

基于轻量化YoloV5s的血癌细胞检测

乔冬*, 何利文

南京邮电大学物联网学院, 江苏 南京

收稿日期: 2024年3月5日; 录用日期: 2024年4月19日; 发布日期: 2024年4月29日

摘要

为了解决人工检测血癌细胞耗费人力, 且容易出现漏检、误检的情况, 提出一种新型的检测血癌细胞检测算法——YoloV5s-FasterNet。该算法在原本的YoloV5s框架上, 通过将主干网络中的C3检测层改进为FasterNet轻量化网络用来提取特征信息, 降低整体网络模型的复杂程度, 并在主干网络中增加坐标注意力(coordinate attention, CA)机制模型更好的定位和识别目标信息来提高检测精度。通过在血癌数据集上进行大量实验结果表明: 与原模型相比, 改进的模型平均精度提升2.2%, 参数量减少25.7%, 检测速度提高15%, 实验结果良好, 证明该算法对血癌细胞检测具有实用性。

关键词

深度学习, 血癌细胞检测, YOLOV5s, 轻量化网络, 注意力机制

Detection of Blood Cancer Cells Based on Lightweight YoloV5s

Dong Qiao*, Liwen He

School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu

Received: Mar. 5th, 2024; accepted: Apr. 19th, 2024; published: Apr. 29th, 2024

Abstract

In order to solve the problem that manual detection of blood cancer cells is labor-intensive, and it is easy to miss and misdetect, a new detection algorithm for blood cancer cells is proposed- YoloV5s-FasterNet. Based on the original YoloV5s framework, this algorithm improves the C3 detection layer in the backbone network into a FasterNet lightweight network to extract feature information, reduce the complexity of the overall network model, and add coordinate attention to the backbone network. This model can better locate and identify target information to improve detec-

*通讯作者。

tion accuracy. A large number of experimental results on the blood cancer data set show that: compared with the original model, the average accuracy of the improved model is increased by 1.2%, the number of parameters is reduced by 25.7%, and the detection speed is increased by 15%. The experimental results are good, proving that the algorithm is effective in detecting blood cancer cells.

Keywords

Deep Learning, Blood Cancer Cell Detection, YOLOV5s, Lightweight Network, Attention Mechanism

Copyright © 2024 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

传统的血癌细胞识别方法通常依赖于人工检查和显微镜观察。但这种方法存在操作复杂、耗时长、主观性强等问题[1] [2]。而基于图像处理和机器学习的血癌细胞识别技术则能够提高诊断的准确性和效率,帮助医生更早地发现和治理血癌,提高患者的存活率和生活质量。此外,随着计算机硬件的不断进步和深度学习算法的优化,血癌细胞识别技术将在未来持续发展,并为临床医学带来更多的突破和创新[3] [4] [5]。

利用深度学习技术,可以对目标进行较为准确的检测,在医疗影像领域中可以加以应用。主流的目标检测法大致分为两类:一类是两阶段检测,比如 R-CNN 算法[6],该算法在第 1 阶段生成候选对象区域放入深度网络中提取特征,在第 2 阶段将候选区域进行分类,这样的两阶段检测虽然运行速度缓慢但精度很高。另一类是单阶段检测:比如 YOLO 系列[7] [8]和 SSD 算法[9],其相较两阶段检测的检测速度有所提升,准确率也在逐渐提高。

综上,本文提出了 YOLOv5s-FasterCA 算法来进行血癌细胞检测,该算法在 YoloV5s 的基础上,将主干网络中的 C3 检测层替换为效率更高的 C3_FasterNet,在不降低精度的同时提高检测速度,并在主干网络中添加坐标注意力机制[10]来增强模型提取特征的能力。经过实验,本文的模型对血癌细胞的检测效果较原模型显著提高,应用于血癌细胞检测的效果很好。

2. YoloV5-7.0 算法

YOLOv5 算法在目标检测中应用相对较多,包括了多个版本,如 1.0、6.0 等,同时网络规模也有很多类型。在 YOLOv5-7.0 算法当中,其整体布局和 v3、v4 较为类似,和 YOLOv5 也不会有很大的差距。在算法的结构组成上,包括了输入、分类器、颈部及主干网等几个主要组成部分,具体的整体结构如图 1 所示。

2.1. 输入端

在输入端当中,输入图片的尺寸会有一些的差异,需要采取缩放的方法将尺寸调整到一个合适的大小,之后再利用算法对其进行填充处理,并在网络当中进行训练。由于图像的尺寸会存在着分布不够均匀的问题,故需要对图像进行训练和增强,这样就能够解决这一问题。在对图像进行增强的过程中,可以对图像的亮度进行随机调整,同时也可以对图像进行翻转,并且提高网络鲁棒性。

2.2. 主干网络

在主干网络当中, 需要采取到卷积结构, 包括了卷积核、步长、填充等几个主要的功能模块。和传统的 Focus 模块相比, 采取卷积层能够取得更高的效率, SPPF 的结构如图 1 所示。

