CFM); contextual prior aided branch (CPAB); real affinity matrix

0 引言

人工进行染色体核型分析(罗纯龙和赵屹, 2023)包含4个主要阶段:选择中期图像、提取单条染色体和重叠接触染色体、分割重叠接触的染色体以及染色体分类(王广杰, 2022),最终根据分割和分类结果生成医学报告。其中,重叠染色体由于重叠区域形态多样、像素叠加直接影响分割和分类的准确度,进而影响临床诊断结论。因此,近年来,许多研究人员提出了一些针对重叠染色体的计算机辅助

或全自动方法以提高染色体分割和核型分析的准确度和效率。

1)传统重叠染色体分割方法。重叠染色体分割的传统方法主要基于阈值、轮廓以及几何等方法或其中几种的组合实现。其中,由于重叠染色体在直方图中共享一个区域,仅简单使用全局或局部阈值的方法适用性受限,需要人工交互式的干预对重叠染色体进行分离。敬雪等人(2008)使用阈值处理辅助几何形态学进行重叠染色体分割,但实验表明此方法的精度不高。Grisan等人(2009)提出一种局部自适应阈值法,基于图像和几何信息多次递归得到

多个分割结果,但此方法需要大量的人工筛查,因而仅能用于辅助或半自动的核型分析系统。与阈值方法不同,基于轮廓的方法不需要对图像进行预处理。范继红等人(2009)将区域生长法和边缘检测算法相结合以进行重叠染色体分割,但此方法对边缘检测结果的准确度依赖极高,对染色体形态的变化也比较敏感。基于几何的方法能够处理更加复杂的重叠染色体,Minaee等人(2014)提出了一个两阶段方法,第1阶段采用多个几何标准检测染色体簇并在边界上找出交叉点;第2阶段基于交叉点分离染色体簇。此方法通过多次执行第2阶段来处理多条染色体交叉形成的重叠染色体簇,但无法区分染色体接触和重叠的不同现象。Madian等人(2015)基于染色体非重叠区域和重叠区域的轮廓,使用曲率函数识别轮廓中的凹凸点以搜索重叠区域,该方法可以适应多条重叠染色体的分割。此外,不同于关注重叠染色体是否被正确分离,Yilmaz等人(2018)旨在提高重叠染色体的分割质量。首先对输入图像执行灰度阈值处理和形态学操作去除复杂边缘和背景噪声;然后通过灰度重建法(Vincent, 1993)和分水岭方法(Grau等, 2004)分离对象区域;最后通过结合轮廓切点和染色体测地线分割重叠染色体。然而此方法由于基于曲率确定染色体重叠部位,因此存在凹点的漏检和误检问题。

总的来说,这些传统方法依赖于人工构造的特征提取器和先验处理,然而染色体的非刚性特点导致难以通过固定的特征提取器获取高质量、普适的特征,样本特征如灰度值等稍微变化就使得分割性能急剧下降,鲁棒性和泛化能力上存在较大的不足。

2)基于深度学习的重叠染色体分割方法。随着深度学习和计算机视觉的迅速发展,许多研究人员致力于研究深度卷积神经网络搭建的语义分割或实例分割模型解决重叠染色体分割不精准的问题。

由于深度学习模型需要大量数据,为了解决染色体图像样本量不足的问题,许多研究工作通过数据增强使得模型充分训练。Saleh等人(2019)提出在推理阶段使用测试时增强(test time augmentation, TTA)(Wang等, 2019)的方法以解决U-Net易过拟合的缺陷,并通过简化结构、降低通道数使得分割结果更加平滑。Altinsoy等人(2019)基于U-Net对G带中期细胞染色体图像进行语义分割分离染色体与背景,在训练和测试时采用旋转 $\pm 45^\circ$ 、水平或垂

直翻转、缩放10%比例大小等7种数据增强手段对样本扩充,此方法大幅优于所有基于阈值的方法。与通用的数据增强方法不同,Chen等人(2020)提出了一种数据模拟生成策略来扩充训练集,还采用Zernike矩阵(Gorelick等, 2006)表示染色体的形状以度量生成染色体的质量。此外,通过增加标注信息,Klivanec等人(2022)利用染色体的方向信息分离重叠染色体,预测染色体的方向并将其作为区分不同实例的依据。

在深度学习模型应用方面,Sharma等人(2017)首次使用深度学习模型进行染色体核型分析,但其无法实现端到端的自动染色体分割。鉴于染色体和细胞成像过程的相似性,Hu等人(2017)首次采用U-Net分割两条重叠染色体,相比之前方法的效率性能有所提升,但模型不能分割3个及以上的染色体形成的重叠染色体簇。张林等人(2022)提出一种引入跳跃连接和空洞卷积的残差U-Net网络的染色体图像分割方法。受到传统方法的启发,Cao等人(2020)采用了两阶段方法进行重叠染色体分割,在第1阶段中使用改进的U-Net++对背景类、染色体类和重叠区域类别进行语义分割,在第2阶段中使用图像处理算法拼接每条染色体实例。为了增加网络卷积层的感受野,Sun等人(2021)在U-Net的编码器的每个层级应用改进的空洞卷积以提升网络的多尺度特征提取能力,在分割多条染色体的重叠染色体簇时仍具有较为稳定的性能。

针对多条重叠染色体的分割,研究人员逐渐开始研究基于实例分割网络的方法。中国科学院广州电子技术有限公司(2022)改进了实例分割网络Mask R-CNN(mask region-based convolutional neural network),引入角度加权交并比(intersection over union, IoU)对非极大值抑制(Hosang等, 2017)进行改进,以解决漏检和边界框重叠问题。林成创等人(2020)基于路径增强网络(path aggregation network, PANet)(Liu等, 2018),添加细胞分裂中期图像中染色体数量的分支,并在特征金字塔网络(feature pyramid network, FPN)(Lin等, 2017)中插入多条路径以缩短特征传输距离。Wang等人(2021)提出了一种增强旋转模型来提高接触和重叠染色体的分割和分类准确度,但该方法依赖加入角度标注信息,实用性略显不足。Lin等人(2021)提出了一种两阶段模型,首先使用迁移学习技术分离染色体簇,然后使用经

预训练的ResNeXT(residual next)网络(Xie等,2017)分割重叠染色体。

综上所述,基于深度学习的分割方法具有更强的泛化性能,更容易实现多个染色体交叉形成的重叠染色体簇的分割。但若采用实例分割网络解决重叠染色体分割问题,因发生重叠的染色体并无语义差别,进而导致前期数据标注工作复杂,且模型训练更具挑战性。

对此,本文构建了一个两阶段的重叠染色体语义分割模型SC-Net(skip connection network)。该模型首先在U-Net++分割网络中增加混合池化模块(hybrid pooling module, HPM)捕获重叠染色体的局部上下文信息,在解码器中并联上下文融合模块(context fusion module, CFM)和上下文先验辅助分支(context prior auxiliary branch, CPAB),融合通道和空间的上下文信息。其次,利用已标注样本的类别先验信息生成真实亲和矩阵,在亲和力损失的作用下监督重叠染色体的空间关系。最后,提出染色体实例重建算法实现单条染色体的拼接。SC-Net充分提取和利用了重叠染色体图像的上下文信息和类别先验信息,显著提升了形态多样的重叠染色体簇的分割性能。

1 本文方法

1.1 问题描述和定义

在核型分析中,将染色体分离是首要任务。因为染色体的非刚性特点,在中期图像中很容易出现弯曲交叉重叠和聚集等情况,这是造成人工核型分析研究难度大和工作效率低的重要原因。所以,准确分割重叠染色体具有很大的临床应用价值,这确保了高质量核型分析的智能化和自动化,为后续染色体分类、异常检测和临床诊断的准确性和可靠性打下基础。

1.1.1 重叠染色体图像分析

由于染色体的长短和粗细不一,不同尺度的染色体导致重叠区域大小也不同,其中小尺度染色体的重叠区域往往不到大尺度的1/10。因此,基于深度学习的分割网络需要具备多尺度特征提取的能力,以获得不同尺度的细节和语义特征。而且,染色体往往聚集在中期图像中的小区域并多条交叉形成染色体簇,当面对不同深浅的条带颜色,需更准确地

区分染色体的重叠与非重叠区域。对于专家而言,他们通常会结合染色体在重叠区域附近的轮廓信息,绘制出被重叠区域截断的曲线,4条曲线所包围区域即为染色体的重叠区域。在利用语义分割模型时,则需通过利用目标图像的全局和局部上下文信息更好地理解染色体的多样形态和重叠关系。

1.1.2 上下文融合

全卷积神经网络(fully convolutional neural network, FCN)(Long等,2015)由于去除全连接层使得输入图像不需要为固定尺寸,使其能够灵活应用于不同的语义分割任务中。基于重叠染色体的图像分析,除需获取基本的轮廓、纹理以及高级语义信息外,还要考虑染色体的多样形态,其分割重点依赖于边界等细节特征。因此,语义分割网络需要融合丰富的上下文信息以实现重叠染色体的准确分割。最近的研究工作也致力于改进上下文融合策略,其中,PSPNet(pyramid scene parsing network)更改FCN的骨干网络为具有更强大特征提取能力的ResNet-101(residual network)(Zhao等,2017),并提出了金字塔池化模块(pyramid pooling module, PPM)融合多尺度上下文。为了在不增加模型参数数量的前提下扩大感受野,DeepLabV2(Chen等,2018)提出了空洞空间金字塔池化(atrous spatial pyramid pooling, ASPP)以聚合多尺度上下文。又因上述方法捕获的上下文信息较为粗糙,学者们在解决医学图像分割问题(张学峰等,2023)中引入注意力机制(罗恺锺等,2021),研究人员基于注意力机制提出了卷积注意力模块(convolutional block attention module, CBAM)(Woo等,2018)。总之,在解决重叠染色体的分割问题中,语义分割网络需要融合目标图像的上下文信息,提升网络的多尺度感受野,捕获目标的全局和局部特征信息,提高模型对染色体重叠场景和结构的理解能力,提升网络的分割性能。

1.2 模型设计

本文使用存在多条染色体构成的重叠染色体簇的数据集,选择基于语义分割的深度学习方法解决重叠染色体的分割问题。本文研究工作主要为:将染色体重叠区域、非重叠区域和背景区域视为3类,前期使用语义分割网络得到染色体图像中的重叠和非重叠区域,后期通过图形学算法重建染色体实例。本文模型的改进主要在增强目标的上下文表示信息,并通过实验结果证明本文模型可以提高重叠染

