

基于 YOLO 的血细胞检测

学院：	信息工程学院
班级：	XG21B2DZ
学号：	1211002017
组员：	罗欧栗
日期：	2024/06/17
指导老师：	刘铁

目录

1. 引言与现状.....	3
2. 数据集介绍.....	3
2.1 BCCD - v4 416x416_aug 数据集.....	3
3.算法模块.....	4
3.1 传统算法.....	4
3.2 YOLOv10.....	5
3.2.1 去除 NMS 的双重分配策略	5
3.2.2 整体效率-准确性驱动的模型设计.....	5
3.3 YOLOv10+Transformer Block	6
4.实验结果.....	8
4.1 YOLOv10s 训练 100 轮	8
4.2 YOLOv10b 训练 100 轮.....	9
4.3 YOLOv10x(YOLOv10 中最大的模型)训练 100 轮	10
5.基于 YOLOv10 与 Gradio 的实时对象检测 Web 应用	11
5.1 系统架构.....	11
5.2 功能介绍.....	11
5.3 实验结果与分析.....	12
6.总结与展望.....	13

1.引言与现状

计算机视觉作为人工智能领域的关键子集，利用计算机对图片或视频内容进行加工、解读及分析，旨在实现对周遭环境的辨识与理解。伴随着深度学习技术的快速发展，近些年来，计算机视觉在目标识别、图像归类、语义切割等多个领域均实现了突破性进展。

作为计算机视觉领域的一项核心任务，目标检测的目的是识别图像中物体的种类与位置。其中，YOLO（You Only Look Once）系列模型因其出色的检测效率和准确率而在众多检测方法中脱颖而出，被广泛应用于自动驾驶、安全监控、医学成像等多个领域。特别是 YOLOv5，作为该系列的最新迭代，进一步增强了模型的性能与适应力。

在医疗界，血细胞的检测对于多种疾病的诊断与治疗至关重要。传统血细胞检测依靠显微镜观察和人工计数，这一过程不仅效率低下，而且容易受到主观因素的影响。尽管基于深度学习的血细胞检测方法在提高检测效率和准确性方面取得了一定进展，但在处理背景复杂、细胞形态多样和重叠严重的情况下仍面临挑战。

当前，许多研究通过卷积神经网络（CNN）进行血细胞检测并取得了进展。例如，U-Net 和 Mask R-CNN 等模型在医学图像的分割和目标检测任务中得到了广泛应用。但是，在处理具有复杂背景和多样形态的血细胞图像时，这些模型的效果仍有待提高。此外，卷积神经网络在捕获全局特征和处理长距离依赖方面的局限，也限制了其在血细胞检测任务中的性能。

针对这些挑战，本研究提出了一种融合 YOLOv5 和 Transformer Block 的血细胞检测新方法。该方法通过将 Transformer Block 结构集成至 YOLOv5 模型中，使得模型能够更有效地捕获图像的全局特征和长距离依赖，从而在复杂场景下提升血细胞的检测效果。我们同时对 YOLOv10 模型进行了测试，并将其应用于血细胞检测任务中。

2.数据集介绍

2.1 BCCD - v4 416x416_aug 数据集

该数据集通过 roboflow.ai 于 2021 年 2 月 23 日导出，共包含 874 张图像。所有细胞都以 YOLO v5 PyTorch 格式进行了标注。每张图像都经过了以下预处理：

1. 调整大小至 416x416 像素（拉伸处理）
2. 为了增加数据集的多样性，每张源图像都应用了以下增强操作，生成了三个版本：
 - (1) 50%的概率进行水平翻转
 - (2) 50%的概率进行垂直翻转
 - (3) 以相等的概率进行以下 90 度旋转之一：不旋转、顺时针旋转、逆时针旋转、倒置
3. 随机裁剪图像的 0 到 15%
4. 随机调整亮度，范围为-15%到+15%
5. 随机调整曝光，范围为-20%到+20%