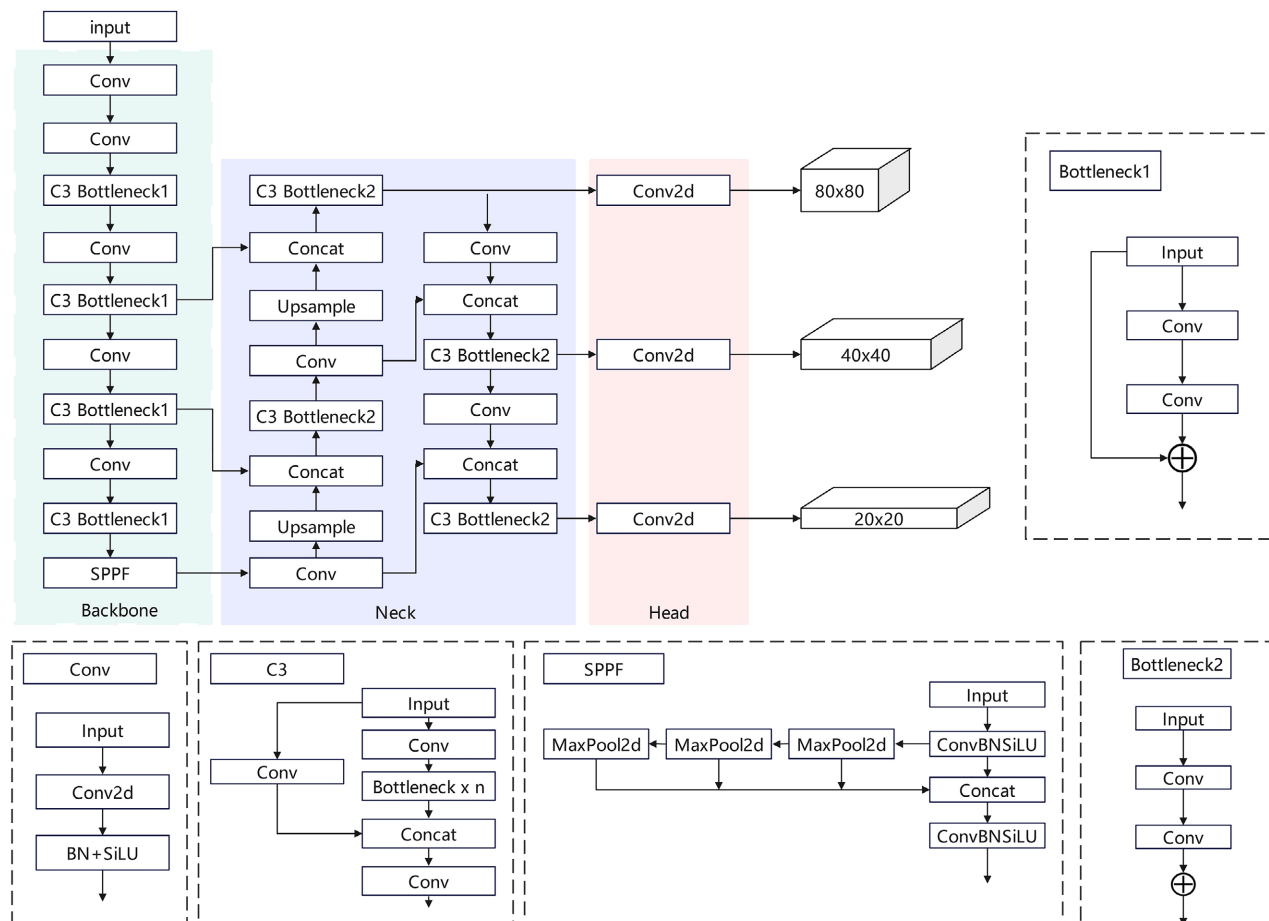


Figure 1. YoloV5-7.0 network structure
图 1. YoloV5-7.0 网络结构

2.3. 颈部网络

YOLOv5-7.0 的 Neck 部分采用 FPN + PAN 的结构形式, 其中 PAN 代表路径聚合网络, FPN 代表特征金字塔网络。通过这种结构形式, 能够取得较好的应用效果。

2.4. 预测分类器

YOLOv5 的 head 用 3 个 1×1 的卷积层替代全连接层(FC)进行预测和分类, 分别生成 20×20 , 40×40 , 80×80 这 3 个不同尺度的特征图, 输出结果为 $4 + 1 + C$, 分别为预测框坐标、目标置信度、类别概率值。

3. 改进的 YOLOv5s 算法

本实验基于 YOLOv5-7.0 算法提出了改进网络 YOLOv5s-FasterCA。将原 YOLOv5s 模型中主干网络的 C3 检测层替换为 FasterNet, 该网络通过将常规卷积替换为部分卷积(Partial Convolution, PConv)与逐点卷积(Pointwise Conv, PWConv)来减少计算冗余[11], 大大降低了模型复杂度和参数量来提高计算速度,