色体的分割准确性。

1.2.1 网络整体结构

U-Net++(Zhou 等, 2020)是基于经典语义分割网络 U-Net(Ronneberger 等, 2015)的改进。它实现了不同尺度和语义特征的融合,但它限制了网络在复杂场景下的表现,因此会导致特定信息丢失,如边界等细节特征被不断弱化。由于染色体重叠区域的识别依赖于准确的边界,针对重叠染色体分割问题,需要在网络中更多地保留原始图像的局部特征以及提高分割掩码的连续性。因此,本文设计了两阶段分割模型 SC-Net,包含对染色体重叠区域、非重叠区域和背景区域的分割网络 SC-UNet++,以及将重叠区域和非重叠区域重建为单条染色体的染色体实例

重建算法,如图 1 所示。其中,SC-UNet++使用 4 层编码器—解码器结构以达到浅层细节特征和深度语义信息的平衡,网络结构如图 2 所示。SC-UNet++在基线模型 U-Net++中增加了 HPM、CFM 和 CPAB 3 个模块信息,获取重叠染色体全局和局部的上下文信息。此外,SC-UNet++利用标注样本的类别信息,作为网络训练时的附加监督信息,区分目标图像中易混淆的空间特征。具体地说,在卷积节点 $X^{3,0}$ 之后加入一个 HPM 以捕获局部上下文,减轻下采样造成的低层信息丢失。另外,在卷积节点 $X^{2,2}$ 后加入一个 CFM 用于捕获局部和全局上下文。最后,在 CFM 之后加入一个 CPAB,利用 CFM 提取的上下文信息,在亲和力损失的深度监督下学习类内和类间相关

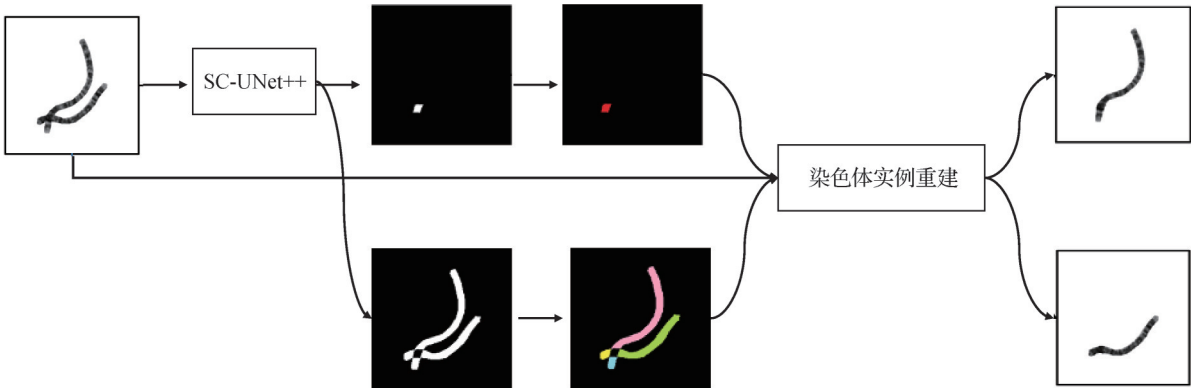


图 1 SC-Net 的两阶段框架示意图
Fig. 1 Schematic diagram of the two-stage framework of SC-Net

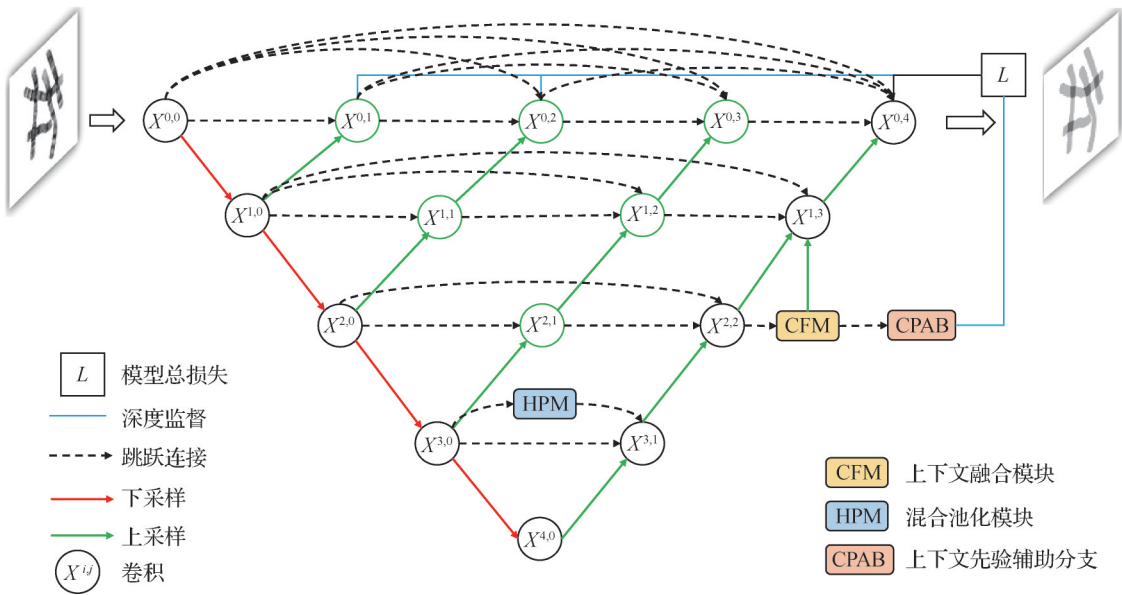


图 2 SC-UNet++网络结构图
Fig. 2 Network structure diagram of SC-UNet++

性,使网络能够更好地处理边界,产生更加自然的分割结果。

1.2.2 上下文融合模块

前文论述验证了上下文信息对重叠染色体分割的重要性。由于染色体重叠情况的复杂性,模型需要更加丰富和针对性的上下文信息。受到卷积注意力的启发,本文发现先后在通道和空间上应用注意力机制有利于语义分割网络的上下文提取。

因此,本文提出上下文融合模块CFM,它由高效通道注意力模块(efficient channel attention module,

ECAM)、混合池化模块HPM和空间注意力模块(spatial attention module, SAM)部分序列化组成,如图3所示。在重叠染色体分割任务中,CFM的任务是提供全局和局部上下文信息,其中,SAM专注于捕获图像的全局信息,包括染色体的总体排列和周围环境。HPM专注于局部上下文,以捕获染色体的细节和内部结构。CFM将这些不同层次的信息整合在一起,以提供更全面的上下文信息,有助于模型更准确地分割重叠染色体。这种综合性的上下文理解使模型能够应对复杂染色体排列,提高分割的准确性。

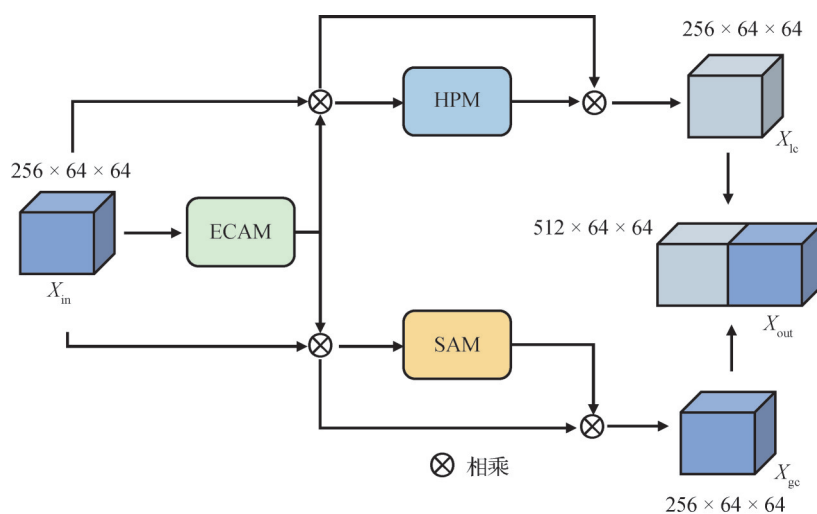


图3 上下文融合模块(CFM)结构图

Fig. 3 Context fusion module (CFM) structure diagram

在分割网络SC-UNet++中,CFM并联至解码器的中间层,并且使用高效通道注意力模块ECAM减少通道损失以捕获更清晰的上下文信息。具体地说,SC-UNet++解码器节点 $X^{2,2}$ 的输出特征进入CFM之后,首先经过ECAM提取输入特征的通道相关性,ECAM的输出和输入特征相乘后分别送入HPM和SAM。其中,HPM探索像素和周围区域的相关性以获取局部上下文;SAM则通过全局池化操作提取全局上下文。最后,将HPM输出特征下采样后与SAM输出特征融合,作为解码器节点 $X^{1,3}$ 的输入特征,继续进行网络的学习。

通道注意力分为挤压和激励两个步骤,挤压指将输入特征转化为一个固定值,激励指固定值进入全连接层以促使通道间的信息交互,最后通过激活函数产生通道权重(朱旗,2022)。一般需要通过全连接层降维以减少参数和计算量。Wang等人

(2019)发现对特征降维会造成信息损失。因此,他们提出了高效通道注意力ECA。因此,本文在CFM中使用ECA减少通道损失以捕获更清晰的上下文信息,如图4所示。其中,自适应的一维卷积的卷积核大小 k 为

$$k(C_{in}) = \left\lfloor \frac{1 + \log_2 C_{in}}{2} \right\rfloor \quad (1)$$

式中, C_{in} 为输入特征 X_{in} 的通道数。ECAM对输入特征 X_{in} 通道注意力权重 w_c 的计算过程为

$$w_c = \sigma \left(\text{Conv}_{a-1d} \left(\text{GAP} (X_{in}), k(C_{in}) \right) \right) \quad (2)$$