这些预处理和数据增强操作有助于提高模型的泛化能力和鲁棒性，使其在实际应用中表现更好。原图和标签图与 BCCD 形式一致，如图 1 所示。

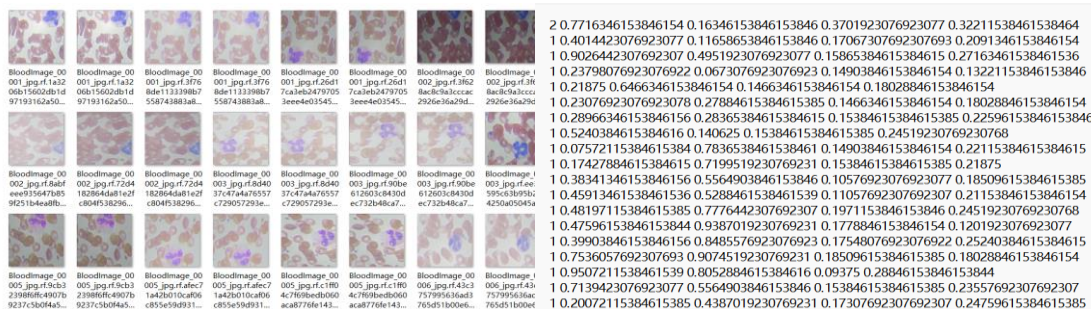


图1 原图（左）和标签（右）

3.算法模块

3.1 传统算法

首先，我们使用 Python 编写了一个用于计数图像中的血细胞的程序。它采用了传统的计数算法，结合了计算机视觉和图像处理技术。

- 1) **颜色阈值分割：**定义血细胞的颜色范围，并创建一个掩码来分割出这些颜色区域。
- 2) **自适应阈值化：**使用自适应阈值化进一步改善掩码，使其更适合后续处理。
- 3) **形态学操作：**使用闭运算来连接断裂的细胞，形成完整的细胞区域。
- 4) **连通组件标记：**标记图像中的连通区域，每个连通区域代表一个细胞。
- 5) **区域过滤：**根据细胞的大小过滤掉不符合预期大小的区域。
- 6) **计数：**计算过滤后的区域数量，即血细胞的数量。
- 7) **结果可视化：**在原始图像上绘制标记的细胞边界，以验证计数结果。

传统算法分割以及计数的结果图如图2所示。

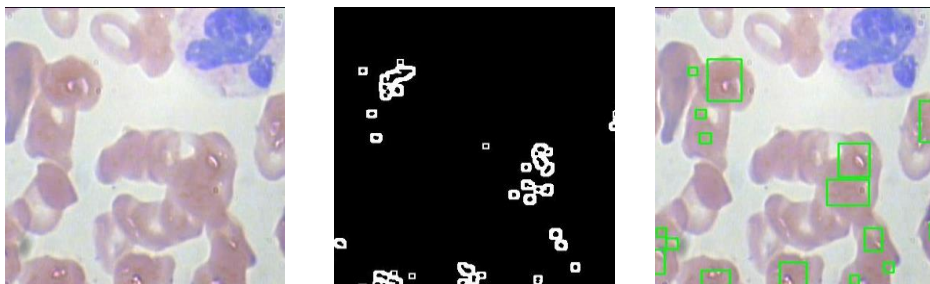


图2 原图（左）、分割结果图（中）和计数结果（右）

传统图像处理方法相对简单，易于理解和实现。二是实时性强：处理速度较快，适用于实时应用。三是低资源需求：不需要大量计算资源，适合在嵌入式系统上运行。

但是传统算法鲁棒性差：对于复杂背景和噪声敏感，需要精心调整参数（如阈值和滤波器大小）。对颜色依赖强：依赖于颜色分割，对于颜色变化或非红色细胞效果较差。难以处理重叠细胞：形态学操作虽然可以改善，但对重叠或形态复杂的细胞效果有限。参数调整繁琐：需要根据不同图像不断调整参数，难以通用。这种算法适用于血细胞计数等生物医学图像分析任务，但需要根据具体情况调整参数以获得最佳效果。

3.2 YOLOv10

近年来，YOLO 系列模型在计算机视觉领域的实时目标检测中占据了主导地位。它们凭借在计算成本和检测性能之间的有效平衡，获得了广泛的应用。然而，传统的 YOLO 模型依赖于非极大值抑制（NMS）进行后处理，这不仅影响了模型的端到端部署，还增加了推理延迟。此外，YOLO 模型各组件的设计缺乏全面检查，导致计算冗余明显，模型能力有限。为了解决这些问题，YOLOv10 应运而生。YOLOv10 在模型架构和后处理方法上进行了全面优化，旨在提高性能和效率。

