

YOLO-MaskRCNN：一种改进的血细胞分类统计模型

王家琛

福建师范大学数学与统计学院，福建 福州

收稿日期：2025年10月13日；录用日期：2025年11月4日；发布日期：2025年11月14日

摘 要

血液中血细胞的数量对疾病诊断具有重要意义。尽管全自动血细胞计数仪已普及，但仍需人工涂片和显微镜形态学检查，工作量较大。为此本文提出一种改进的血细胞识别和计数模型：在早期YOLO模型中引入颜色注意力机制，利用染色后不同血细胞在颜色上的差异，通过小型全卷积网络对特征图进行加权，使模型更加关注颜色特征，从而提升血细胞的检测与分类效果，并保持较低的计算开销。另外，针对血液涂片中粘连和重叠红细胞的计数难题，采用Mask R-CNN算法实现实例分割，相较于传统的分水岭和轮廓检测方法，能更准确地地区分粘连红细胞。实验证明改进后的模型训练成本更低且识别准确度更高，最后，开发了一个基于Python的开源桌面程序，实现血细胞识别、计数与分割的一体化，降低了工作者的负担。

关键词

生物统计，细胞计数，目标检测，颜色注意力，YOLO，Mask R-CNN

YOLO-MaskRCNN: An Improved Statistical Model for Blood Cell Classification

Jiachen Wang

School of Mathematics and Statistics, Fujian Normal University, Fuzhou Fujian

Received: October 13, 2025; accepted: November 4, 2025; published: November 14, 2025

Abstract

The quantity of blood cells plays a crucial role in disease diagnosis. Although automated hematology

analyzers have become widely used, manual smear preparation and microscopic morphological examination are still required, resulting in heavy workloads. To address this issue, this paper proposes an improved blood cell detection and counting model. A color attention mechanism is introduced into an early YOLO model to exploit the color differences among various blood cells after staining. A lightweight fully convolutional network is employed to weight the feature maps, enabling the model to focus more on color features and thereby enhance detection and classification performance while maintaining low computational cost. In addition, to tackle the challenge of counting adherent and overlapping red blood cells in blood smear images, the Mask R-CNN algorithm is adopted for instance segmentation. Compared with traditional watershed and contour-based methods, this approach can more accurately distinguish adherent red blood cells. Experimental results demonstrate that the improved model achieves lower training cost and higher recognition accuracy. Finally, a Python-based desktop application has been developed to integrate blood cell detection, counting, and segmentation, effectively reducing the workload of laboratory personnel.

Keywords

Biostatistics, Cell Counting, Object Detection, Color Attention, YOLO, Mask R-CNN

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

血细胞存在于人体血液当中，能随血液流动遍及全身。血细胞计数结果是衡量人体健康状况的重要指标之一，它可以提供有关血液中不同类型细胞数量的信息。主要的血细胞包括红细胞、白细胞和血小板。红细胞(RBC)计数用于评估负责输送氧气到身体各部分的红细胞数量。低于正常范围可能表示贫血或其他血液疾病。白细胞(WBCs)是免疫系统的一部分，负责抵抗感染和疾病。高于正常水平可能表示感染、炎症或其他疾病。血小板计数(Platelet)血小板是帮助止血和凝血的细胞，高或低于正常范围的血小板计数可能与出血或凝血异常有关。

过去对于病人的血细胞计数都是采用人工方法，在光学显微镜下观察病人血液样品制成的血涂片(Peripheral Blood Smear)借助专门的细胞计数板肉眼计数，这种方法速度慢而且过程繁琐，也容易出错。而现在大多数医院均配备了血细胞分析仪，该类仪器可以通过细小的管道来吸取少量的标本。在这种管道内，有负责对经过的细胞的数量加以计数的光学和电阻抗传感器，并且借助细胞大小对细胞类型加以鉴别[1]。尽管血细胞分析仪能够提供血细胞的数量和一些基本特征信息，但计数任务其依然需要人工辅助，特别是在复杂病例的诊断中，例如对于某些遗传性疾病或特定器官功能异常的诊断需要其他类型的检测或影像学方法。而某些情景下需要相关工作人员肉眼对血液涂片中的细胞进行识别并计数，此工作量巨大，且准确度完全依赖工作人员本身，而这类任务在今天可以通过目标检测自动完成。

现如今深度学习及计算机视觉领域蓬勃发展，各种目标检测算法也在生物统计当中，目标检测可以应用在血涂片检查中，用神经网络模型实现对血细胞的分类并计数，成本较低，配合光学显微镜即可完成计数和形态学检查，相比血细胞分析仪目标检测可以提供更丰富的数据和信息，例如除了细胞数量之外，还可以获取细胞大小、形态、分布等更详细的特征信息，这对于研究细胞功能和病理特征有更大的帮助。一些目标检测方法可以实现实时监测和分析，例如针对细胞培养的实时细胞计数，或者通过图像处理实时分析组织切片中的细胞数量和特征。

2. 国内外研究现状

目标检测的任务是识别图像中的物体，不仅要知道它们的位置，还要知道它们是什么。早期阶段的目标检测主要是手工设计特征，再通过这些特征对图像进行边缘、纹理颜色分析，然后结合 SVM 等分类器进行分类。这类方法鲁棒性较差，准确度和速度较慢。随着深度学习发展和硬件的提升，2010 年之后出现了很多优秀的目标检测算法。比如 R-CNN 算法，将卷积神经网络应用于目标检测中。该方法先通过候选区域提议再用卷积神经网络提取特征，最后使用 SVM 进行目标分类。后续的 Fast R-CNN 和 Faster R-CNN 进一步优化了速度和准确性，引入了区域兴趣池化和区域生成网络等技术。随后又出现了 YOLO (You Only Look Once)、SSD (Single Shot MultiBox Detection) 等单阶段目标检测算法，将目标分类和检测任务只通过一个深度卷积网络回归完成，大大提高了目标检测的速度。各种目标检测算法的出现促使细胞计数任务可以通过计算机程序自动完成，提高了医护人员的工作效率，且计数和识别的质量和速度正在不断提升。国内外研究人员在该领域展开了诸多研究，创新和改进了很多算法模型，以求实现更加准确、高效的细胞自动计数。

丁媛媛[2]对传统的密度回归估计的方法进行了改进，通过直接学习预测计数图到注释图的映射，这种方法避免了使用高斯核生成真实密度图的传统步骤。这一创新有效减少了因高斯核选取不当带来的误差，提高了细胞计数的准确性。王治[3]使用了 RetinaNet (使用 Focal Loss 损失函数消除了类别不平衡) 算法对骨髓红系细胞进行目标检测，而四类红细胞相对白细胞而言小且密集，其提取了红细胞的 mask 图像并提出了细胞形状注意力模块来改善红系细胞细粒度分类效果。崔兆文[4]在原本的 YOLOv5 的基础上加入了注意力机制 SENet 网络和密集网络(DenseNet)来对血细胞进行检测和计数，同时提出一种凹点检测方法对黏连在一起的血细胞进行有效分割和计数，而对于种类较多的白细胞利用多分支残差网络进行进一步分类。林跃飞[5]提出了两种细胞计数方法，一种是基于最大类间方差法(OTSU)与标记控制分水岭分割相结合的方法，另一种是基于不均匀光照校正和分块 OTSU 阈值法结合的荧光显微细胞图像计数方法。设计并实现了荧光显微细胞图像分析系统。