式中,GAP表示全局平均池化,Conv_{a-1d}表示自适应的一维卷积。 σ 为sigmoid激活函数。

相比于ECAM关注输入特征的通道相关性,SAM通过关注不同区域的相对关系以捕获空间上下文,如图5所示。

ECAM输出的 w_c 与 X_{in} 相乘后,作为SAM的输入

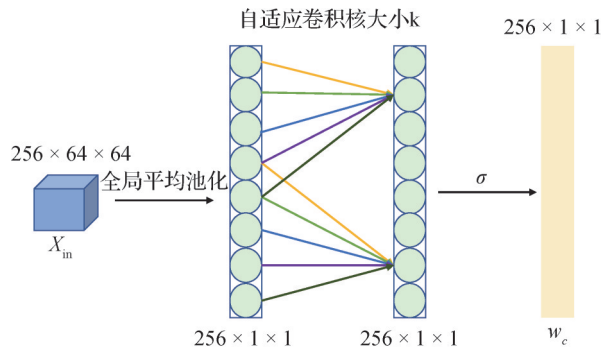


图4 高效通道注意力模块(ECAM)结构图

Fig. 4 Efficient channel attention module (ECAM) structure diagram

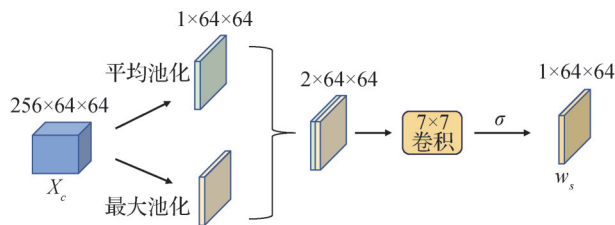


图5 空间注意力模块(SAM)结构图

Fig. 5 Spatial attention module (SAM) structure diagram

特征 X_c 。SAM 首先进行通道平均池化和最大池化以获得两个空间描述特征,将 X_{in} 特征拼接后使用 7×7 卷积对通道压缩,并通过 sigmoid 激活函数 σ 得到空间注意力权重 w_s 。最后 X_c 与 w_s 相乘以得到富含全局上下文的特征 X_{gc} 。具体流程为

$$X_{gc} = \sigma \left(Conv_{7 \times 7} \left(AP(X_c) + MP(X_c) \right) \right) \times X_c \quad (3)$$

式中, $Conv_{7 \times 7}$ 表示 7×7 的标准卷积, AP 为平均池化 (average pooling), MP 为最大池化 (max pooling), X_c 的计算式为

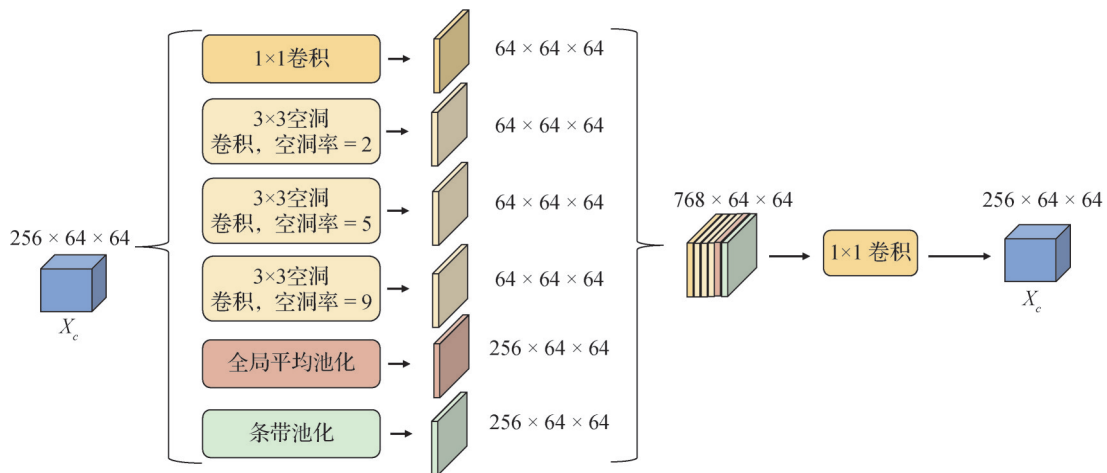


图6 混合池化模块(HPM)结构图

Fig. 6 Hybrid pooling module (HPM) structure diagram

$$X_c = w_c \times X_{in} \quad (4)$$

CFM 中还包含提取局部上下文的混合池化模块 HPM。另外,在 SC-UNet++ 的编码器卷积节点 $X^{3,0}$ 后也同时加入了 HPM 以提取多尺度局部上下文,因此,接下来对 HPM 单独展开论述。

1.2.3 混合池化模块

SAM 对于细节丰富的区域会因池化操作丢失局部空间信息。由于染色体的重叠区域同时包含两个染色体的部分特征,而不同的染色体的颜色、粗细和条带等细节特征不同,包含这些特征的局部上下文信息也很重要。因此,本文提出 HPM 使模型学习重叠染色体的局部上下文特征。其具体实现方式为:首先,将输入特征图分割成多个条带,每个条带在行和列维度上对应细长物体的一个部分;其次,对每个条带在行维度与列维度上应用池化操作,帮助模型理解染色体的横向关系、纵向关系和周围环境;之后,将每个条带在行和列维度上池化后的结果合并在一起,以生成条带池化的输出特征;最后,将条带池化的输出特征与 ASPP 模块的输出特征并联或连接,以综合利用它们的信息。HPM 结构如图 6 所示,其中,空洞卷积、条带池化与其他层并行,输入特征在经过不同池化和卷积之后,通道拼接并经过 1×1 卷积压缩通道后输出特征。

具体来讲,HPM 基于空洞空间金字塔池化 ASPP 进行改进。其中,空洞卷积是 ASPP 的核心,ASPP 通过不同空洞率的卷积核来实现多尺度感受野,将这些不同空洞率的特征图融合,有助于网络更好地理解不同尺度的对象。当卷积核大小为 k ,空

洞率为 r 时,感受野大小 n 的计算式为

$$n = k + (k - 1) \times (r - 1) \quad (5)$$

对于输入特征 X ,空洞卷积的输出特征 X_{dc} 的计算过程为

$$X_{dc}[i] = \sum_{k=1} X[i + r \times k] \times w[k] \quad (6)$$

式中, w 表示卷积核。式(6)表示在连续的卷积核间插入数量为 $r - 1$ 的零值,基于新的卷积核进行卷积,空洞卷积因而通过改变空洞率来改变感受野的大小。

虽然空洞卷积提升了感受野,但由于卷积核的

大小限制,卷积和池化操作只能在局部的方形区域内进行,而固定形状的局部特征会导致在提取细长物体(如染色体)的上下文时受到区域内无关因素干扰。Hou 等人(2020)设计条带池化(strip pooling)操作,通过对输入特征在行和列的不同形态下分别进行池化,有效地捕获细长物体的上下文信息。其中,细长物体通常具有复杂的边界,结合ASPP和条带池化可以提供更多的上下文信息,帮助模型更好地识别和分割物体的边界,能够适应不同的情境,提高在各种情况下的鲁棒性。其结构如图7所示。

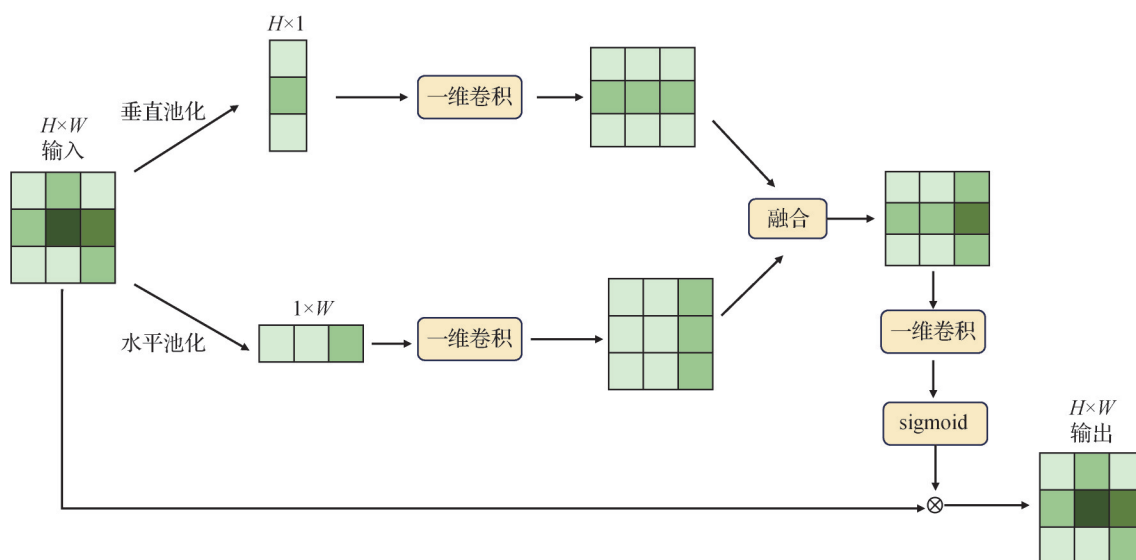


图7 条带池化结构图

Fig. 7 Strip pooling structure diagram

对于尺寸为 (H, W) 的输入特征 X ,条带池化首先分别沿着水平和垂直方向池化,计算过程为

$$Y_i^h = \frac{1}{W} \sum_{0 \leq j \leq W} X_{ij} \quad (7)$$

$$Y_j^v = \frac{1}{H} \sum_{0 \leq i \leq H} X_{ij} \quad (8)$$

式中, Y_i^h 与 Y_j^v 分别表示 X 第 i 行水平方向与第 j 列垂直方向池化后特征相应位置值。 $X_{i,j}$ 表示 X 在 (i, j) 位置处的值。之后通过一维卷积分别对 Y_i^h 和 Y_j^v 沿着水平和垂直方向扩展尺寸为 (H, W) 并进行融合,得到与 X 相同尺寸的特征 Y ,具体为

$$Y_{ij} = Y_i^h + Y_j^v \quad (9)$$

最后,将 Y 通过一维卷积 $Conv_{1d}$ 和sigmoid激活函数 σ 以生成条带池化的特征描述符,代表 X 中像素的空间相关性。输出特征 Z 为 X 与特征描述符的

哈达马积,具体为

$$Z_{ij} = X_{ij} \times \sigma(Conv_{1d}(Y)) \quad (10)$$

因此,本文在ASPP中加入条带池化,并将空洞率分别设置为2、5和9的大小以适应染色体的不同尺寸。

将HPM分别加入上述的CFM和解码器卷积节点 $X^{3,0}$ 后以获得多尺度的局部上下文提取能力。在CFM中,HPM的输入特征为CFM经过通道注意力对 X_m 加权得到的特征 X_c ,HPM的输出特征为 X_{lc} 。而在SC-UNet++的跳跃连接中,节点 $X^{3,0}$ 的输出特征为HPM的输入特征,并与HPM的输出特征拼接作为节点 $X^{3,1}$ 的输入特征。HPM可在多尺度感受野下全面捕捉局部上下文,在空间上进行细粒度的处理,提高模型对细节特征和空间特征的感知。