同时, 在主干网络中引入 CA 模块来提高模型对血癌细胞中三类细胞的特征提取能力, 改进后的 YOLOv5s-FasterCA 模型如图 2 所示。

3.1. 部分卷积与逐点卷积

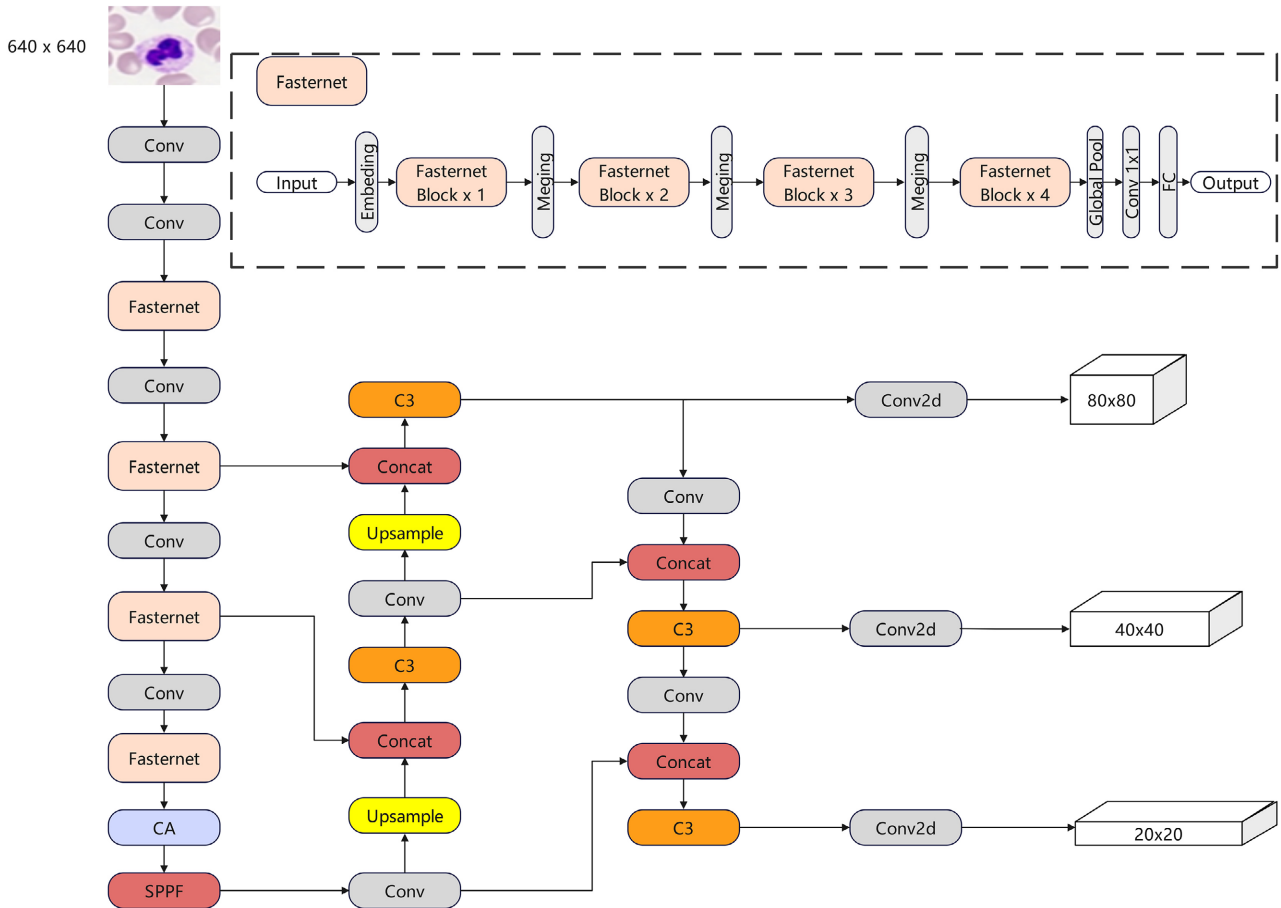


Figure 2. Improved YOLOv5s network structure
图 2. 改进后的 YOLOv5s 网络结构

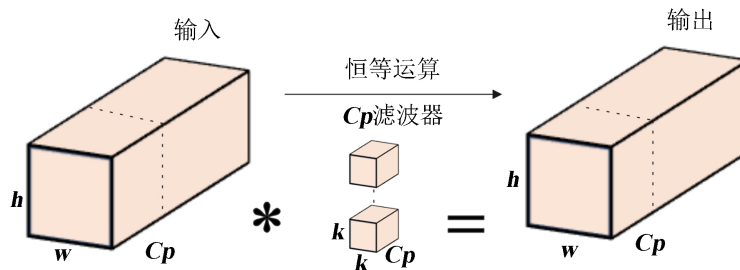


Figure 3. Partial convolution structure
图 3. 部分卷积结构

在原 YOLOv5s 模型中, 采用了 C3 模块对输入图片进行卷积操作并提取特征信息, C3 模块是由三个 Conv 块构成的, 其中第一个 Conv 块的步幅为 2, 可以将特征图的尺寸减半, 第二个 Conv 块和第三个 Conv 块的步幅为 1。C3 模块中的 Conv 块采用的都是 3×3 的卷积核。在每个 Conv 块之间, 还加入

了 BN 层和 LeakyReLU 激活函数, 以提高模型的稳定性和泛化性能。但是此模块在对特征较为单一的输入图片进行特征提取时会产生大量计算冗余, 因此导致模型的计算速度降低。因此本文的改进算法借鉴的部分卷积与逐点卷积的思想来提高模型的计算速度。