Thanh Tran 等人[6]对血细胞计数时，首先进行语义分割，输入数据由原始图片改为了对红白细胞掩码后的图片，再对其进行分类和计数。Roberto Morelli 等人[7]将 ResNet 和 U-net 中的网络架构结合起来，解决了荧光显微镜细胞计数过度拥挤和伪影问题。Inayatul Haq 等人[8]采用将 YOLOv2 模型和 ResNet 残差网络相结合的方法解决图像中重叠拥挤的癌细胞计数问题，首先用较快速度的 YOLOv2 检测出细胞可能出现的位置，然后对检测出的区域使用 ResNet-18 和 ResNet-50 进行特征提取和分类，通过使用 ResNet 的残差连接，可以有效解决深度学习中常见的梯度消失问题，提高模型的训练效率和准确性。Leda Kamal 等人[9]提出了结合 K 均值算法和图像处理技术的增强型卷积神经网络(ECNN)模型，用于自动化血细胞分割和分类。通过中值滤波去噪和 K 均值分割技术，显著提高了图像质量和分类准确性。Najmeddine Dhieb 等人[10]出了一种基于 Mask R-CNN 和深度学习的自动化血细胞计数和分类框架，通过结合实例分割技术和 ResNet-101 特征提取网络，实现了对显微镜血涂片图像中淡化重叠的红细胞和白细胞的高精度识别、分类和计数。

3. 本文工作

目标检测等算法在血细胞统计任务上已有大量尝试，但先前诸多工作依旧存在不足之处：(1) 血液作为主体的红细胞数量较多容易发生粘连重叠，给计数造成了诸多不便。(2) 很多血细胞统计直接套用通用目标检测模型，缺少对血细胞数据处理的针对性，故训练和预测较慢。

针对以上两个问题，文本提出了一种新的模型——YOLO-MaskRCNN：(1) 在原来的 YOLOv5 目标检测模型中加入颜色注意力机制，让神经网络在计算过程中更加注重颜色特征，提高了模型识别的准确

度,且相比先前模型达到同等效果(mAP@0.5)耗时更短。(2) 利用 Mask R-CNN 模型单独对粘连重叠在一起的红细胞区域进行“分割”后计数,很好地解决了重叠红细胞无法计数的难题。以上也是本文的两个主要创新点。

为了验证此模型血细胞分类计数效果,本文在公开血细胞数据集(8种血细胞)上进行对照试验,并将该数据集训练好的模型整合为开源血细胞统计桌面应用。

4. YOLO-MaskRCNN 模型

4.1. YOLOv5

YOLO (You Only Look Once)算法最早是由华盛顿大学的 Joseph Redmon 在 2015 年提出[11],也深度学习目标检测领域中第一个单阶段算法,其推理速度极快。

在初代版本的推理预测阶段,图像将会被分为 $S \times S$ 个网格(grid cell),不同版本的 YOLO 算法 S 可能不同,每个网格将会预测出 B 个不同尺寸的预测框 Bounding boxes (的位置 x, y, w, h)以及每个 Bounding box 包含物体的概率 $P(\text{object})$,即置信度。与此同时每个网格将会预测出属于某个类别的条件概率 $P(\text{class}_i | \text{object})$ 。所以一张被划分为 $S \times S$ 个网格的图片,共需要预测出 $S \times S \times (B \times 5 + \text{class_count})$ 个值来完成定位和分类任务,也就是将目标检测任务转化为了深度学习中的回归任务。完成这一回归任务依靠复杂的卷积神经网络,YOLOv1 用了 GoogleNet, v2 版本用了 DarkNet19, v3, v4, v5 主要采用了 DarkNet53 及其改进版,通过这些卷积网络可以得到一个含有全部预测框位置和类别信息的张量,即 A 阶段,见图 1。

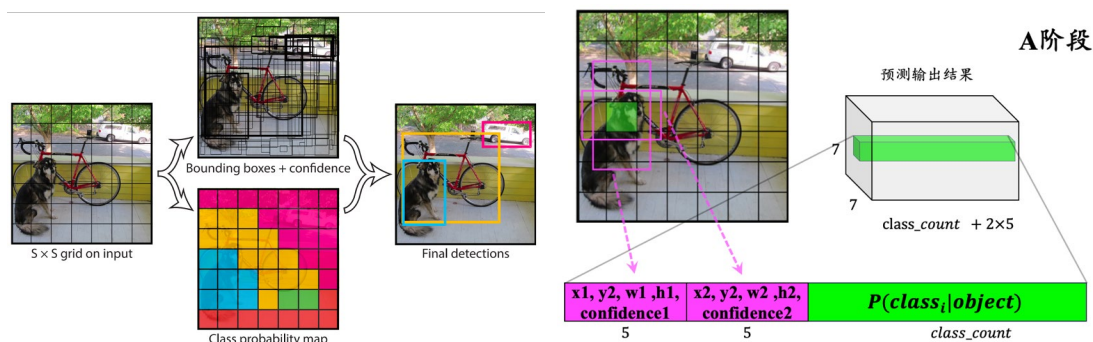


Figure 1. YOLO model target detection principle: predict target location and category information through one deep network

图 1. YOLO 模型目标检测原理: 通过一个深度网络预测目标位置和类别信息

初代 YOLO 基于每个网格直接预测边界框及其类别,但该机制存在一定局限:一方面,每个 grid cell 只能预测一个类别,若多个不同类别的目标中心落在同一 cell 内,则难以实现多目标检测;另一方面, Bounding boxes 预测范围过大,可能导致模型不稳定。为解决这些问题,从 YOLOv2 起至 YOLOv5,算法引入了 Anchor boxes (锚框)机制。Anchor boxes 是一组预定义的不同尺寸和长宽比的候选框,每个 grid cell 可以对应多个 Anchor box,每个 Anchor box 预测一个类别,从而有效缓解了重叠目标检测和类别冲突问题。同时,Anchor box 的偏移范围有限,不会在整幅图像上随意漂移,提高了检测稳定性。若设定 K 个 Anchor box,则卷积网络的输出张量为 $(5 + \text{class_count}) \times K$,其中包含边界框坐标、置信度以及类别信息。虽然增加 Anchor box 数量可以提升检测精度,但同时也会带来更高的计算开销。

Anchor box 的尺寸和比例既可以人工设定,也可以通过聚类等方法自动学习。YOLOv5 在此基础上引入了自适应锚框(Adaptive Anchors)机制:在训练阶段,通过 k 均值聚类对数据集中目标的尺寸与比例