1.2.4 结合亲和力损失的深度监督

应用于模型训练和评估的标注训练集,还包含样本间的类别先验信息,即可以从标注中了解到像素间的区别。Yu 等人(2020)发现利用类别先验可以有效区分易混淆的空间信息,设计了 CPNet 以提升网络对复杂场景的理解。又鉴于重叠染色体分割和场景分割在对象复杂关系上的相似性,本文将上下文先验嵌入 CFM 之后并作为一个辅助

分支。

为了利用重叠染色体图像中的类别先验信息,本文基于训练样本的标注生成真实亲和矩阵作为训练时的附加监督信息。对于输入网络的图像 X ,经过骨干网络和上下文聚合模块得到尺寸均为 (H, W) 的全局上下文特征 X_{gc} 、局部上下文特征 X_{lc} 和融合全局—局部上下文的特征 X_{out} 。通过标注生成真实亲和矩阵 M_{ia} ,其计算过程如图 8 所示。

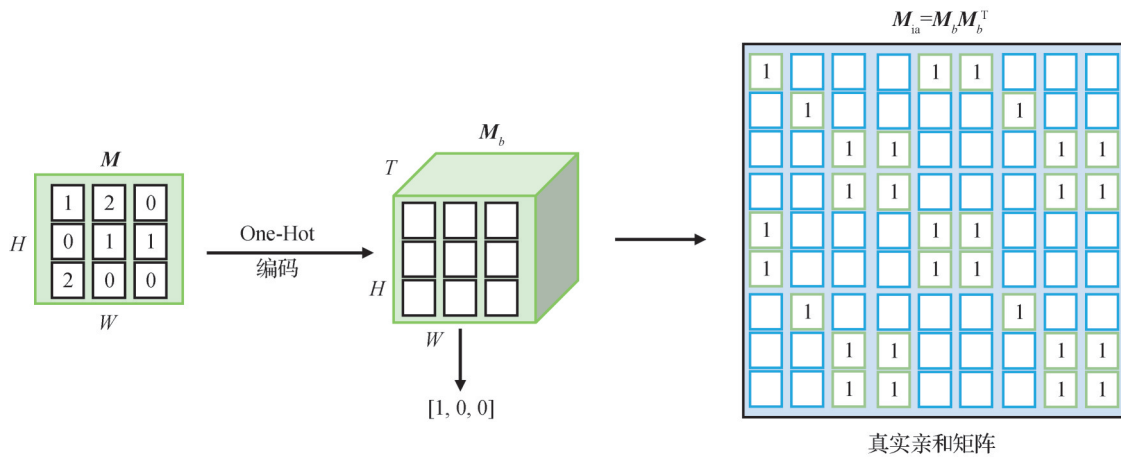


图 8 真实亲和矩阵的计算示意图

Fig. 8 Schematic diagram of the calculation of the real affinity matrix

首先,根据掩码标注 G 进行最近邻插值的下采样得到矩阵 M 。之后,使用 One-Hot 方法编码 M 中每个类别以得到尺寸为 (T, H, W) 的 0-1 矩阵 M_b 。其中 T 为类别数,对于重叠染色体分割问题, T 为 3。然后将 M_b 尺寸重塑为 (T, N) ,其中 $N = H \times W$ 。最后,计算真实亲和矩阵,具体计算为

$$M_{ia} = M_b M_b^T \quad (11)$$

式中, M_{ia} 是一个 0-1 对称阵,每行包含 M 中某位置像素与其他像素的关系,相同类别时对应行中位置的值为 1,否则为 0。 M_{ia} 对同类像素和非同类像素进行编码,包含丰富的类内先验信息, $1 - M_{ia}$ 则含有丰富的类间先验信息(叶妍青, 2021)。

图 9 展示了上下文先验辅助分支 CPAB 在网络中与 CFM 的结构关系。网络输入的重叠染色体图像经过 CFM 后,生成的局部上下文 X_{lc} 和全局上下文 X_{gc} 均为 $(256, 64, 64)$ 形状的张量。由于上下文先验需要局部空间信息推理类别关系,因此基于 X_{lc} 生成上下文先验矩阵 M_{cp} 。

具体地说, CPAB 通过卷积将 X_{lc} 的通道数增加

至 4 096,并重整为尺寸 $(4\ 096, 4\ 096)$ 的上下文先验矩阵 M_{cp} 。在亲和力损失的监督下, M_{cp} 可以对类内和类间像素编码,类内和类间关系可表示为

$$Y = M_{cp} \times \text{Reshape}(X_{lc}) \quad (12)$$

$$\bar{Y} = (1 - M_{cp}) \times \text{Reshape}(X_{gc}) \quad (13)$$

式中, Reshape 表示重整图像至一维。对于同类像素主要考虑与周围区域的连续性,因此类内关系基于局部上下文 X_{lc} 生成;而对于不同类别,与 CPNet 不同,本文考虑和所有区域的相关性以提升模型对重叠情况的理解,因此类间关系基于全局上下文 X_{gc} 生成。之后, Y 与 $\bar{Y}$ 经过拼接和卷积后生成尺寸 $(64, 64)$ 的辅助预测掩码 P_{cp} 。

根据 M_{cp} 和 M_{ia} ,将亲和力损失定义为

$$L_{aff} = \lambda_u L_u + \lambda_g L_g \quad (14)$$

式中,对于尺寸为 $N \times N$ 的 M_{cp} 和 M_{ia} , p_i 和 y_i 分别表示其中的元素。与 CPNet 相同, λ_u 和 λ_g 都设置为 1。 L_u 表示 M_{cp} 和 M_{ia} 的二元交叉熵损失,具体为

$$L_u = -\frac{1}{N^2} \sum_{i=1}^{N^2} (y_i \log p_i + (1 - y_i) \log (1 - p_i)) \quad (15)$$

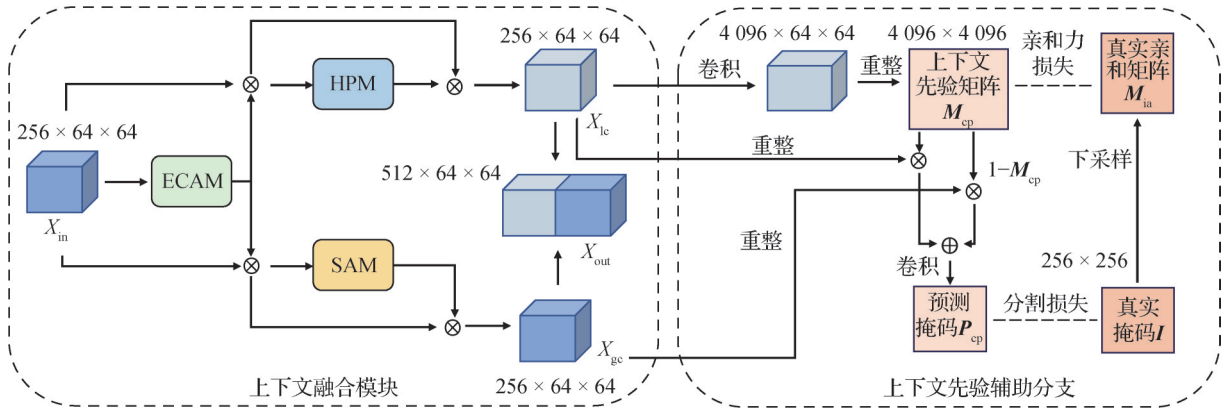


图9 CFM和上下文先验辅助分支(CPAB)结构图

Fig. 9 CFM and context prior auxiliary branch (CPAB) structure diagram

L_u 仅孤立地考虑了 M_{cp} 中的预测误差, 这会导致像素相关性丢失。由于 M_{ia} 的每行像素都与 I 某像素相关, 因此可以将 M_{ia} 中的像素分为类内像素和类间像素, 通过全局损失进行优化, 以使模型更好地推理染色体重叠关系。具体为

$$L_g = \frac{-\sum_{j=1}^N (E_j^p + E_j^r + E_j^s)}{N} \quad (16)$$

式中, E_j^p 、 E_j^r 和 E_j^s 分别表示 M_{ia} 中第 j 行的类内精度、类内召回率和特异度, 计算式为

$$E_j^p = \log \frac{\sum_{i=1}^N y_{ij} p_{ij}}{\sum_{i=1}^N p_{ij}} \quad (17)$$

$$E_j^r = \log \frac{\sum_{i=1}^N y_{ij} p_{ij}}{\sum_{i=1}^N y_{ij}} \quad (18)$$

$$E_j^s = \log \frac{\sum_{i=1}^N (1 - y_{ij})(1 - p_{ij})}{\sum_{i=1}^N (1 - p_{ij})} \quad (19)$$

由于染色体重叠区域的像素占比很小, 因此使用交叉熵损失会使得优化方向偏向占比大的类别, 如染色体非重叠区域和背景, 从而不能获得准确的分割结果。基于目标和预测重叠部分占比来度量分割准确性, Dice 损失可以关注像素级别的相似度, 能够很好地应对类别不平衡问题 (Milletari 等, 2016)。从这一点考虑, 本文模型使用三类别的 Dice 损失作为像素分类损失, 具体为

$$L_{Dice} = 1 - \frac{1}{C} \sum_{k=1}^C \frac{2 \sum_{i=1}^N y_{i,k} p_{i,k} + 1}{\sum_{i=1}^N y_{i,k} + \sum_{i=1}^N p_{i,k} + 1} \quad (20)$$

式中, C 在重叠染色体分割任务中为 3, $y_{i,k}$ 表示第 i 个像素点是否属于第 k 个类别, $p_{i,k}$ 表示模型预测该像素点属于第 k 个类别的概率, N 表示像素点数量。

由于反向传播的梯度消失问题, 模型梯度可能无法更新。而深度监督可以通过在多个分支上添加损失函数以提高梯度信号, 其在每一个分割分支中添加 Dice 损失函数, 可以强制 SC-UNet++ 学习更严格的特征表示。基于多尺度特征的深度监督, 网络的总损失为各个损失的加权之和, 具体为

$$L = w_d L_{dec} + w_{cp} L_{cp} + w_a L_{aff} + w_{a1} L_{aux1} + w_{a2} L_{aux2} \quad (21)$$

式中, L_{dec} 为解码器的分割损失, L_{cp} 为 CPAB 的分割损失, L_{aff} 为亲和力损失, L_{aux1} 和 L_{aux2} 分别为节点 $X^{0,1}$ 和 $X^{0,2}$ 辅助分支的分割损失。根据相关领域的经验, 权重 w_d 、 w_{a1} 和 w_{a2} 设置为 1、0.4、0.4。通过超参数实验研究, 权重 w_{cp} 和 w_a 设置为 0.75 和 0.5 时模型性能最优。