部分卷积只对部分通道进行特征提取, 对于其他的通道, 则保持不变, 其结构如图 3 所示。在对内存访问中, 计算代表可以选择第 1 个或者是最后 1 个 c_p 通道, 这样计算效率也会相对较高, 输入及输出的通道数量可以保持一致。因此, 部分卷积的 FLOPs 仅有:

$$h \times w \times k^2 \times c_p^2 \quad (1)$$

其中 h 为输入图片的高, w 为宽, k 为卷积核大小。一般情况下, 部分卷积的通道数 c_p 为常规卷积 c 的 1/4, 所以部分卷积的 FLOPs 仅为常规卷积的 1/16。此外, 部分卷积的内存访问量更小, 仅为常规卷积的 1/4, 为:

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (2)$$

通道当中的信息需要得到充分的利用, 并且需要对每一个点进行卷积计算。它们看起来像一个 T 形卷积, 如图 4 所示, 与均匀处理一个点的常规卷积相比, 它更关注中心位置。

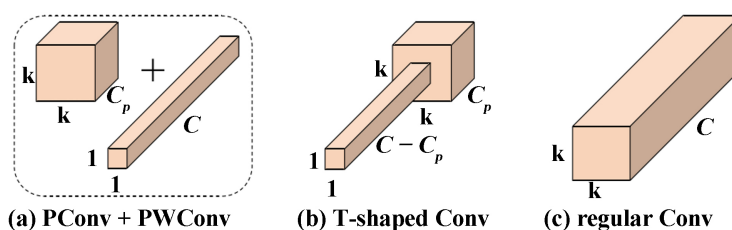


Figure 4. T-shaped convolution structure
图 4. T 形卷积结构

虽然 T 形 Conv 可以直接用于高效计算, 但结果表明, 将 T 形 Conv 分解为 PConv 和 PWConv 更好, 因为分解利用了内部层之间的冗余可以进一步节省 FLOP 对于相同输入 $I \in R^{c \times h \times w}$ 和输出 $O \in R^{c \times h \times w}$, T 型 Conv 的 FLOPs 可以计算为:

$$h \times w \times (k^2 \times c_p \times c + c \times (c - c_p)) \quad (3)$$

而 PConv 与 PWConv 的 FLOPs 为:

$$h \times w \times (k^2 \times c_p^2 + c \times c_p) \quad (4)$$

公式 3 中 $c > c_p$, 并且 $c - c_p > c_p$ 。此外, 我们可以很容易地利用常规的 Conv 进行两步实现。

3.2. FasterNet

FasterNet [11]是在 PConv 和 PWConv 的基础上构筑而成的, 本文将其用于替换 YOLOv5 原模型主干网络中的 C3 模块。FasterNet Block 的结构如图 5 所示, 他有四个分层阶段。它有四个分层阶段, 针对每一个阶段, 都会存在着嵌入层或者是合并层。在嵌入层当中, 是 Conv 4×4 。而是合并层当中, 是 Conv 2×2 , 其步幅也不一致。前者的步幅是 4, 后者的步幅是 2, 这样也能够将其用于对通道数量的扩展当中。如果对 FLOPs 更为倾向, 则可以设置数量更多的 Fasternet 块, 并且给其分配更多的计算。

在每个 FasterNet 块当中, 都会有 PConv 层和 PWConv 层, 整体呈现出残差块。如果在网络系统中较多使用这些层, 就会导致特征多样性受到影响, 并且进一步降低了网络系统的整体性能。为此需要采

取必要的措施来保证网络系统的特征多样性, 以便取得较快的计算速度。在激活层选用 ReLU 来平衡运行时间和效率。最后三层为全局平均池化层, 一个 Conv 1 × 1 层和全连接层, 这些层用于特征转换和分类。

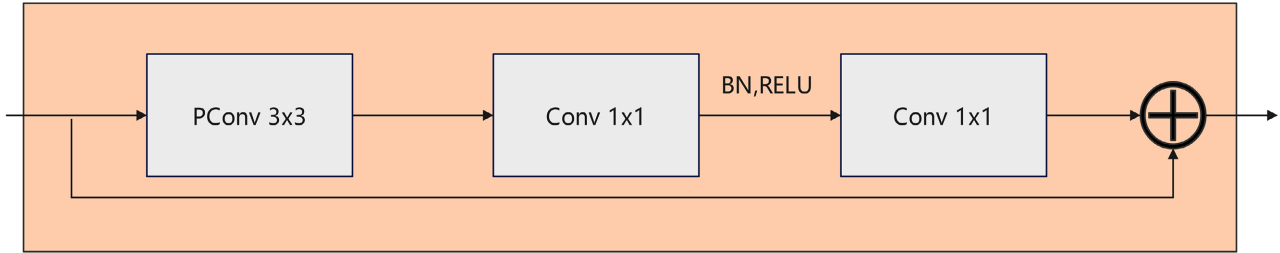


Figure 5. FasterNet Block structure
图 5. FasterNet Block 结构

原 YOLOv5s 模型通常在对特征较为复杂, 检测目标较大的数据集具有较好的效果(例如 COCO 数据集), 但在对血癌细胞检测时, 由于目标特征较少并且目标较小, 例如癌变细胞中的血小板和粘连的白细胞, 经过多次卷积容易导致丢失目标特征并且速度较慢。所以本文用 FasterNet 模块替换原 YOLOv5s 模型中的 C3 模块, 从而在提高精度的同时加快检测速度。