进行分析, 依据 IoU 作为匹配度指标迭代优化, 生成与数据分布高度匹配的最优锚框集合。这种自适应方式减少了人工调整的负担, 使模型能够根据不同数据集自动优化锚框参数, 显著增强了泛化能力和检测精度。

另外, YOLOv5 还首次采用了 **letterbox** 自适应缩放技术。早期版本通常将输入图像强制缩放到统一尺寸, 导致目标形态发生畸变或信息丢失。而 **letterbox** 缩放通过保持原图纵横比进行等比例缩放, 并在空余区域填充灰色像素补齐至指定尺寸, 既保证了网络输入的一致性, 又最大限度保留了图像的有效信息, 从而进一步提升了检测的准确性和稳定性, 其网络结构见图 2。

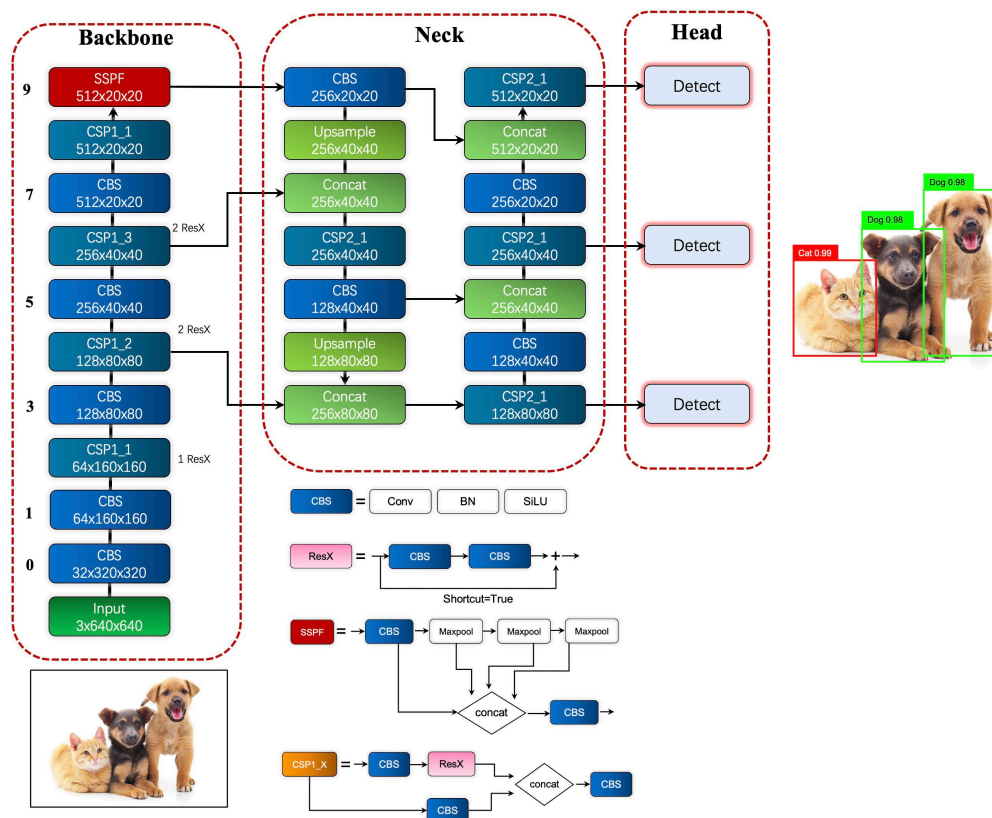


Figure 2. YOLOv5 internal network structure

图 2. YOLOv5 模型内部网络结构

4.2. Mask R-CNN

另 Mask R-CNN 是一个二阶段的目标检测算法并且还能完成实例分割任务, 由 Kaiming He 等人[12]提出, 其在 COCO 数据集实例分割任务上取得了很好的效果, 领先于其他实例分割模型。其主要的在先前的 Faster R-CNN 的基础上加入了用于实例分割的分支网络预测掩码, 该分支和现有的分类和边框回归并行, 添加了该分支仅仅增加了很小的运算负担。该分支将区域建议网络得到的 ROI align 输入一个小型的全卷积网络 FPN 中, 输出该区域图片的二进制掩码, 且对于每一个类别独立预测其掩码, 使用的是 Sigmoid 和二元交叉熵损失函数(Binary Cross-Entropy Loss), 这样避免了多个类别彼此竞争。最后实例具体是哪一个类别由分类分支的结果决定, 此时输出该预测类别的掩码。而对于每一个实例的类别和对应预测框的预测一般通过 ResNet 或 ResNeXt 网络进行预测。所以 Mask R-CNN 共输出三个结果: 预测类别、预测框和掩码, 三部分组合得到我们想要的结果, 也就是带有掩码和类别的图片。Mask R-CNN 的优势在于像

素级别的实例分割和精细化特征对齐，这使它在处理粘连、重叠的目标时更准确、更鲁棒，见图 3。

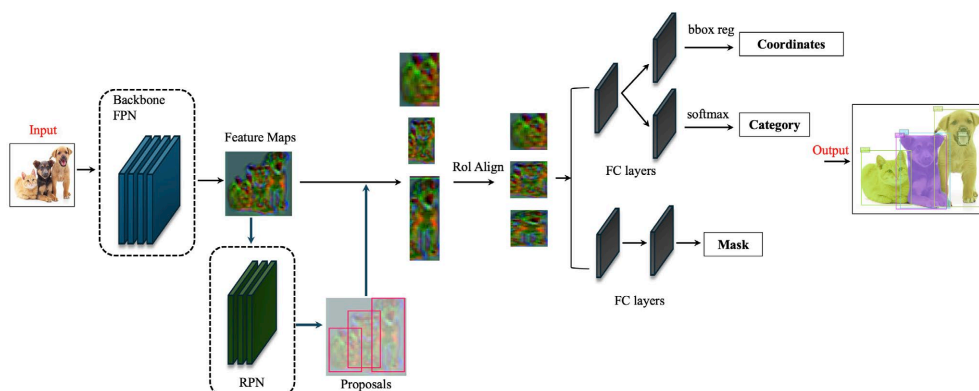


Figure 3. Mask R-CNN model internal structure and operation process

图 3. Mask R-CNN 模型内部结构以及运作流程

4.3. 颜色注意力机制

注意力机制模仿人类视觉和认知系统，使模型在处理输入数据时聚焦于相关部分。通过注意力机制，神经网络可以选择性关注重要信息，提高模型的泛化能力和训练收敛速度。在图像分类任务中，注意力机制可以使模型集中于对分类决策最重要的图像区域。在目标检测任务当中，加入注意力机制可以提高模型对小目标或遮挡重叠目标的识别能力。通过聚焦于图片中的特定区域，模型可以更精确地定位和识别各种物体。