1.2.5 染色体实例重建

由于 SC-UNet++ 仅分离了重叠区域和非重叠区域, 而没有得到完整的单条染色体, 且分割边界可能不完整或出现细小的错误, 需要对模型分割结果进行后处理。如图 10 所示。图 10(a)(b) 分别表示重叠区域和非重叠区域粘连, 这些错误难以通过改进网络得到解决, 但是可以通过简单的形态学运算处理。

本文分别对 SC-UNet++ 输出的染色体重叠区域

和非重叠区域掩码进行开运算以将粘连的同类区域断开。开运算是图像处理中的一种形态学运算,由腐蚀操作和膨胀操作组成。基于开运算得到的区域,本文提出了染色体实例重建算法以构建单条染色体。算法输入染色体重叠区域集合 A 和非重叠区域集合 B ,以及倾斜角差别范围 β 。首先计算每个相邻非重叠和重叠区域边界两端连成直线的倾斜角,与相应的非重叠区域和重叠区域组合后形成三元组保存至 T 。然后迭代 T 中元素:将非重叠区域相同的元组,以及重叠区域相同且任意两组间边界倾斜角之差小于 β 的元组合并为同组,根据共同的重叠区域和相近倾斜角配对另一侧的非重叠区域。最后取每组中元素的并集,得到单条染色体区域。

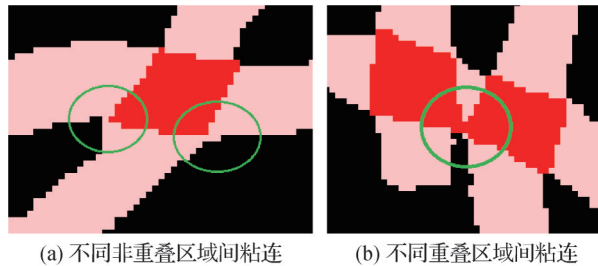


图10 分割边界的细微错误示意图

Fig. 10 Schematic illustration of a subtle error in the segmentation boundary ((a) adhesion between different non-overlapping regions; (b) adhesion between different overlapping regions)

使用相近的倾斜角来配对非重叠区域的原因在于染色体在重叠区域的粗细和弯曲方向变化很小,染色体被另一条染色体遮挡的两侧曲线的形状十分相似。算法中的倾斜角为弧度。通过对数据集的研究,本文发现重叠的两条染色体的夹角均大于 30° ,因此参数 $\beta < 0.53$ 时非重叠区域不会配对错误。经验证, β 取值为 0.4 时配对的准确率最高。使用该算法,本文发现绝大多数复杂的重叠染色体都可以有效重建,使得在分割网络不区分染色体实例的情况下仍然可以分离重叠染色体。

2 实验

2.1 实验方案

为了验证本文所构建的融合上下文信息的两阶段分割模型 SC-Net 的算法性能,在公共数据集 ChromSeg 上进行对比实验和消融实验。

在 SC-Net 第 1 阶段的重叠染色体分割任务中,采用交并比 (IoU) 作为模型的评价标准。考虑到染色体重叠区域像素较小导致评估指标显著降低,引入非重叠区域的 IoU 指标和所有类别的均值 mIoU (mean IoU) 指标,共同评价网络的分割性能。为了与现有模型的分割性能进行比较,本文还计算了重叠染色体的分割平均准确率 mACC 和检测准确率 ACC_{ov} 。

mACC 评估分割正确的像素与总像素数之比,计算过程为

$$mACC = \frac{1}{C} \sum_{c=1}^C \frac{TP_c}{TP_c + FP_c + FN_c} \quad (22)$$

式中, C 为类别数,在本文的研究中为 3。 TP_c 表示属于第 c 类并且被正确预测的像素数量, FP_c 表示不属于第 c 类但被错误预测为第 c 类的像素数量, FN_c 表示属于第 c 类但被错误预测为其他类别的像素数量。

ACC_{ov} 评价重叠染色体检测的准确率,具体为

$$ACC_{ov} = \frac{N_1 \times 100\%}{N} \quad (23)$$

式中, N_1 表示检测出的重叠染色体图像数, N 表示图像总数。

在 SC-Net 第 2 阶段的染色体实例重建算法中,参考 ChromSeg (Cao 等, 2020) 模型的评估,采用拼接准确率 ACC_{sp} 作为评价指标衡量算法性能。计算过程为

$$ACC_{sp} = \frac{N_2 \times 100\%}{N_1} \quad (24)$$

式中, N_2 表示第 2 阶段拼接形成单条染色体的图像。SC-Net 模型的总体准确率表示为 $ACC_{all} = ACC_{ov} \times ACC_{sp}$ 。

2.1.1 数据集构建

本文在 ChromSeg 数据集上进行实验,数据集中有 345 个交叉重叠染色体图像,来自于香港大学的 535 个尺寸为 256×256 像素的中期细胞图像中提取的重叠染色体。对图像进行两类标注:标注染色体重叠区域和非重叠区域,用于重叠染色体图像的重叠—非重叠的分割结果评估;标注单条染色体区域,用于评估染色体检测准确率。原始图像和两类标注分别如图 11 和图 12 所示。表 1 为原始中期图像中的相关统计数据,由于所使用的数据集的样本数量较小,本文对原始数据集随机采样并按照 2:1 的比

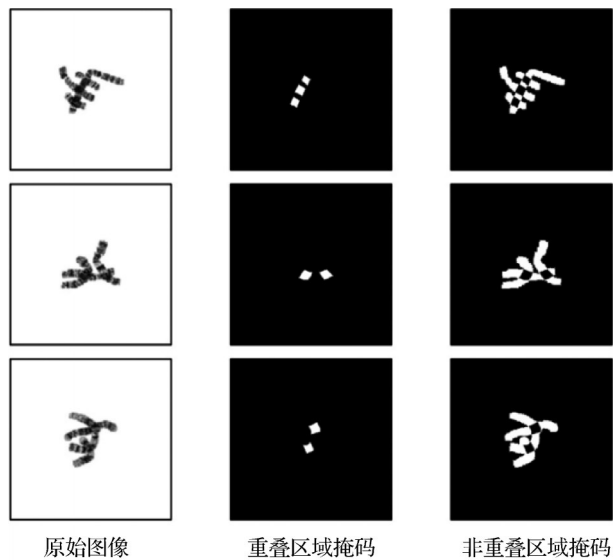


图 11 第 1 类标注的若干样本示意图

Fig. 11 Some sample diagrams of the first type of labeling

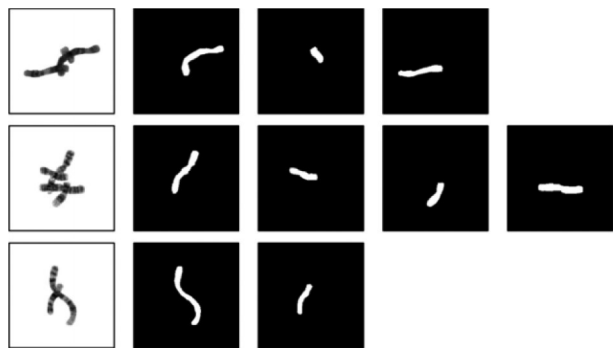


图 12 第 2 类标注的若干样本示意图

Fig. 12 Some sample diagrams of the second type of labeling

表 1 ChromSeg 数据集的中期图像中重叠染色体统计数据

Table 1 Statistics of overlapping chromosomes in metaphase images derived from the ChromSeg dataset		
重叠染色体	染色体数量	占比/%
2	2 920	11.86
3	1 542	6.27
4	880	3.58
≥5	945	3.84
总计	6 287	25.55

例划分训练集和测试集,且训练时使用随机水平和垂直翻转方法进行数据增强。

2.1.2 实验环境及参数设置

本文所有计算步骤和实验所用的硬件资源为内存 32 GB, CPU 是 3.3 GHz Intel Xeon, GPU 是 NVIDIA RTX 3070 的台式服务器,操作系统为

Ubuntu 18.04。本文所有模型和实验都基于版本为 0.30.0 的语义分割工具包 MMSegmentation 实现,并在 Python 3.8.10 和 PyTorch 1.10.0 的软件环境下运行。

采用 3 折交叉验证方法进行实验以验证模型在样本上的表现。其中,采用 ImageNet (Deng 等, 2009)数据集上的预训练模型进行权重初始化,使用随机梯度下降 (stochastic gradient descent,SGD) 优化器 (Robbins 和 Monro, 1951),学习率设置为 0.01,训练的最大轮次设置为 800,并通过比较分别为 8、16 和 32 的批次大小的模型分割结果,发现使用 16 为批次大小时的分割效果最好。

2.2 实验结果和分析

2.2.1 总体结果

为了验证 SC-Net 模型的分割和染色体重建性能,使用 DeepLabV3+ (Firdaus-Nawi 等, 2011)、U-Net (Ronneberger 等, 2015)、PSPNet (Zhao 等, 2017)、U-Net++ (Zhou 等, 2020)、ChromSeg (Cao 等, 2020)、UNet-2022 (Guo 等, 2022)、Klvanec (Klvanec 等, 2022)、SAN (side adapter network) (Xu 等, 2023) 的模型进行比较。表 2 展示了不同模型的对比结果。

表 2 SC-Net 与其他模型的对比实验

Table 2 Comparative experiments of SC-Net and other models				/%
模型	重叠 IoU	非重叠 IoU	mIoU	ACC _{ov}
DeepLabV3+	79.5	85.3	89.6	98.8
U-Net	78.7	87.1	88.6	98.2
PSPNet	79.9	89.7	89.8	98.6
U-Net++	79.6	88.6	88.7	98.5
ChromSeg	80.8	89.2	90.8	99.1
UNet-2022	80.4	91.2	90.4	98.8
Klvanec	80.2	88.4	90.4	98.7
SAN	80.6	90.9	90.7	99.0
SC-Net (本文)	83.5	90.3	91.3	99.4

注:加粗字体表示各列最优结果。

实验结果表明,本文提出的 SC-Net 在 ChromSeg 数据集上的性能在所有指标上均优于其他模型,在分割任务的性能上,该模型分割阶段获取的重叠区域 IoU 值达到了 83.5%,比最佳的单分割模型 SAN

高出 2.9%。在染色体实例重建任务的性能上,拼接的准确率达到 92.9%,模型的总体准确率达到 92.3%,比最佳的方法 ChromSeg 分别高出 2.7% 和 1.8%。SC-Net 优于这些模型的原因主要在于其能够提取类别相关性,从而使模型对重叠区域有更好的理解。