3.3. 注意力机制模块

在深度学习技术体系当中, 需要对感兴趣的目标或者是对象加以关注, 也就是采取注意力机制。神经网络中的注意机制有通道注意[12]、空间注意[13]和混合注意[14]等多种类型。Hou [10]等的研究成果表明, 通道注意力(例如, SE 注意力)虽然可以提升模型的性能, 但它们通常会忽略位置信息, 而 CA 模块将通道注意力按维度分解, 并分别沿不同的空间方向聚合特征, 捕获跨通道信息和位置信息。

采用 CA 模块, 能够使得对目标的定位更为准确, 同时定位的精度也会更高, 当加入了 FasterNet 之后, 就可以使得网络系统的表达能力更强。当使用这种算法进行血癌细胞的检测时, 就能够取得更高的检测效率, 同时检测的精度也会更高, CA 结构如图 6 所示。

CA 对输入的特征图 X。首先使用两个一维的池化核分别沿水平与垂直方向对每个通道进行编码, 第 c 个通道在其对应的高度和宽度输出可以分别表示为:

$$z_c^h(h) = \frac{1}{w} \sum_{0 \leq i \leq w} x_c(h, i) \tag{5}$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq i \leq H} x_c(i, w) \tag{6}$$

$$f = \delta(F_1([Z^h, Z^w])) \tag{7}$$

利用激活函数将 f 分解为 2 个单独的张量 f^h 和 f^w 。用 2 个 1 × 1 卷积变换分别将 f^h 和 f^w 变换为与输入 X 具有相同通道数的张量 $g^h = \sigma(F_h(f^h))$ 和 $g^w = \sigma(F_w(f^w))$ 。最后, 输出为特征图与两个权重图相乘:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \tag{8}$$

CA 模块增强了特征图的识别能力, 可以帮助轻量级的 YOLOv5s 模型更加精确的定位和识别感兴趣的目标。YOLOv5s-FasterNet 在 Backbone 中加入了两个 CA 模块。

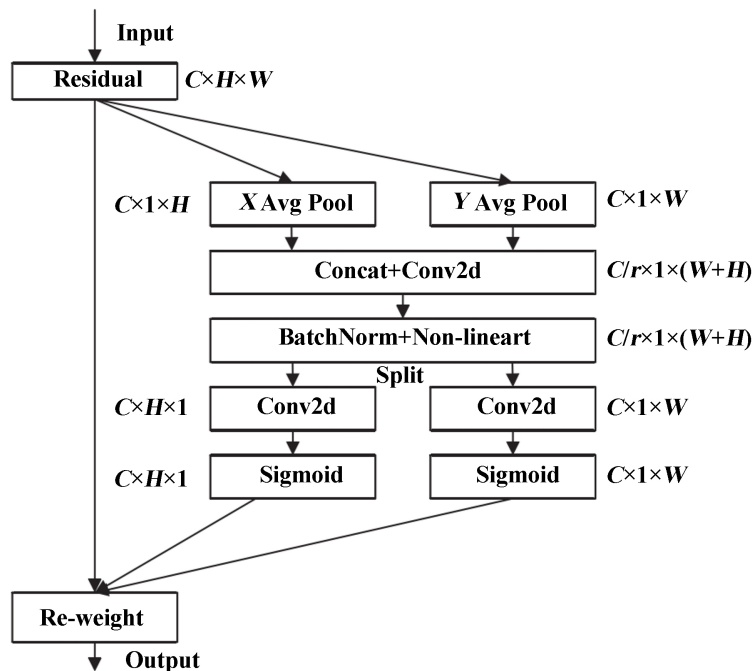


Figure 6. CA structure
图 6. CA 结构图

4. 实验分析

4.1. 实验环境

本实验的处理器为 AMD RYZEN 7745h, GPU 为 NVIDIA GeForce RTX4060, 深度学习框架为 Pytorch, 开发环境为 Pytorch 2.0.1, Python 3.10, Cuda 11.8, 优化器采用 Adam 优化器, 实验参数设置如表 1 所示。