通常在进行血细胞计数前，都会人为对血细胞进行染色，以便更好的用显微镜观察。而不同种类的血细胞在染色后会呈现不同的颜色和纹理特征，比如白细胞的颜色相对红细胞更深，红细胞由于其双凹盘结构中心颜色较浅接近于背景色，相比让神经网络盲目通过训练来寻找特征，我们可以让模型在训练过程中主动更加关注血细胞的颜色特征，使其更快收敛，降低了计算成本，也可以提高血细胞识别分类精度。为了实现让模型提高对颜色特征的关注度，本文提出了颜色注意力机制，并将该模块其加入到 YOLOv5 的主干网络中，其内部结构如图 4：

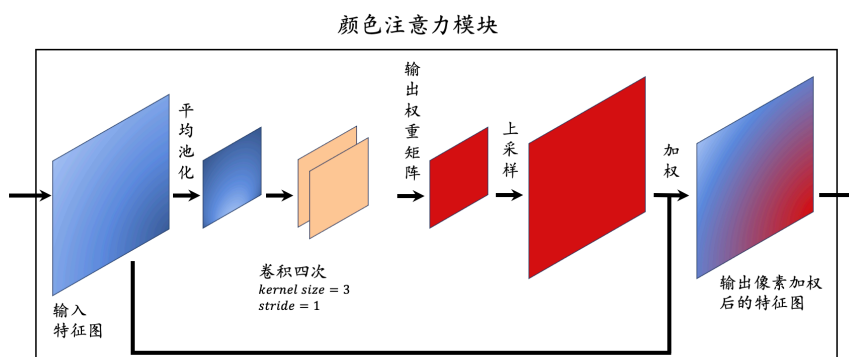


Figure 4. Internal structure of color attention module

图 4. 颜色注意力模块内部结构

该模块的输入为卷积网络产生的特征图，模块内先经过平均池化将输入特征图尺寸压缩 1/2，以降低计算负担，后经过两次卷积运算，输出权重矩阵，而权重矩阵中的每一个权重通过反向传播求得。然后

通过上采样将权重矩阵尺寸恢复为原始输入尺寸,此时特征图的每一个像素颜色都得到了一个对应权重,最后将其加权(Hadamard 积)到输入特征图中输出。

4.4. YOLO-MaskRCNN 模型结构

YOLO 单阶段目标检测算法速度快,但对于复杂图像精细化分类能力不足;Mask R-CNN 虽算法预测速度慢但精细化程度高。所以将两种算法结合取长补短综合为新的模型。对血细胞一般区域通过改进 YOLO 算法进行检测,而对于其中的粘连红细胞区域单独采用 Mask R-CNN 进行实例分割精细化计数。具体方法实现是通过将粘连红细胞单独作为一个类 RBC2 别加入到 YOLO 模型训练中,预测阶段 YOLO 算法首先将粘连区域识别出来,然后根据边界框在图片中的坐标将该区域图像单独裁剪出来作为 Mask R-CNN 的输入,而对于 YOLOv5 的改进是将颜色注意力模块加入到主干网络第 8 层 CSP 模块之后,其目的在于对卷积之后的特征图进行逐个像素地加权。模型流程见图 5,模型全过程通过 Python 实现。

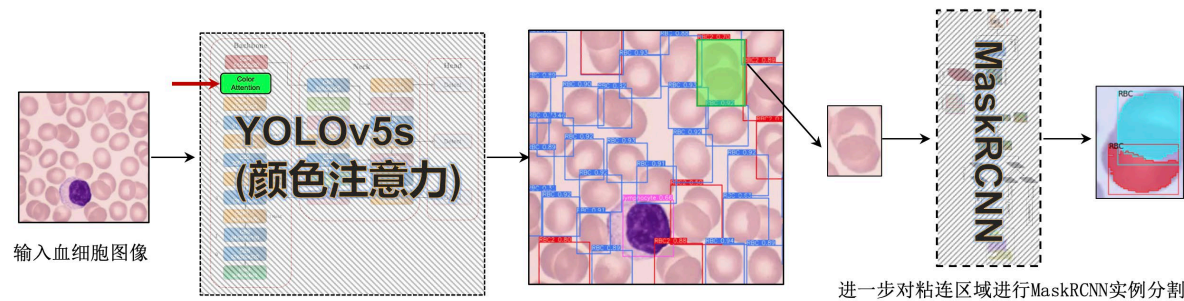


Figure 5. YOLO-MaskRCNN model basic structure
图 5. YOLO-MaskRCNN 模型基本结构

5. 实验数据以及预处理

实验所用数据集共包含 621 张正常细胞图像,这些图像由巴塞罗那医院诊所核心实验室的 CellaVision DM96 分析仪采集。数据集分为以下八组:中性粒细胞、嗜酸性粒细胞、嗜碱性粒细胞、淋巴细胞、单核细胞、未成熟粒细胞(早幼粒细胞、中幼粒细胞和晚幼粒细胞)、有核红细胞以及血小板或血栓细胞。图像尺寸为 360×363 像素, JPG 格式,并由专业临床病理学家进行注释。这些图像采集自未感染、血液病或肿瘤疾病且在采血时未接受任何药物治疗的个体,图 6 为数据集部分图像。

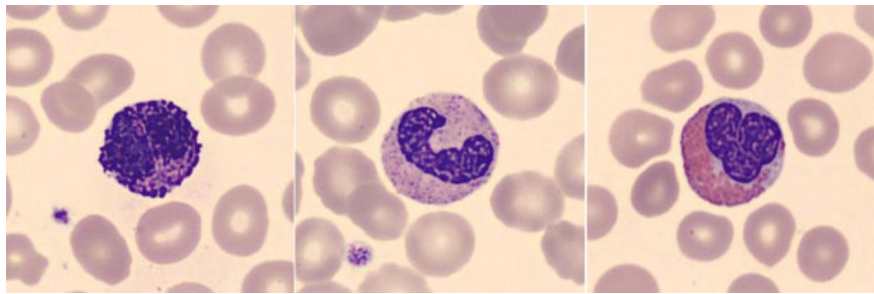


Figure 6. Some images in this article's dataset, including three types of white blood cells, blood cells, and platelets
图 6. 本文数据集中的部分图像, 其中包含 3 种白细胞、血细胞和血小板

为了后续训练 YOLO 模型和 Mask R-CNN 模型,用 labellmg 工具还有 labelme 工具事先进行目标标注,其中将粘连区域的红细胞单独标注为一个类别。

6. 模型训练及实验结果

实验使用的操作系统为 macOS 11.5.2, 处理器为 1.4 GHz 四核 Intel Core i5, 显卡为 Intel Iris Plus Graphics 645, 项目环境 Python 3.8 (conda), 使用的编译器为 Pycharm CE 2021。

6.1. 模型 YOLOv5 部分的训练

(1) YOLO 模型加入颜色注意力机制前后对比:

加入前后进行两组实验, 均使用 621 张标注后的细胞图像, 训练 200 轮次, batch size 为 8, 训练占比 80%。以 $mAP@0.5$ ¹ 作为模型指标, 实验结果如图 7。每一种血细胞的平均精度 AP 见表 1。

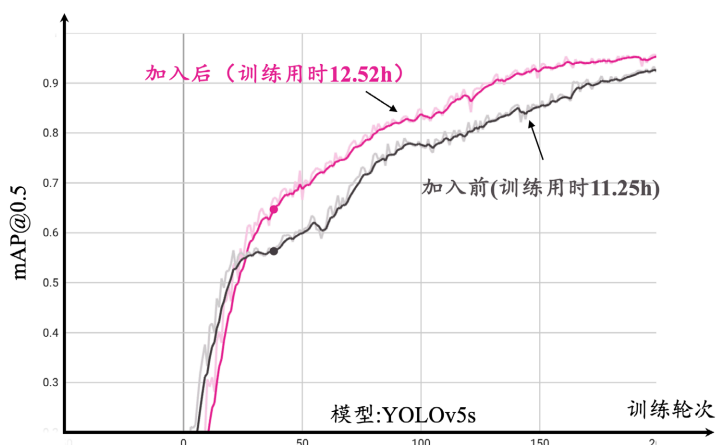


Figure 7. Comparison of $mAP@0.5$ before and after adding color attention mechanism to YOLOv5 model
图 7. YOLOv5 模型加入颜色注意力机制前后 $mAP@0.5$ 对比

Table 1. AP of the two models for identifying different blood cells and $mAP@0.5$ of the overall model

表 1. 两种模型识不同种类血细胞的 AP 以及模型总体的 $mAP@0.5$

不同类别血细胞 AP											
模 型	BA	RBC	NE	PL	EO	IG	LY	MO	RBC2	耗时	$mAP@0.5$
YOLOv5s	0.987	0.975	0.978	0.760	0.995	0.995	0.974	0.986	0.765	11.25h	0.935
YOLOv5s + 颜色注意力	0.990	0.977	0.984	0.879	0.986	0.996	0.971	0.973	0.800	12.52h	0.951

通过两次实验发现, 在同样的实验环境下, 在 YOLOv5s 模型中加入颜色注意力机制后, 模型在训练的大部分轮次中 $mAP@0.5$ 均高于原始模型。同时观察模型对每一类血细胞的识别精度, 可以发现加入颜色注意力后模型对血小板(PL)和粘连红细胞(RBC2)识别效果有明显提升, 这说明颜色注意力加入后通过对颜色特征的关注来帮助模型识别小目标以及重叠目标难识别的问题。

(2) 加入颜色注意力机制同加入其他注意力机制对比:

在很多文章中对 YOLO 算法的目标检测模型改进加入了 SENet 通道注意力机制(Squeeze-and-Excitation Networks) [13], 通过引入压缩激励模块, 自适应地重新加权每个通道的特征, 提高网络的表示能力和分类性能。具体来说, SENet 通过全局平均池化压缩每个通道的空间信息, 然后通过全连接层和 Sigmoid 激活函数生成每个通道的权重, 这些权重用于调整原始特征图, 使得网络更关注有用的特征。本

¹ $mAP@0.5$ 是目标检测中最常用的性能指标, 表示当预测框与真实框的重叠度(IoU) 0.5 时, 计算每个类别的平均精度(AP), 再对所有类别取平均值, 反映模型在“检测到目标”层面的整体准确性。

文也将其加入实验进行对比。实验分为三组，对照组为普通的YOLOv5s，实验组一为YOLOv5s+SENet(通道注意力)，实验组二为YOLOv5s+颜色注意力模块，实验组三(对照组)为原始YOLOv5s模型。训练200轮后，图8为实验结果。

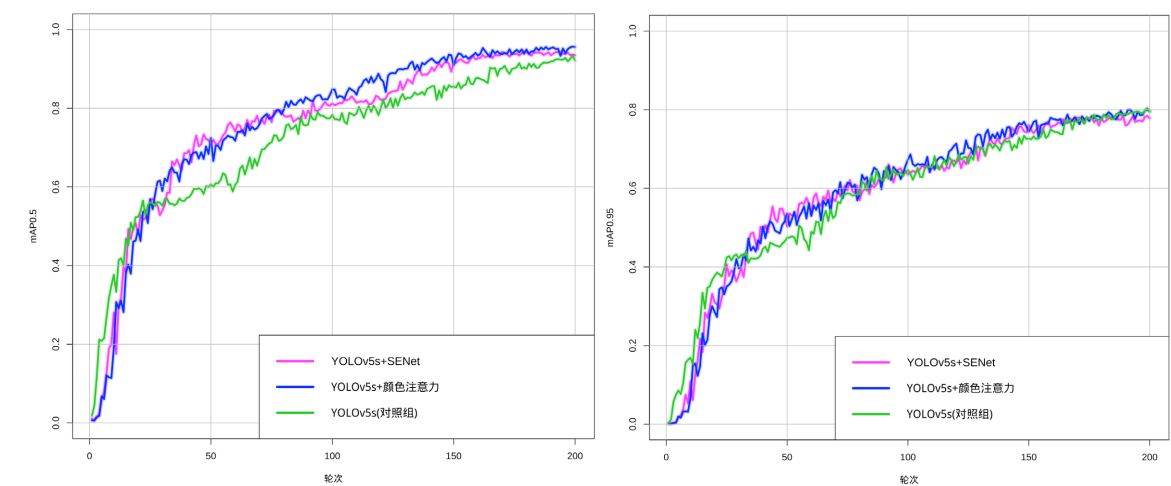
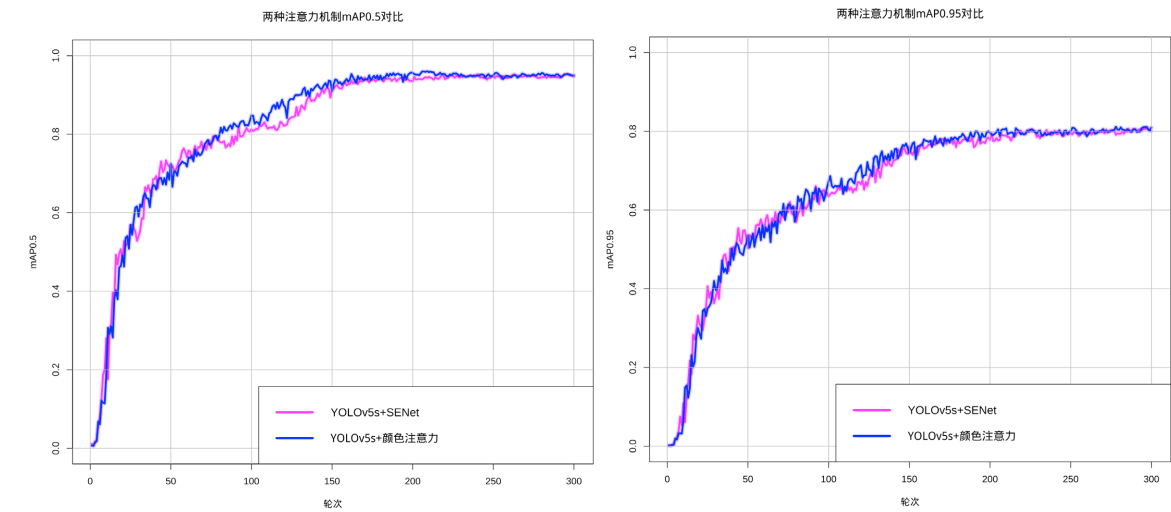