2.2.2 分割结果展示

首先展示 SC-Net 对若干幅重叠染色体图像分割并重建的结果,如图 13 所示。

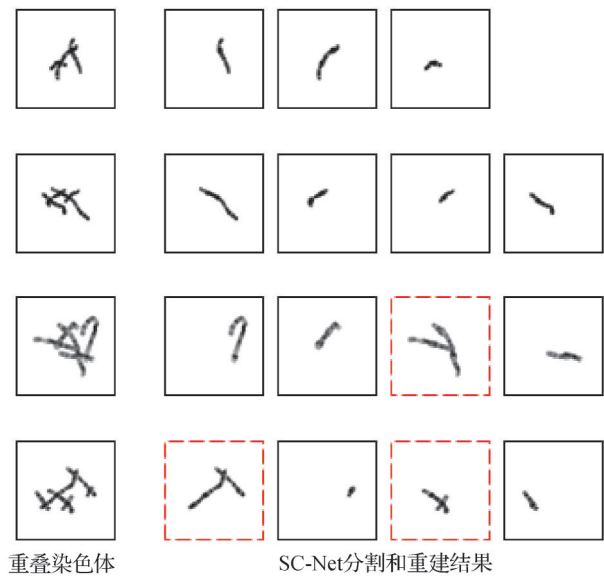


图 13 SC-Net 的染色体重建结果
Fig. 13 Chromosome reconstruction results of SC-Net

图 13 中,实线框包围的图像表示单条染色体正确重建,虚线框表示重建发生错误。从图中可以看出,对于较少的重叠染色体图像(3~4 条染色体),模型能够很好地分离单条染色体。但对于较多的重叠染色体图像(>4 条),结果中会出现未分离的染色体。这主要是由两方面因素造成的:

1)当图像中存在很多重叠染色体时,由于染色体的非刚性特点,染色体的弯曲折叠情况会更加严重,而若重叠染色体属于接触重叠、T 型重叠或 L 型重叠时,重叠染色体的形态和弯曲折叠的单条染色体会高度相似。因此语义分割模型会更加容易错误判别重叠的两条染色体为同一条。

2)染色体实例重建算法生成准确的单条染色体依赖于准确的重叠区域和非重叠区域掩码,以及准确的重叠区域和非重叠区域边界,若语义分割网络输出了不准确的掩码,会导致染色体实例重建算法

在配对两个非重叠区域时发生错误。

此外,将 SC-Net 和多个性能优秀的模型分割结果进行可视化对比,如图 14 所示。从结果中可以看出,若仅存在两条重叠染色体且在重叠区域不改变其轴线方向时,所有模型的分割结果都很好。但当染色体发生较大程度的弯曲时,U-Net++ 出现了严重的分割断裂,UNet-2022 在区域边界处也发生了分割错误。若图像中为多条重叠的染色体,U-Net++ 和 UNet-2022 出现了相近重叠区域的掩码粘连和缺失。ChromSeg 虽然改善了掩码缺失现象,但是掩码粘连仍然存在。这是由于这些模型缺少对复杂重叠情况的理解。而 SC-Net 通过特征的上下文融合,除了获得全局上下文信息以获取图像的细节特征外,还通过局部上下文更好地理解空间关系。具体地说,相较于其他几个模型,SC-Net 的分割结果中重叠与非重叠区域的边界更完整,并将非重叠区域准确划分为不同的封闭区域,这对染色体实例的成功重建非常关键。

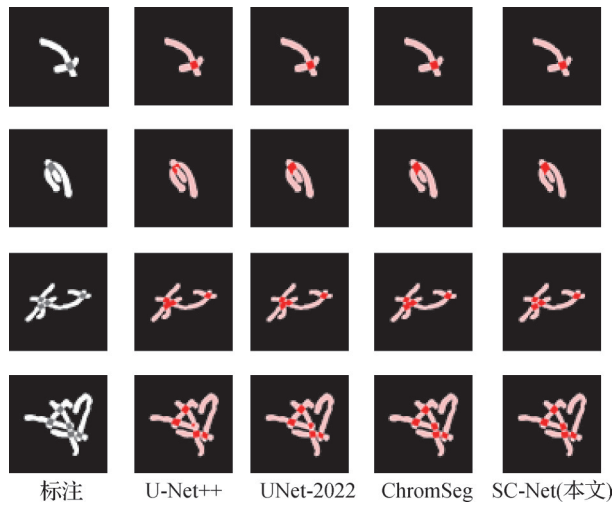


图 14 本文 SC-Net 与其他网络的语义分割结果对比

Fig. 14 Comparison of semantic segmentation results of SC-Net and other networks

除此之外,由于所使用的公开数据集的样本数量较小,训练时使用随机水平和垂直翻转方法进行数据增强,同时根据染色体和杂质的分离、背景分离和中期图像模拟的算法流程设计染色体模拟数据扩充算法进行实验,结果如表 3 所示。其中 ChromSeg 的重叠染色体簇是由香港大学的 535 幅中期图像所提取得到,Metaphase 数据集是本文从北京某医院收集了样本量为 100 的人类细胞中期图像,Metaphase+

表示经过模拟数据扩充算法扩充原始中期图像的训练集之后的数据集,之后的+1k、+2k、+3k、+5k 分别表示模拟数据扩充生成 1 000、2 000、3 000 和 5 000 幅图像。从表中可以看出,采用染色体模拟数据扩充算法提升了精度,并且随着扩充样本数量的增加,精

度也随之增加,扩充样本数量达到 2 000 时,模型的精度最高,总体准确率达到了 82.6%,与在 Chrom-Seg 数据集中 83.5% 的差距仅有不到 1% 的差距。实验分析表明,SC-Net 具有很好的泛化能力,能够在不对模型进行任何改变的前提下适应不同数据集。

表 3 SC-Net 在不同数据集的性能对比
Table 3 Performance comparison of SC-Net on different datasets

数据集	重叠 IoU	非重叠 IoU	mIoU	ACC _{ov}	ACC _{sp}	/%
						ACC _{all}
ChromSeg	83.5	90.3	91.3	99.4	92.9	92.3
Metaphase	79.6	89.1	89.9	98.5	90.8	89.4
Metaphase+1k	81.2	89.6	90.5	99.0	92.2	91.3
Metaphase+2k	82.6	89.8	90.9	99.2	92.5	91.8
Metaphase+3k	82.6	89.6	90.8	99.1	92.5	91.7
Metaphase+5k	82.6	89.6	90.8	99.1	92.5	91.7

2. 2. 3 消融实验

本文进行了消融实验以研究各个改进模块对分割结果的影响。由于染色体实例重建算法不改变分割结果,因此消融实验在语义分割阶段进行。在基线模型 U-Net++ 中设计 7 个模型,如表 4 所示。表中第 1 列代表应用不同改进的模型,BCE(binary cross entropy)为用于分割的二元交叉熵损失函数。HPM 表示在网络的跳跃连接处加入。CFM 包含的 HPM 不参与此消融实验。另外,由于 CPAB 基于 CFM 提取的局部和全局上下文生成上下文先验矩阵 M_{cp} ,因此 CPAB 仅在模型加入 CFM 时引入。

本文基于表 4 中的 7 个模型进行消融实验,用箱

线图展示模型分割结果中染色体重叠区域的 IoU 值和非重叠区域的 IoU 值,如图 15 所示。图中每个箱子的高度代表数据的四分位数分布范围,箱子上下延伸出去的触须代表极值。

本文对 7 种网络在多个评价指标上统计分析,如表 5 所示。具体来说,从染色体重叠区域和非重叠区域分割结果的对比角度出发,U-Net++(C)引入了 CFM 获得全局一局部的上下文信息,从而显著提升了重叠区域的 IoU 值,但对于非重叠区域反而恶化,性能不如基线模型。这是由于染色体非重叠区域占比更大,基线模型已经能够从局部感受野中获取对于分割而言足够的信息,而 CFM 引入的全局上下文包含噪声等不必要的信息。

因此,U-Net++(CH)在 U-Net++(C)的基础上引入 HPM。HPM 和 CFM 的共同使用大幅提升了重叠区域和非重叠区域的 IoU 值。这是由于 HPM 在 SC-UNet++ 的编码器阶段应用,虽然 CFM 中的局部上下文已经采用了 HPM 以进行局部区域上下文的提取,但由于 CFM 仅对尺寸为 (64, 64) 的特征执行上下文融合,而在网络的更深层会更容易丢失细节信息,因此,HPM 的引入使得 U-Net++(CH)获得了多尺度感受野,降低解码器多次下采样造成的信息丢失对获取局部上下文信息的影响。

U-Net++(CP)在重叠区域的分割性能略高于 U-Net++(CH),但对于非重叠区域,U-Net++(CP)

表 4 消融实验的网络配置表
Table 4 Network configuration table for ablation experiments

模型	CFM	HPM	CPAB	分割损失
U-Net++	-	-	-	BCE
U-Net++(C)	√	-	-	BCE
U-Net++(H)	-	√	-	BCE
U-Net++(CH)	√	√	-	BCE
U-Net++(CP)	√	-	√	BCE
U-Net++(CHP)	√	√	√	BCE
SC-UNet++	√	√	√	Dice

注:“√”表示采用,“-”表示未采用。

的 IoU 值远低于 U-Net++(CH)和 U-Net++(H)。由于 CPAB 需要准确的局部区域特征以生成高精度的上下文先验矩阵,而这依赖于 CFM 输入特征的质量。U-Net++中没有从编码器直接输入解码器对

应层级特征的跳跃连接,而是用密集的卷积节点、上采样和下采样代替,而 HPM 的引入使得模型在上采样阶段能够通过多尺度感受野主动地挖掘局部上下文,从而提升上下文先验矩阵 M_{cp} 的质量。