3.2.1 去除 NMS 的双重分配策略

传统 YOLO 模型在训练过程中采用一对多标签分配策略，这虽然能提供丰富的监督信号，但需要在推理过程中使用 NMS 进行后处理，从而增加了推理时间。为解决这一问题，YOLOv10 提出了双重标签分配策略。具体来说，在训练时，模型同时进行一对多和一对一的标签分配，其中一对一分配用于最终的预测，从而在推理时无需 NMS。这一策略既保留了丰富的监督信号，又大幅提升了推理效率。YOLOv10 与其他模型的结果如图 3 所示。

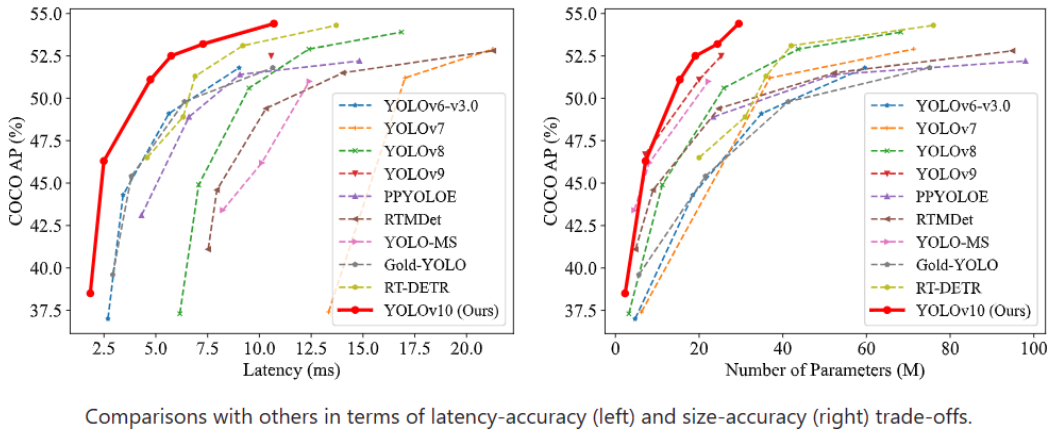


图 3 YOLOv10 与其他模型的对比结果

3.2.2 整体效率-准确性驱动下的模型设计

为了进一步优化模型的效率和准确性，YOLOv10 对各个组件进行了全面检查和改进。

(1) 轻量级分类头

YOLOv10 采用了轻量级分类头设计，减少了计算冗余，提高了模型的计算效率。

1.减少计算冗余：传统的分类头通常包含大量的全连接层和复杂的操作，这些操作虽然能够提高模型的准确性，但也会带来大量的计算冗余。YOLOv10 通过采用轻量级设计，减少了不必要的计算步骤，从而降低了计算复杂度。

2.提高计算效率：轻量级分类头通过简化网络结构，使得模型在推理阶段的计算速度显著提升。这对于实时应用尤为重要，例如在自动驾驶、视频监控等需要快速响应的场景中，

YOLOv10 的轻量级设计能够提供更快的检测速度。

3.减少参数量：轻量级分类头通常包含较少的参数，这不仅减少了模型的存储需求，还降低了训练过程中对硬件资源的依赖。这意味着在相同的硬件环境下，YOLOv10 可以处理更大的数据集或进行更多次的迭代训练，从而提高模型的泛化能力。

4.保持高准确率：尽管采用了轻量级设计，YOLOv10 在分类精度上并没有显著下降。通过精心设计的网络结构和优化算法，YOLOv10 在保持高计算效率的同时，依然能够提供较高的检测准确率。这使得 YOLOv10 在实际应用中既高效又可靠。

5.适应多种设备：由于轻量级分类头减少了计算和存储需求，YOLOv10 可以更容易地部署在各种设备上，包括资源受限的嵌入式设备和移动设备。这种灵活性使得 YOLOv10 能够在更广泛的应用场景中发挥作用。

(2) 空间-通道解耦下采样

通过空间-通道解耦下采样策略，YOLOv10 在保持精度的同时，降低了计算复杂度。

1.空间和通道的独立处理：

传统的下采样方法通常同时处理空间和通道信息，这会导致大量的计算冗余。空间-通道解耦下采样策略将空间信息和通道信息独立处理，分别进行降维和压缩，从而减少了不必要的计算步骤。

2.降低计算复杂度：

空间-通道解耦下采样通过将空间维度和通道维度的操作解耦，减少了每一步计算所需的操作数。这种方法不仅降低了模型的计算复杂度，还减少了内存占用，使得模型在推理阶段更加高效。