Figure 8. Comparison of mAP@0.5 and mAP@0.95 of three groups of experiments during training
图8. 训练过程中三组实验 mAP@0.5 以及 mAP@0.95 对比

200轮训练数据可以看出，无论使用SENet通道注意力机制还是用本文提出的颜色注意力机制来改进YOLOv5模型，改进后的mAP0.5和Precision均明显大于原始模型，说明加入注意力机制确实可以提高模型整体性能。为了更好地对比两种注意力机制，将加入两种注意力机制的YOLOv5s各训练300轮，训练结果如表2和图9。

Table 2. Preliminary comparison of the two attention mechanisms
表2. 两种注意力机制初步对比

模型	耗时	mAP@0.5
YOLOv5s + SENet	25.13h	0.947
YOLOv5s + 颜色注意力	18.26h	0.950



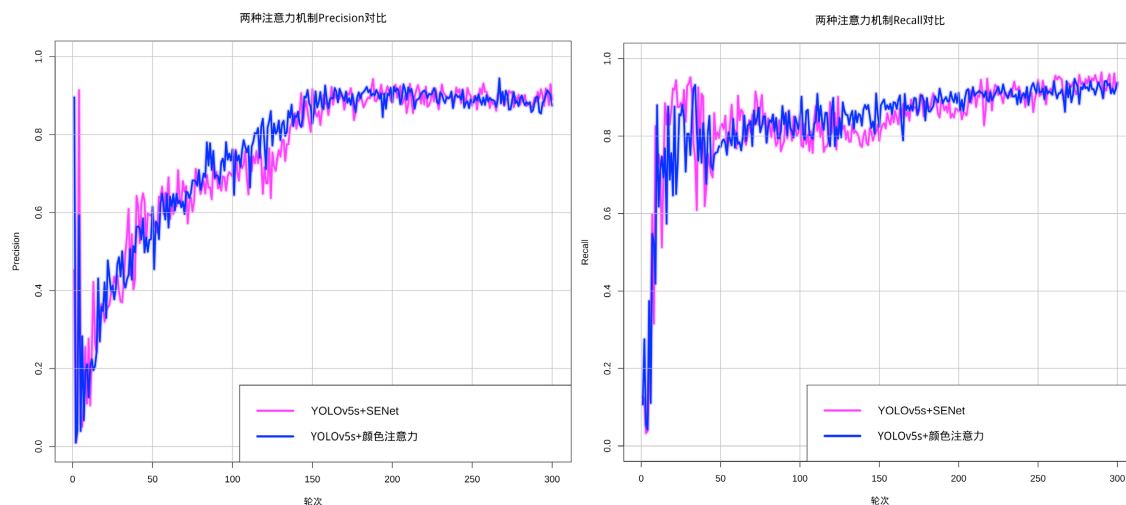


Figure 9. Comparison of mAP@0.5, mAP@0.95, Precision and Recall metrics of two attention mechanisms
图 9. 两种注意力机制 mAP@0.5, mAP@0.95, Precision 和 Recall 指标对比

利用 TensorBoard 获得模型损失、mAP 和 Precision 等实时训练数据，无论是以 0.5 还是以 0.95 为阈值计算模型的 mAP，大多数情况下加入了颜色注意力机制的 YOLOv5s 模型要略微高于加入了通道注意力机制 SENet 的 YOLOv5s。但是颜色注意力机制要相比通道注意力机制训练耗时更短，在大规模数据集上前者消耗的计算资源相对更少。前者在测试集(125 张 $1024 \times 1024 \times 3$ 的血细胞图像)上的推理检测速度为 113.216 s，后者为 118.599 s，颜色注意力机制略快。训练好的带有颜色注意力机制的 YOLO 模型不仅在测试集上表现良好，在其他数据上的泛化效果也不错，见图 10：

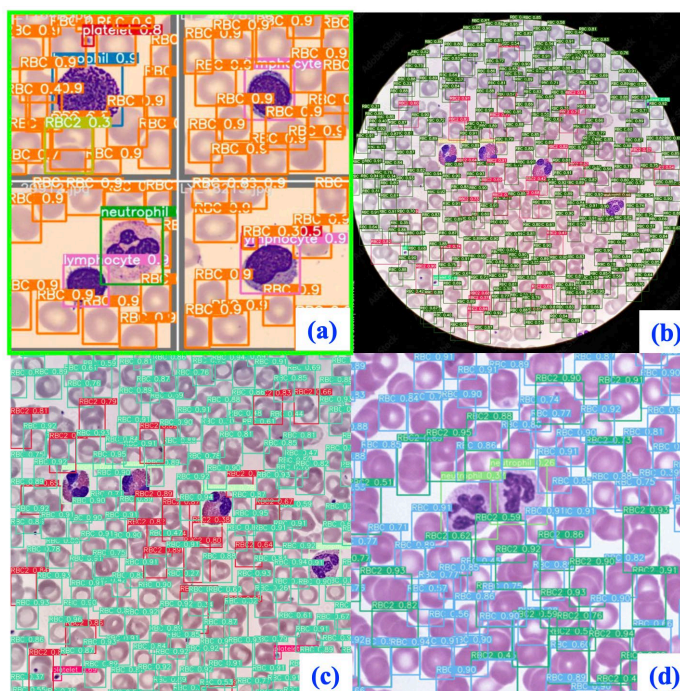


Figure 10. YOLOv5 target detection effect with color attention mechanism: images from the test set (a) and from other datasets (b, c, d)

图 10. 带颜色注意力机制的 YOLOv5 目标检测效果：图像来自测试集(a)和来自其他数据集(b, c, d)

6.2. 模型 MaskRCNN 部分的训练

从原始的 621 张数据集中提取 180 张带有粘连红细胞的图片作为数据集。训练占 80%，batch size 为 8，100 轮次。用训练好的模型对测试集中的重叠红细胞进行实例分割，取得了较好的效果，效果如图 11：