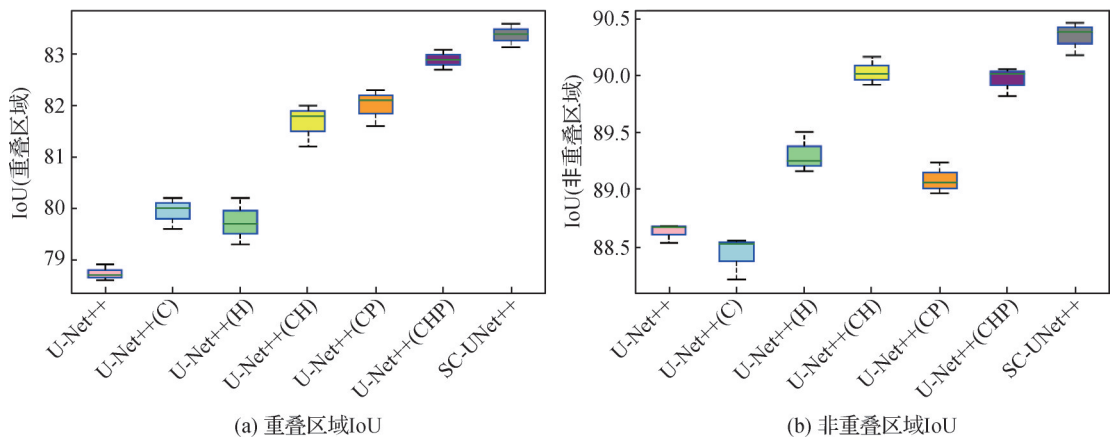


图 15 7 种网络 3 折交叉验证的 IoU 箱线图
Fig. 15 IoU boxplot of seven network 3-fold cross-validation((a) overlap area IoU;(b)non-overlapping area IoU)

U-Net++(CHP)的 IoU 值和 SC-UNet++相当,而对于非重叠区域的 IoU 值低于 U-Net++(CH)。这是由于 Dice 损失对于类别不平衡问题的适应性更佳,它关注正确像素数目占总像素数目的比例。对于重叠染色体分割问题,重叠区域的像素数目较少,这导

致交叉熵损失函数在训练过程中偏向于优化像素占比更多的类别分类结果。

总体来看,SC-UNet++在所有评价指标上均优于所比较的模型,显著提高了重叠染色体分割准确率,证明了本文方法的有效性。

表 5 7 种网络在不同评价指标下的性能对照表
Table 5 Performance comparison table of seven networks under different evaluation indicators

模型	mIoU	重叠 IoU	非重叠 IoU	背景 IoU	mACC
U-Net++	88.72	78.73	88.62	99.59	94.56
U-Net++(C)	89.23	79.91	88.42	99.58	94.88
U-Net++(H)	88.96	79.69	89.30	99.60	94.96
U-Net++(CH)	90.54	81.69	90.03	99.63	95.08
U-Net++(CP)	90.30	82.23	89.08	99.59	95.17
U-Net++(CHP)	90.82	82.84	89.96	99.62	95.38
SC-UNet++	91.31	83.47	90.34	99.65	95.54

2. 2. 4 模型复杂度分析

为了考虑模型的复杂度对训练和推理时间的影响,本文在相同的实验硬件设备,即硬件资源为内存 32 GB, CPU 是 3. 3 GHz Intel Xeon, GPU 是 NVIDIA RTX 3070 的台式服务器下进行实验,对模型的推理速度进行了补充,具体结果如表 6 和表 7 所示。

表 6 结果显示,在处理相同任务时,完全不同的网络结构,其模型推理速度相差很大,如推理速度相对 U-Net 模型比本文两阶段模型 SC-Net 要快 21. 8 ms;而相同的网络结构,如同两阶段网络模型 ChromSeg 也因模块分支的差别比 SC-Net 快 12. 3 ms。

表 6 SC-Net 与其他模型的推理速度对比
Table 6 Comparison of inference speed between SC-Net and other models

模型	推理速度 (/ms/幅)
DeepLabV3+	66.2
U-Net	33.4
PSPNet	64.8
U-Net++	37.9
ChromSeg	42.9
UNet-2022	36.1
Klivanec	44.0
SC-Net(本文)	55.2

表 7 考虑不同的输入对 SC-Net 推理速度的影响, 根据图 2 中 SC-UNet++ 的网络结构图所示, 若 CFM 输入的是 $X^{3,1}$, 即表明输入特征尺寸减半, 推理速度明显加快, 但模型性能随之降低。综上所述, 模型的复杂度对训练和推理时间都有显著影响。

2. 2. 5 超参数选取分析

由于模型中的亲和力辅助分支除了引入亲和力损失外, 还附加了与其余分支相同的 Dice 损失。其中, Dice 损失的权重设置会影响模型对分割准确性和边缘细节的关注程度, 较高的权重会使模型更加关注分割边缘的细节, 较低的权重则会导致模型高度关注整体准确性而忽略细节。如果对亲和力损失设置过高的权重, 模型会过于关注同类别像素的一致性, 进而导致分割结果出现过度平滑, 损失细节信息。

表 7 模型输入对 SC-Net 推理速度的影响
Table 7 The effect of model inputs on SC-Net inference speed

模型	重叠IoU/%	非重叠IoU/%	mIoU/%	ACC _{ov} /%	推理速度/(ms/幅)
SC-Net	83.5	90.3	91.3	99.4	55.2
SC-Net($X^{3,1}$)	79.2	87.6	89.0	98.5	45.1

因此, 本文对 SC-UNet++ 中的损失权重进行研究, 设计不同组合进行实验。涉及的超参数有 CPAB 的分割损失权重 w_{cp} 以及亲和力损失权重 w_a 。选择重叠区域 IoU 值、非重叠区域 IoU 值以及总体准确率 ACC_{all} 作为评价指标, 结果如表 8 所示。由于辅助分支的损失权重通常小于解码器权重, 本文在实验中将 w_{cp} 的取值范围设置为小于 1 的值。从表中可以看出, SC-Net 的超参数 w_{cp} 和 w_a 取值分别为 0.75 和 0.5 时, 该模型获得了最高的染色体重叠区域 IoU 值和总体准确率 ACC_{all}。总体上看, SC-Net 的模型性能对超参数具有较低的敏感程度, 模型的总体准确率均高于 90%。

3 结 论

本文基于重叠染色体图像的分析, 为解决多条形态多样的染色体形成的重叠染色体簇的分割问题, 构建了一种融合上下文信息的两阶段网络模型 SC-Net。该模型第 1 阶段的分割网络在 U-Net++ 中增加 CFM 提取全局和局部上下文信息, 在 CFM 后和解码器中引入 HPM 以提取多尺度局部上下文, 并在

表 8 不同超参数组合下 SC-Net 的分割和重建性能
Table 8 Segmentation and reconstruction performance of SC-Net under different hyperparameter combinations

w_{cp}	w_a	重叠区域 IoU/%	非重叠区域 IoU/%	ACC _{all} /%
0.25	0.5	81.15	89.60	90.9
0.5	0.5	81.71	88.12	90.8
0.75	0.5	83.47	90.34	92.3
1	0.5	82.67	90.75	92.0
0.25	1	81.44	88.89	90.9
0.5	1	80.50	88.13	90.3
0.75	1	81.94	89.82	91.1
1	1	82.15	90.46	91.2
0.25	2	81.45	89.03	90.8
0.5	2	82.29	89.10	91.4
0.75	2	82.22	89.63	91.6

注: 加粗字体表示各列最优结果。

CFM 后引入 CPAB, 在亲和力损失的监督下关注染色体的重叠关系, 实现了重叠染色体的准确分割。该模型第 2 阶段提出了用于拼接单条染色体的染色

体实例重建算法,在不同的重叠染色体分割场景中准确完成单条染色体拼接。

实验结果表明,SC-Net充分提取和利用了重叠染色体图像的上下文信息和类别先验信息,并在不同数量的重叠染色体的分割场景中都表现出很好的性能。然而在细胞分裂中期图像中,染色体还存在接触现象,后续研究将考虑扩充新型重叠和接触染色体数据集,重构新的分割模型实现重叠与接触染色体的完整分割,并设计一个新的染色体实例重建的后处理过程。其次,考虑实际应用中,因染色体准备技术、成像设备性能等导致染色体图像的质量差异,通过专家干预、评估图像特征、多模态融合技术等手段处理低质量图像,保证分割的性能。之后,将运用模型压缩、迁移学习和知识蒸馏等方法降低模型复杂度,保证模型轻量化,从而提高临床的可用性。

参考文献(References)

- Altinsoy E, Yilmaz C, Wen J, Wu L Q, Yang J and Zhu Y M. 2019. Raw G-band chromosome image segmentation using U-Net based neural network//Proceedings of the 18th International Conference on Artificial Intelligence and Soft Computing. Zakopane, Poland: Springer: 117-126 [DOI: 10.1007/978-3-030-20915-5_11]
- Cao X, Lan F Z, Liu C M, Lam T W and Luo R B. 2020. ChromSeg: two-stage framework for overlapping chromosome segmentation and reconstruction//Proceedings of 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Seoul, Korea (South): IEEE: 2335-2342 [DOI: 10.1109/BIBM49941.2020.9313458]
- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A L. 2018. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4): 834-848 [DOI: 10.1109/TPAMI.2017.2699184]
- Chen P J, Cai J Z and Yang L. 2020. Chromosome segmentation via data simulation and shape learning//Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Montreal, Canada: IEEE: 1637-1640 [DOI: 10.1109/EMBC44109.2020.9176020]
- Deng J, Dong W, Socher R, Li L J, Li K and Li F F. 2009. ImageNet: a large-scale hierarchical image database//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE: 248-255 [DOI: 10.1109/CVPR. 2009.5206848]
- Fan J H, Qian Z and Qiao S Q. 2009. Research on chromosome image segmentation. Journal of Qiqihar Medical College, 30(23): 2927-2928 (范继红, 钱真, 乔树清. 2009. 染色体图像分割的研究. 齐齐哈尔医学院学报, 30(23): 2927-2928) [DOI: 10.3969/j.issn.1002-1256.2009.23.042]
- Firdaus-Nawi M, Noraini O, Sabri M Y, Siti-Zahrah A, Zamri-Saad M, and Latifah H. 2011. DeepLabv3+ _encoder-decoder with Atrous separable convolution for semantic image segmentation. Pertanika Journal of Tropical Agricultural Science, 34(1): 137-143
- Gorelick L, Galun M, Sharon E, Basri R and Brandt A. 2006. Shape representation and classification using the Poisson equation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(12): 1991-2005 [DOI: 10.1109/TPAMI.2006.253]
- Grau V, Mewes A U J, Alcaniz M, Kikinis R and Warfield S K. 2004. Improved watershed transform for medical image segmentation using prior information. IEEE Transactions on Medical Imaging, 23(4): 447-458 [DOI: 10.1109/TMI.2004.824224]
- Grisan E, Poletti E and Ruggeri A. 2009. Automatic segmentation and disentangling of chromosomes in Q-band prometaphase images. IEEE Transactions on Information Technology in Biomedicine, 13(4): 575-581 [DOI: 10.1109/ITTB.2009.2014464]
- Guangzhou Electronic Technology Co., Ltd., Chinese Academy of Sciences. 2022. Chromosome image instance segmentation method and device based on improved Mask RCNN. China, 201911021995.9 (中国科学院广州电子技术有限公司. 2022. 基于改进 Mask RCNN 的染色体图像实例分割方法及装置. 中国, 201911021995.9)
- Guo J S, Zhou H Y, Wang L S and Yu Y Z. 2022. UNet-2022: exploring dynamics in non-isomorphic architecture [EB/OL]. [2023-08-10]. <http://arxiv.org/pdf/2210.15566.pdf>
- Hosang J, Benenson R and Schiele B. 2017. Learning non-maximum suppression//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 6469-6477 [DOI: 10.1109/CVPR.2017.685]
- Hou Q B, Zhang L, Cheng M M and Feng J S. 2020. Strip pooling: rethinking spatial pooling for scene parsing//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 4002-4011 [DOI: 10.1109/CVPR42600.2020.00406]
- Hu R L, Karnowski J, Fadely R and Pommier J P. 2017. Image segmentation to distinguish between overlapping human chromosomes. [EB/OL]. [2023-08-10]. <http://arxiv.org/pdf/1712.07639.pdf>
- Jing X, Teng Q Z, Tang T and Wang S. 2008. Efficient method for segmentation of chromosome images. Computer Engineering and Applications, 44(35): 182-184, 187 (敬雪, 滕奇志, 唐棠, 王姝. 2008. 有效的染色体分割新方法. 计算机工程与应用, 44(35): 182-184, 187) [DOI: 10.3778/j.issn.1002-8331.2008.35.054]
- Klivanec D, Phillips T B, McCaffrey K J W and Al Moubayed N. 2022. Using orientation to distinguish overlapping chromosomes//Proceedings of the 31st International Conference on Artificial Neural Net-