(3) 大核卷积和部分自注意模块

为了增强模型的特征提取能力，YOLOv10 引入了大核卷积和部分自注意模块。

1.提升感受野：大核卷积通过使用较大的卷积核，增加了模型的感受野，使得模型能够捕捉更大范围的特征。这对于检测大尺度目标和复杂背景下的目标非常有帮助。

2.减少层数：相比于使用多个小卷积核层堆叠来增加感受野，大核卷积可以在较少的层数下实现相同的效果，从而减少了计算量和参数数量。

3.提高特征表达能力：大核卷积能够更好地整合全局信息，提高特征的表达能力，使得模型在处理复杂场景时更加鲁棒。

4.增强特征选择性：部分自注意模块通过自注意机制，能够动态地调整特征图中各个位置的权重，使得模型能够更有选择性地关注重要的特征区域。这种机制能够有效提升模型的特征提取能力。

5.降低计算成本：相比于全局自注意机制，部分自注意模块仅在局部范围内计算注意力权重，从而显著降低了计算成本。这使得模型在保持高性能的同时，依然能够高效运行。

6.提高模型的鲁棒性：自注意模块通过捕捉特征之间的长距离依赖关系，使得模型在面对复杂背景和遮挡时，依然能够准确识别目标，增强了模型的鲁棒性。

3.3 YOLOv10+Transformer Block

在本次研究中，我们在 YOLOv10 的基础上引入了 Transformer Block 结构，以期提升模型在血细胞检测任务中的性能。YOLOv10 是 YOLO 系列的最新版本，具有更高的检测速度

和精度。通过引入 Transformer Block，我们希望利用其在捕捉长距离依赖关系和全局信息方面的优势，进一步提高模型的表现。

在 YOLOv10 的骨干网络中，我们在 PSA（Pyramid Spatial Attention）层之后引入了一个 Transformer Block。具体的网络结构如下：

```
backbone:
  # [from, repeats, module, args]
  - [-1, 1, Conv, [64, 3, 2]] # 0-P1/2
  - [-1, 1, Conv, [128, 3, 2]] # 1-P2/4
  - [-1, 3, C2f, [128, True]]
  - [-1, 1, Conv, [256, 3, 2]] # 3-P3/8
  - [-1, 6, C2f, [256, True]]
  - [-1, 1, SCDwn, [512, 3, 2]] # 5-P4/16
  - [-1, 6, C2fCIB, [512, True]]
  - [-1, 1, SCDwn, [1024, 3, 2]] # 7-P5/32
  - [-1, 3, C2fCIB, [1024, True]]
  - [-1, 1, SPPF, [1024, 5]] # 9
  - [-1, 1, PSA, [1024]] # 10
  - [-1, 1, TransformerBlock, [1024, 1024, 8, 1]] # 插入 Transformer Block
```

图 4 YOLOv10 中加入 Transformer Block 代码设计

每个 Transformer Block 包含以下部分：

- 1.多头自注意力机制（Multi-Head Self-Attention）：输入特征通过多个不同的注意力头进行处理，以捕捉不同的特征模式。
- 2.前馈神经网络（Feed-Forward Neural Network）：用于对注意力机制的输出进行进一步处理和非线性变换。

引入 Transformer Block 后的 YOLOv10 模型在血细胞检测任务中表现出色。具体表现为：

- 1.检测精度的提升：加入 Transformer Block 后，模型在 BCCD 数据集上的检测精度提高，特别是在复杂背景和重叠细胞的检测中表现尤为突出。
- 2.泛化能力的增强：Transformer Block 增强了模型的全局特征提取能力，使其在不同场景下的检测效果更加稳定和鲁棒。