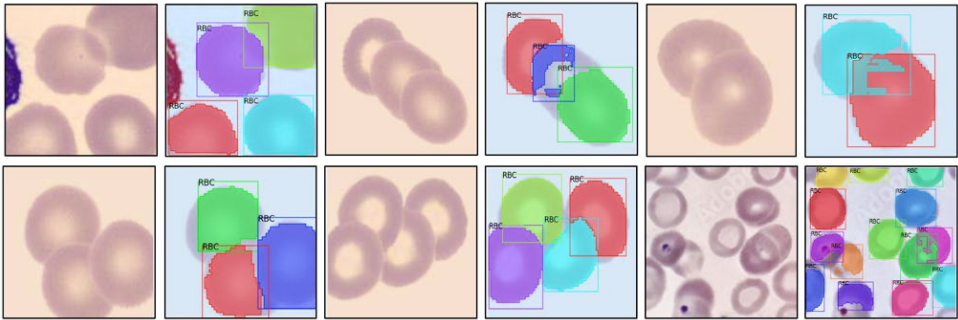


Figure 11. The trained MaskRCNN segmentation effect on adhesion blood cell instances
图 11. 训练好的 MaskRCNN 对粘连血细胞实例分割效果

7. 开源血细胞计数程序

本文主要通过改进后的 YOLOv5 模型结合 Mask R-CNN 模型实现外周血细胞的计数任务，但是由于上文用到的两个算法是分开的两个 python 项目，实际操作起来较繁琐，且容易出错，所以编写一个桌面应用将两个算法整合起来，让细胞计数可以实现流水线式操作。

桌面程序开发环境为 Python 3.6，编译器为 Pycharm CE 2021，桌面程序开发工具为 tkinter。tkinter 是 Python 的标准 GUI 库，它提供了一个简单易用的接口来创建图形用户界面。tkinter 基于跨平台的工具包，支持在 Windows、macOS 和 Linux 上运行。通过 tkinter，开发者可以使用丰富的控件，如按钮、标签、文本框、菜单等，构建事件驱动的桌面应用程序。由于其内置于 Python 标准库中，无需额外安装，所以本文采用 tkinter 开发桌面程序[14]。该程序主要内容和使用方法如下：

(1) 图像输入：在程序主窗口的蓝色路径框中输入所要处理的血细胞图片绝对路径，点击【确认】可以在下方显示该图片，见图 12。

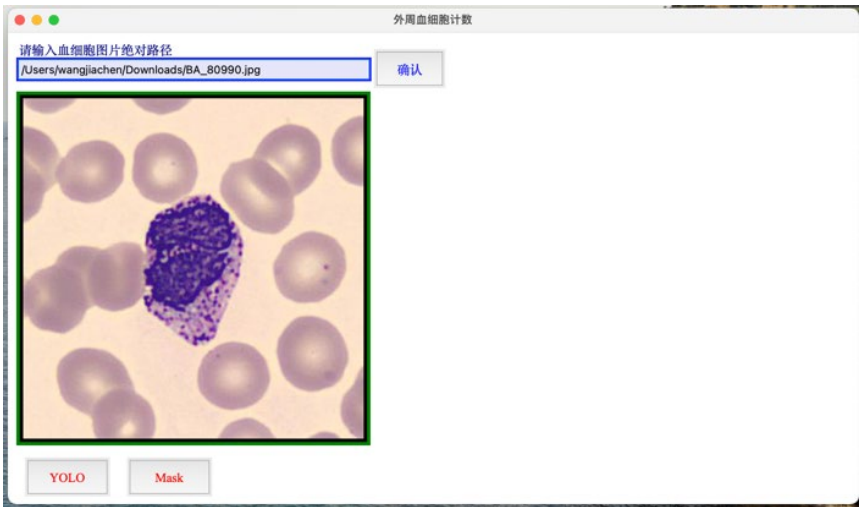


Figure 12. Enter an image of blood cells into the program
图 12. 向程序中输入血细胞的图像

(3) 粘连细胞实例分割：将训练好的 Mask R-CNN 权重以及相关文件嵌入，点击【Mask】按钮，对之前检测到的每一个 RBC2 区域进行实例分割(获取粘连区域 bounding boxes 的坐标，并保存到 txt 中，然后将粘连区域单独从图片中取出并保存)，并且可以查看每一个区域的实例分割带掩码图片，比如下图 4 个按钮对应 4 个重叠的红细胞，点击【0】即可查看第 1 个重叠红细胞实例分割后的图片，如图 14。

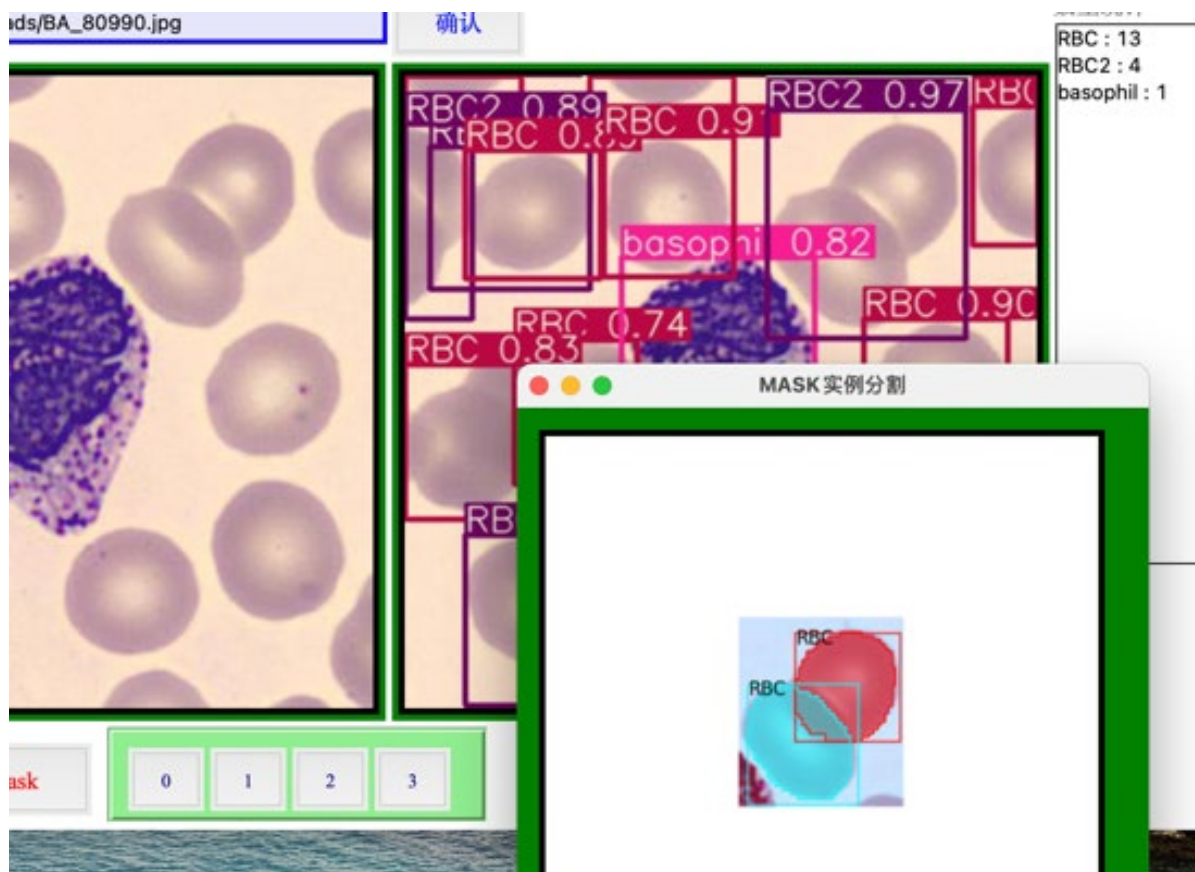


Figure 14. The program segments and counts the red blood cells adhered to the image
图 14. 程序对图像中粘连的红细胞进行分割计数