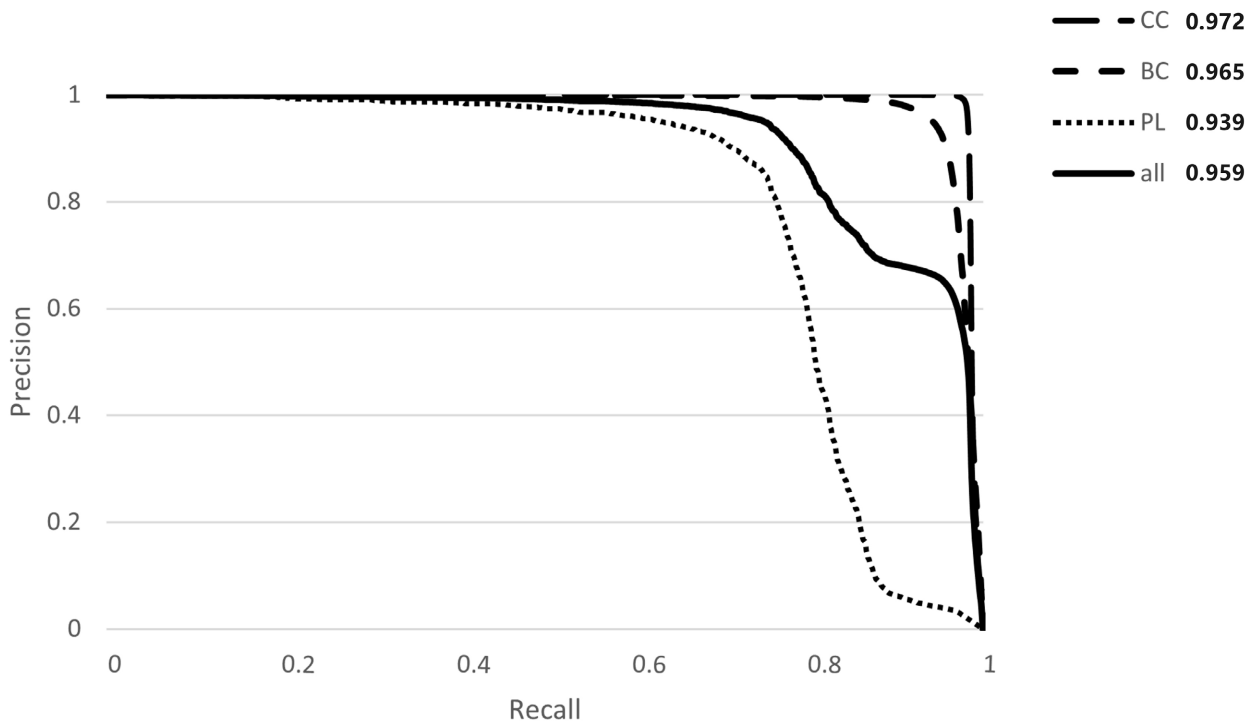
4.2. 数据集

本实验所采用数据集来源于慕尼黑 AML 形态学数据集, 这些图像取自 2014 年至 2017 年间在慕尼黑大学医院诊断为急性髓系白血病的 100 名患者以及 100 名没有血液恶性肿瘤迹象的患者的外周血涂片, 本实验从中选取 5000 中样本较好的图片, 并通过 Labelimg 对其进行人工标注, 标注分为三类: 癌变白细胞(Cancer Cell, CC), 红细胞(Blood Cell, BC)和血小板(Platelets, PL), 并按照训练集、验证集和测试集按 8:1:1 划分, 实验参数如表 1 所示。

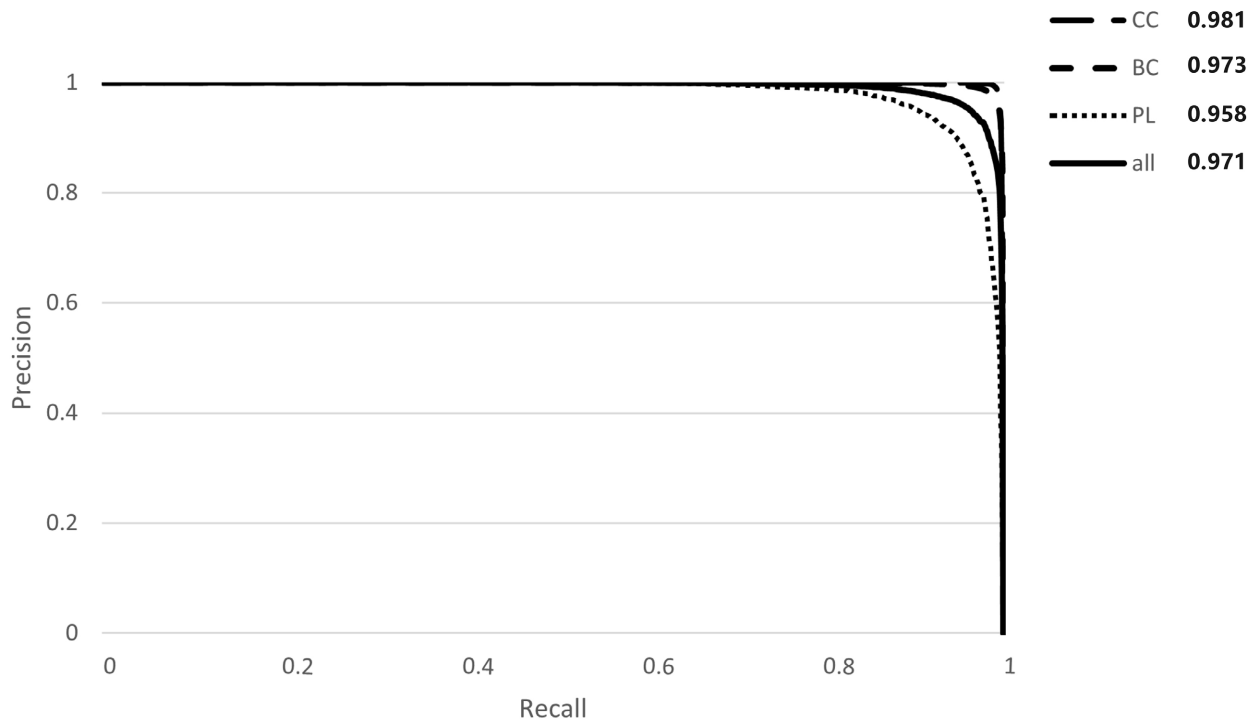
Table 1. Experimental parameters
表 1. 实验参数

实验参数	值
Learning rate	0.01
Train epoch	300
Batch size	16
Momentum	0.92
Weight decay	0.05
Confience	0.5
Nms threshold	0.3