- works. Bristol, UK: Springer: 391-403 [DOI: 10.1007/978-3-031-15919-0_33]
- Lin C C, Zhao G S, Yin A H, Ding B C, Guo L and Chen H B. 2020. AS-PANet: a chromosome instance segmentation method based on improved path aggregation network architecture. *Journal of Image and Graphics*, 25(10): 2271-2280 (林成创, 赵淦森, 尹爱华, 丁笔超, 郭莉, 陈汉彪. 2020. AS-PANet: 改进路径增强网络的重叠染色体实例分割. *中国图象图形学报*, 25(10): 2271-2280) [DOI: 10.11834/jig.200236]
- Lin C C, Zhao G S, Yin A H, Yang Z R, Guo L, Chen H B, Zhao L, Li S Y, Luo H Y and Ma Z H. 2021. A novel chromosome cluster types identification method using ResNeXt WSL model. *Medical Image Analysis*, 69: #101943 [DOI: 10.1016/j.media.2020.101943]
- Lin T Y, Dollár P, Girshick R, He K M, Hariharan B and Belongie S. 2017. Feature pyramid networks for object detection//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 936-944 [DOI: 10.1109/CVPR.2017.106]
- Liu S, Qi L, Qin H F, Shi J P and Jia J Y. 2018. Path aggregation network for instance segmentation//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE: 8759-8768 [DOI: 10.1109/CVPR.2018.00913]
- Long J, Shelhamer E and Darrell T. 2015. Fully convolutional networks for semantic segmentation//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, USA: IEEE: 3431-3440 [DOI: 10.1109/CVPR.2015.7298965]
- Luo C L and Zhao Y. 2023. Review of deep learning methods for karyotype analysis. *Journal of Image and Graphics*, 28(11): 3363-3385 (罗纯龙, 赵屹. 2023. 染色体核型分析深度学习的方法综述. *中国图象图形学报*, 28(11): 3363-3385) [DOI: 10.11834/jig.221094]
- Luo K K, Wang T and Ye F F. 2021. U-Net segmentation model of brain tumor MR image based on attention mechanism and multi-view fusion. *Journal of Image and Graphics*, 26(9): 2208-2218 (罗恺锴, 王婷, 叶芳芳. 2021. 引入注意力机制和多视角融合的脑肿瘤MR图像U-Net分割模型. *中国图象图形学报*, 26(9): 2208-2218) [DOI: 10.11834/jig.200584]
- Madian N, Jayanthi K B and Suresh S. 2015. Contour based segmentation of chromosomes in G-band metaphase images//*Proceedings of 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. Orlando, USA: IEEE: 943-947 [DOI: 10.1109/GlobalSIP.2015.7418336]
- Milletari F, Navab N and Ahmadi S A. 2016. V-Net: fully convolutional neural networks for volumetric medical image segmentation//*Proceedings of the 4th International Conference on 3D Vision (3DV)*. Stanford, USA: IEEE: 565-571 [DOI: 10.1109/3DV.2016.79]
- Minaee S, Fotouhi M and Khalaj B H. 2014. A geometric approach to fully automatic chromosome segmentation//*Proceedings of 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. Philadelphia, USA: IEEE: 1-6 [DOI: 10.1109/SPMB.2014.7163174]
- Robbins H and Monro S. 1951. A stochastic approximation method. *The Annals of Mathematical Statistics*, 22(3): 400-407 [DOI: 10.1214/aoms/1177729586]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Saleh H M, Saad N H and Isa N A M. 2019. Overlapping chromosome segmentation using U-Net: convolutional networks with test time augmentation. *Procedia Computer Science*, 159: 524-533 [DOI: 10.1016/j.procs.2019.09.207]
- Sharma M, Saha O, Sriraman A, Hebbalaguppe R, Vig L and Karande S. 2017. Crowdsourcing for chromosome segmentation and deep classification//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu, USA: IEEE: 786-793 [DOI: 10.1109/CVPRW.2017.109]
- Sun X F, Li J M, Ma J L, Xu H Q, Chen B, Zhang Y F and Feng T. 2021. Segmentation of overlapping chromosome images using U-net with improved dilated convolutions. *Journal of Intelligent and Fuzzy Systems*, 40(3): 5653-5668 [DOI: 10.3233/JIFS-201466]
- Wang G J. 2022. Research on End-to-End Chromosome Segmentation Method based on Adaptive and Regression Correction Network. Xuzhou: China University of Mining and Technology (王广杰. 2022. 基于自适应与回归修正网络的端到端染色体分割方法研究. 徐州: 中国矿业大学) [DOI: 10.27623/d.cnki.gzkyu.2022.000244]
- Wang G T, Li W Q, Ourselin S and Vercauteren T. 2019. Automatic brain tumor segmentation using convolutional neural networks with test-time augmentation//*4th International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Granada, Spain: Springer: 61-72 [DOI: 10.1007/978-3-030-11726-9_6]
- Wang P L, Hu W J, Zhang J P, Wen Y F, Xu C M and Qian D H. 2021. Enhanced rotated mask R-CNN for chromosome segmentation//*Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Mexico: IEEE: 2769-2772 [DOI: 10.1109/EMBC46164.2021.9630695]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//*Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich, Germany: Springer: 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Xie S N, Girshick R, Dollár P, Tu Z W and He K M. 2017. Aggregated residual transformations for deep neural networks//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*

- tion. Honolulu, USA: IEEE: 5987-5995 [DOI: 10.1109/CVPR.2017.634]
- Xu M D, Zhang Z, Wei F Y, Hu H and Bai X. 2023. Side adapter network for open-vocabulary semantic segmentation//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 2945-2954 [DOI: 10.1109/CVPR52729.2023.00288]
- Ye Y Q. 2021. Joint Segmentation of Multiple Types of Effusions in Retinal OCT Images. Suzhou: Soochow University (叶妍青. 2021. 视网膜 OCT 图像中多类积液的联合分割. 苏州: 苏州大学) [DOI: 10.27351/d.cnki.gszhu.2021.000822]
- Yilmaz I C, Yang J, Altinsoy E and Zhou L. 2018. An improved segmentation for raw G-band chromosome images//Proceedings of the 5th International Conference on Systems and Informatics (ICSAI). Nanjing, China: IEEE: 944-950 [DOI: 10.1109/ICSAI.2018.8599328]
- Vincent L. 1993. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. IEEE Transactions on Image Processing, 2(2): 176-201 [DOI: 10.1109/83.217222]
- Yu C Q, Wang J B, Gao C X, Yu G, Shen C H and Sang N. 2020. Context prior for scene segmentation//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 12413-12422 [DOI: 10.1109/CVPR42600.2020.01243]
- Zhang L, Zhang X L, Lu L and Liu H. 2022. Image segmentation of chromosomes based on residual U-Net. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 34(2): 208-215 (张林, 张雪利, 路霖, 刘辉. 2022. 基于残差 U-Net 网络的染色体图像分割方法. 电大学报(自然科学版), 34(2): 208-215) [DOI: 10.3979/j.issn.1673-825X.202010160309]
- Zhang X F, Zhang S, Zhang D H and Liu R. 2023. Group attention-based medical image segmentation model. Journal of Image and Graphics, 28(10): 3231-3242 (张学峰, 张胜, 张冬晖, 刘瑞. 2023. 引入分组注意力的医学图像分割模型. 中国图象图形学报, 28(10): 3231-3242) [DOI: 10.11834/jig.220748]
- Zhao H S, Shi J P, Qi X J, Wang X G and Jia J Y. 2017. Pyramid scene parsing network//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 6230-6239 [DOI: 10.1109/CVPR.2017.660]
- Zhou Z W, Siddiquee M M R, Tajbakhsh N and Liang J M. 2020. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. IEEE Transactions on Medical Imaging, 39(6): 1856-1867 [DOI: 10.1109/TMI.2019.2959609]
- Zhu Q. 2022. Research on Medical Image Segmentation Method based on Improved U-Net. Guangzhou: Guangzhou University (朱旗. 2022. 基于改进 U-Net 的医学图像分割方法研究. 广州: 广州大学) [DOI: 10.27040/d.cnki.ggzdu.2022.001247]

作者简介

- 焦润海,男,教授,博士生导师,主要研究方向为机器学习、数据挖掘和图像识别。E-mail: runhaijiao@ncepu.edu.cn
- 褚佳杰,女,硕士研究生,主要研究方向为机器学习及其在电力系统中的应用。E-mail: chujiajie@ncepu.edu.cn
- 刘嘉骥,女,博士研究生,主要研究方向为计算机视觉和视频图像处理。E-mail: liujiajiran@163.com
- 余济民,男,硕士,主要研究方向为计算机视觉和医学图像处理。E-mail: yujimin4wo@163.com