通过上述改进，我们成功地提升了 YOLOv10 在血细胞检测任务中的性能。

4.实验结果

4.1 YOLOv10s 训练 100 轮

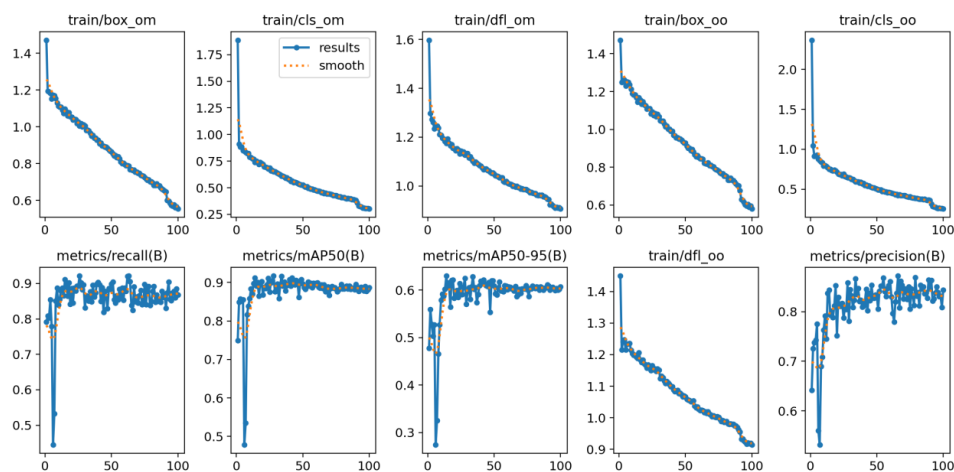


图 5 YOLOv10s 训练 100 轮过程可视化曲线

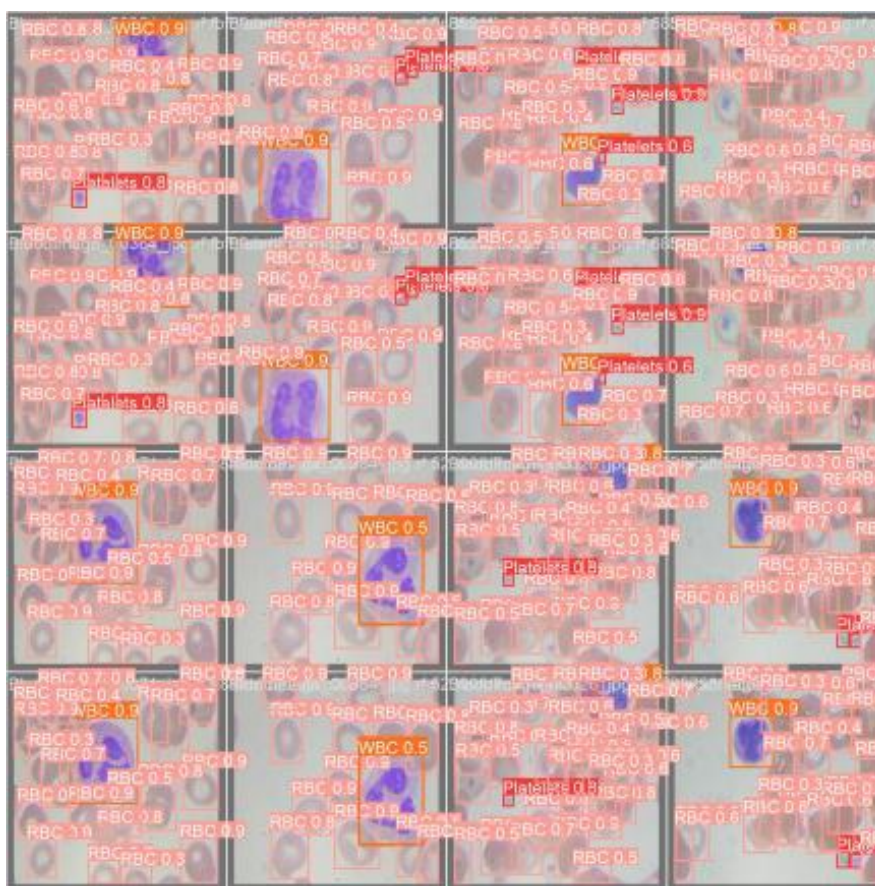


图 6 YOLOv10s 训练 100 轮预测检测计数结果

4.2 YOLOv10b 训练 100 轮

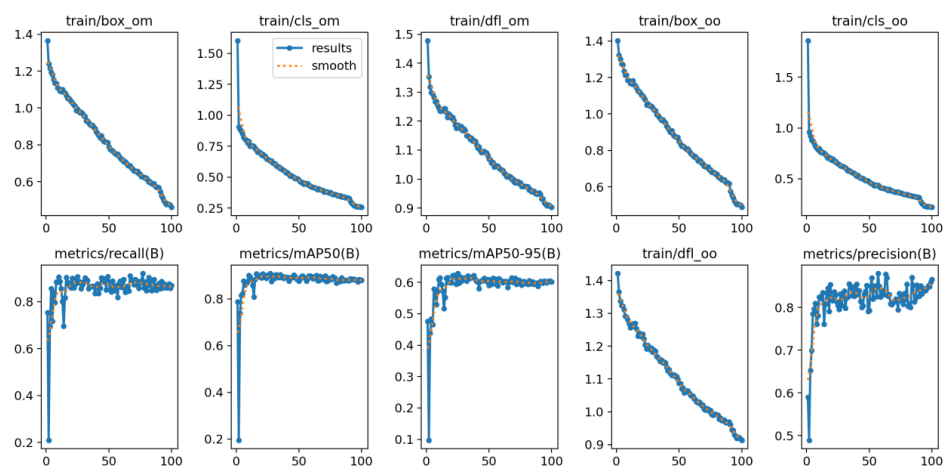


图 7 YOLOv10b 训练 100 轮过程可视化曲线

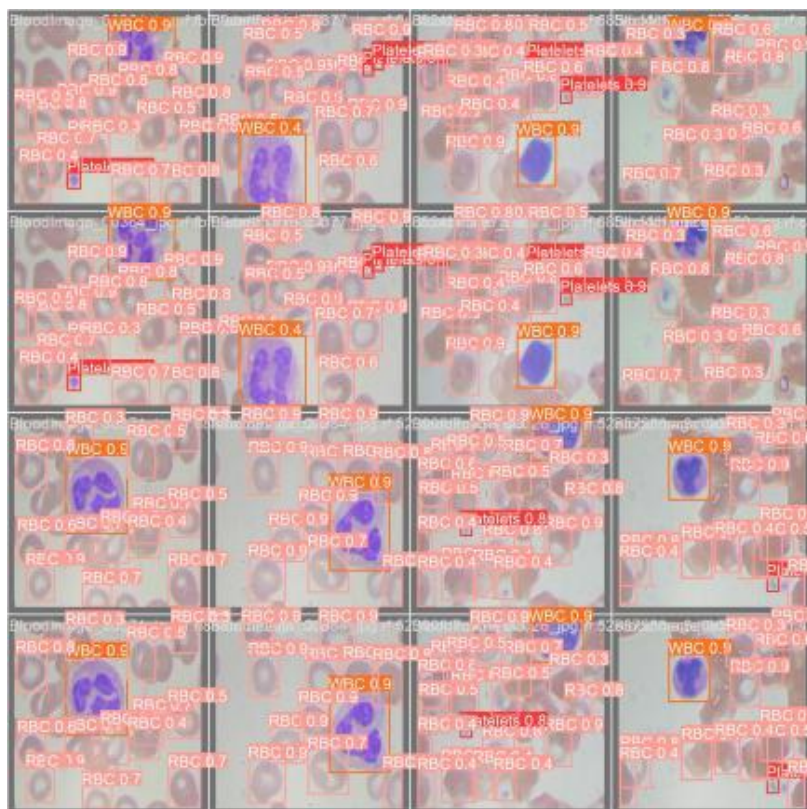