4.3. 评价指标



(a) 改进前的 PR 曲线



(b) 改进后的 PR 曲线

Figure 7. PR curves before and after improvement

图 7. 改进前后的 PR 曲线

为了定量评估模型的检测效果, 使用准确率(Precision)、召回率(Recall)、平均精度(AP)、类平均精度(mAP)、FPS 作为主要评价指标。改进前后的癌变白细胞(CC)、红细胞(BC)和血小板(PL)的 P-R 曲线如图 7 所示。

查准率是衡量在血癌细胞检测任务中的是所有预测为正样本的结果中, 预测某细胞正确的个数(TP)占正确总个数(TP + FP)的比值。

查全率指的是表示血癌细胞的真实类别为正, 最后预测得到的结果也为正的个数(TP)占所有应该被检测到的细胞(TP + FN)的比例。

AP 用来计算单类别模型平均准确度, N 代表类别总数, 血癌细胞检测任务中的类别总数为 3, 各指标具体计算公式如下:

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

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

$$\text{AP} = \int_0^1 \text{PRdR} \quad (11)$$

$$\text{mAP} = \frac{\sum_{i=0}^n \text{AP}_i}{n} \quad (12)$$

4.4. 实验结果分析

本文将改进后的 YOLOv5s-FasterCA 模型与当前主流的一阶段、两阶段算法模型如 Faster-RCNN [15]、SSD 进行对比分析, 结果如表 2 所示。

分析表中数据可以看出本文算法与主流目标识别算法相比检测精度更高, 模型复杂度更小, 速度更快。与原 YOLOv5s 模型相比, 本文算法 mAP@0.5 提高 1.2%, 模型参数量减少 25.7%, 可以有效满足在医疗器械上进行实时检测血癌细胞的需求。

Table 2. Comparative analysis experiments

表 2. 对比分析实验

模型	Precision	Recall	mAP@0.5	模型大小/Mb	FPS
SSD	0.842	0.833	0.856	56.7	28
Faster-RCNN	0.856	0.844	0.869	41.8	13
YoloV3	0.792	0.805	0.827	15.5	34
YoloV4	0.847	0.854	0.876	14.2	36
YoloV5s	0.924	0.939	0.959	16.3	58
本文算法	0.946	0.957	0.971	12.1	67

4.5. 消融实验

为了更加直观验证有效性, 以下开展消融实验, 检测结果如图 8 所示, 具体对比数据如表 3 所示。

由表中数据可知, 在原模型基础上只加入 CA 模块时, 检测精度提高, 速度下降; 只加入 FasterNet 模块时, 检测精度略微提升, 速度大幅提高。同时加入 FasterNet 与 CA 时, 模型精度提升, 速度提升。