该桌面应用可以在不同操作系统下运行，程序全部代码以及两个模型文件已上传到 github 中，
<https://github.com/JiachenWong/Blood-Cells-Count-YOLO-MaskRCNN>

8. 总结和展望

本文主要内容是使用带有颜色注意力机制的 YOLOv5 模型对血细胞进行目标检测，同时对于粘连重叠的红细胞，单独对其所在区域进行实例分割。因为 Mask R-CNN 是多阶段的目标检测算法，其首先要利用 RPN 网络生成图像中感兴趣的区域，然后再利用卷积神经网络进行特征提取，最后还要生成掩码，其速度非常慢，所以主要的血细胞计数和检测任务交给 YOLO 算法完成。

深度学习中的注意力机制是为了让模型能够有针对性地去学习，具体实现方式是通过在神经网络中赋予不同部分不一样的权重，来体现不同区域的重要性。不同血细胞存在颜色上的差异，所以应该关注卷积中特征图的颜色特征。颜色的最小单位是像素，每个像素是一个单独的颜色，但是如果按照像素去

加权计算负担太大, 所以进行池化操作, 得到一个个较大的色块, 然后用全卷积网络对色块加权, 再上采样对每个像素赋予不同权重。

重叠粘连细胞的识别和分割一直都是目标检测中的难点, 很多解决方法是无监督学习的方法比如 Canny 边缘检测或分水岭算法, 这些方法严重依赖阈值和随机数的设置, 计数结果相当不稳定, 算法鲁棒性太差。最关键的问题在于此类方法对细胞形态改变较大(使用腐蚀膨胀操作), 处理后的图片不利于进行细胞形态学检查。而实例分割方法计数准确度高, 粘连细胞分割准确且不会对形态观察产生较大影响。

本文依旧存在诸多问题还没有很好的解决:

(1) 颜色注意力模块的全卷积网络相对简单, 仅使用了 2 个卷积层来用于权重生成。过少的卷积层可能对复杂的血细胞图像的特征提取力不从心, 得到的权重效果不显著, 但是过多的卷积层也可能造成过拟合, 所以依旧需要进行反复试验确定最佳数目。

(2) 本文在 YOLOv5 的网络 neck 层中加入了一个颜色注意力模块, 而具体应该加入几个该模块, 加到什么位置可以发挥最好效果依旧有待反复试验考证。

(3) 对于粘连的红细胞采用 Mask R-CNN 实例分割算法, 该方法虽然常用, 但推理速度太慢, 特别是当图片中存在较多的粘连红细胞时, 将耗费大量的时间。主要是因为该方法主要是用来进行目标检测, 而本文仅用来实例分割。所以应该采用专门进行实例分割的算法代替 Mask R-CNN。

(4) Mask R-CNN 实例分割的准确与否很大程度上取决于样本的数量和质量。因为重叠的红细胞形态各异, 有两两串联形, 也有三五个成团聚集形, 要想将它们精准分开需要相当多的重叠红细胞掩码样本。但是本文仅用了 100 多个样本, 所以模型对于较复杂的粘连细胞束手无策。为了解决样本不足的问题, 可以将多种形态单独的红细胞从已有图片中抠出, 并消除其背景, 然后随机生成重叠的红细胞图片扩充实例分割样本。

参考文献

- [1] 武银刚. VWF 加重小鼠 TBI 后微囊泡介导的血管通透性改变及凝血功能障碍的机制研究[D]: [博士学位论文]. 天津: 天津医科大学, 2018.
- [2] 丁媛媛. 基于密度回归估计的细胞计数方法研究[D]: [硕士学位论文]. 济南: 山东师范大学, 2024.
- [3] 王治. 基于卷积神经网络的红系细胞检测和分类方法研究[D]: [硕士学位论文]. 长春: 吉林大学, 2023.
- [4] 崔兆文. 基于深度学习的血液细胞计数的研究[D]: [硕士学位论文]. 贵阳: 贵州大学, 2022.
- [5] 林跃飞. 基于显微图像的细胞计数方法研究及分析系统的构建[D]. 厦门: 华侨大学, 2024.
- [6] Tran, T., Lam, B.M., Lee, S.H. and Kwon, K.R. (2019) Blood Cell Count Using Deep Learning Semantic Segmentation.
- [7] Morelli, R., Clissa, L., Amici, R., Cerri, M., Hitrec, T., Luppi, M., *et al.* (2021) Automating Cell Counting in Fluorescent Microscopy through Deep Learning with C-ResUnet. *Scientific Reports*, **11**, Article No. 22920. <https://doi.org/10.1038/s41598-021-01929-5>
- [8] Haq, I., Mazhar, T., Asif, R.N., Ghadi, Y.Y., Ullah, N., Khan, M.A., *et al.* (2024) YOLO and Residual Network for Colorectal Cancer Cell Detection and Counting. *Heliyon*, **10**, e24403. <https://doi.org/10.1016/j.heliyon.2024.e24403>
- [9] Kamal, L. and Raj, R.J.R. (2024) Harnessing Deep Learning for Blood Quality Assurance through Complete Blood Cell Count Detection. *e-Prime—Advances in Electrical Engineering, Electronics and Energy*, **7**, Article 100450. <https://doi.org/10.1016/j.prime.2024.100450>
- [10] Dhieb, N., Ghazzai, H., Besbes, H. and Massoud, Y. (2019) An Automated Blood Cells Counting and Classification Framework Using Mask R-CNN Deep Learning Model. *2019 31st International Conference on Microelectronics (ICM)*, Cairo, 15-18 December 2019, 300-303. <https://doi.org/10.1109/icm48031.2019.9021862>
- [11] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788.

<https://doi.org/10.1109/cvpr.2016.91>

- [12] He, K., Gkioxari, G., Dollar, P. and Girshick, R. (2017) Mask R-CNN. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2980-2988. <https://doi.org/10.1109/iccv.2017.322>
- [13] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>
- [14] Janssen, M.E.L. (2024) Initial Development of a Python-Based Graphical User Interface for a Medical Measurement System for Nociceptive (Mal)Function. University of Twente.