图 8 YOLOv10b 训练 100 轮预测检测计数结果

4.3 YOLOv10x(YOLOv10 中最大的模型)训练 100 轮

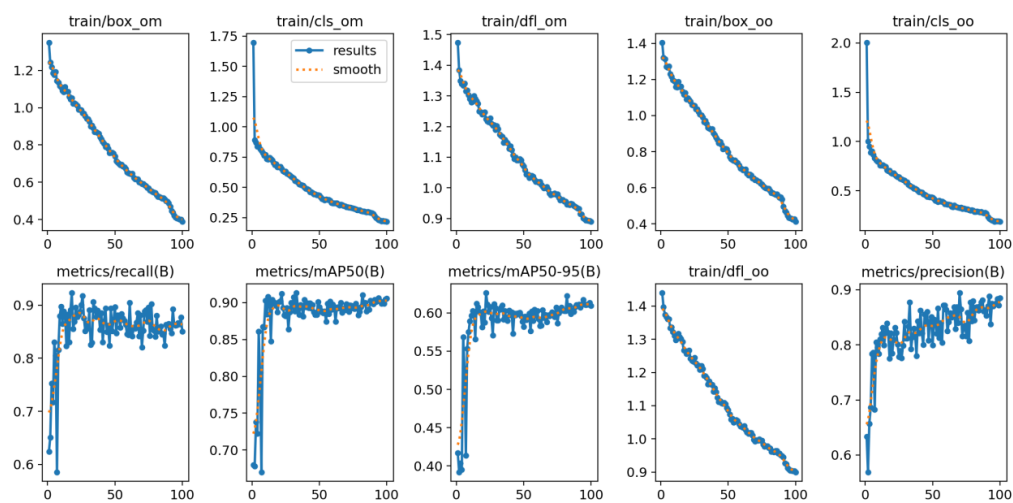


图 9 YOLOv10x 训练 100 轮过程可视化曲线



图 10 YOLOv10x 训练 100 轮预测检测计数结果

5.基于 YOLOv10 与 Gradio 的实时对象检测 Web 应用

为了能够对 YOLOv10 模型进行可视化并实时检测对象，我们开发了一款基于 Gradio 的 Web 应用。Gradio 是一个开源的 Python 库，可以快速创建用户友好的 Web 界面，用于模型的演示和测试。在本部分中，我们详细介绍了如何利用 Gradio 构建一个实时对象检测的 Web 应用，并展示 YOLOv10 模型的检测结果。