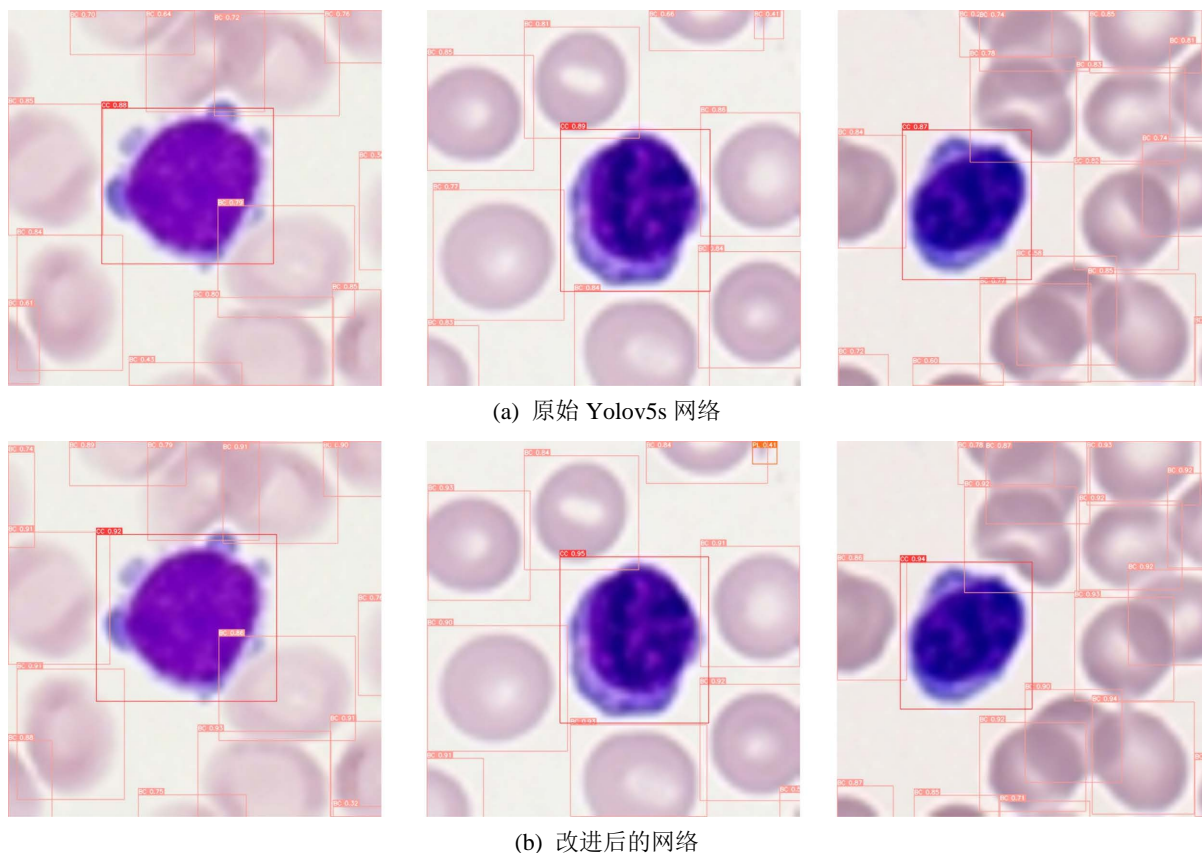


Figure 8. Comparison chart of test results
图 8. 检测结果对比图

Table 3. Ablation experiments
表 3. 消融实验

模型	Precision	Recall	mAP@0.5	FPS
YoloV5s	0.924	0.939	0.959	58
YoloV5s-CA	0.931	0.943	0.964	49
Yolov5s-FasterNet	0.928	0.942	0.962	76
Yolov5s-FasterCA	0.946	0.957	0.971	67

5. 结论

针对血癌细胞在复杂情况下识别难度大、检测精度低和用于医疗器械时速度慢的问题，提出了 Yolov5s-FasterNetCA 的改进网络。通过将 YOLOv5s 主干网络中的 C3 模块改进为 FasterNet，减少计算冗余，降低整体网络模型的复杂程度，大大提高了检测效率，通过在主干网络中加入 CA 模块，提高检测精度。

试验结果表明，本文改进算法 mAP@0.5 达到 97.1%，识别速度为 67 帧/秒，在提高速度保证实时性的同时保证了检测精度，能够满足实际血癌细胞检测时的需求。下一步将在保证检测速度的同时进一步提升检测精度，引入新的卷积模块或注意力机制，进一步提升在复杂环境下图像中的目标检测任务表现和泛化能力。

参考文献

- [1] Nazlibilek, S., Karacor, D., Ercan, T., *et al.* (2014) Automatic Segmentation, Counting, Size Determination and Classification of White Blood Cells. *Measurement*, **55**, 58-65. <https://doi.org/10.1016/j.measurement.2014.04.008>
- [2] Suryani, E., Palgunadi, S. and Pradana, T.P.N. (2017) Classification of Acute Myelogenous Leukemia (AML M2 and AML M3) Using Momentum Back Propagation from Watershed Distance Transform Segmented Images. *Journal of Physics: Conference Series. Iop Publishing*, **801**, 012044. <https://doi.org/10.1088/1742-6596/801/1/012044>
- [3] Shahin, A.I., Guo, Y., Amin, K.M., *et al.* (2019) White Blood Cells Identification System Based on Convolutional Deep Neural Learning Networks. *Computer Methods and Programs in Biomedicine*, **168**, 69-80. <https://doi.org/10.1016/j.cmpb.2017.11.015>
- [4] Liang, G.B., Hong, H.C., Xie, W.F., *et al.* (2018) Combining Convolutional Neural Network with Recursive Neural Network for Blood Cell Image Classification. *IEEE Access*, **6**, 36188-36197. <https://doi.org/10.1109/ACCESS.2018.2846685>
- [5] Zhang, D.D., Zhang, P.F. and Wang, L.S. (2019) Cell Counting Algorithm Based on YOLOv3 and Image Density Estimation. 2019 *IEEE 4th International Conference on Signal and Image Processing*, Wuxi, 19-21 July 2019, 920-924. <https://doi.org/10.1109/SIPROCESS.2019.8868603>
- [6] Girshick, R., Donahue, J., Darrell, T., *et al.* (2014) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 580-587. <https://doi.org/10.1109/CVPR.2014.81>
- [7] Redmon, J., Divvala, S., Girshick, R., *et al.* (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [8] Redmon, J. and Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. 2017 *IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 6517-6525. <https://doi.org/10.1109/CVPR.2017.690>
- [9] Liu, W., Anguelov, D., Erhan, D., *et al.* (2016) SSD: Single Shot Multibox Detector. In: Leibe, B., Matas, J., Sebe, N. and Welling, M., Eds., *Computer Vision-ECCV 2016, Lecture Notes in Computer Science*, Vol. 9905, Springer, Cham, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [10] Hou, Q.B., Zhou, D.Q. and Feng, J.S. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 20-25 June 2021, 13713-13722.
- [11] Chen, J., Kao, S., He, H., *et al.* (2023) Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 12021-12031. <https://doi.org/10.1109/CVPR52729.2023.01157>
- [12] Hu, J., Shen, L., Albanie, S., *et al.* (2018) Squeeze-and-Excitation Networks. *Proceedings of the 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>
- [13] Jaderberg, M., Simonyan, K. and Zisserman, A. (2015) Spatial Transformer Networks. *Proceedings of the 28th International Conference on Neural Information Processing Systems*, Montreal, 2017-2025.
- [14] Woo, S., Park, J., Lee, J.Y., *et al.* (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018. Lecture Notes in Computer Science*, Vol. 11211, Springer, Cham, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [15] Ren, S.Q., He, K.M., Girshick, R.B., *et al.* (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>