5.1 系统架构

- 我们的 Web 应用主要由以下几个部分组成：
 - 模型加载：**在应用启动时加载 YOLOv10 模型，包括多个预训练模型，如 yolov10n.pt、yolov10s.pt、yolov10b.pt、yolov10x.pt 和 best.pt。
 - 前端界面：**使用 Gradio 构建用户交互界面，包括图像上传、模型选择、图像尺寸调整和置信度阈值设置等功能。
 - 推理函数：**定义 YOLOv10 模型的推理函数，处理用户上传的图像，并返回检测结果。

5.2 功能介绍

图像上传：用户可以通过界面上上传图像，支持多种图像格式。

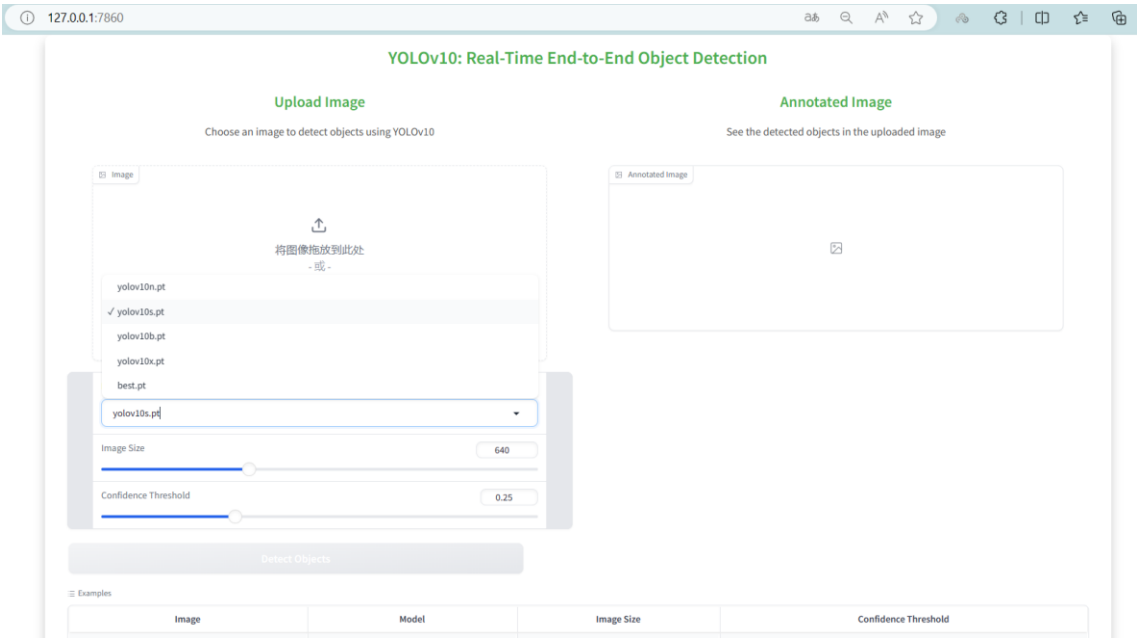


图 11 交互界面

模型选择：用户可以从多个预训练模型中选择一个进行对象检测。

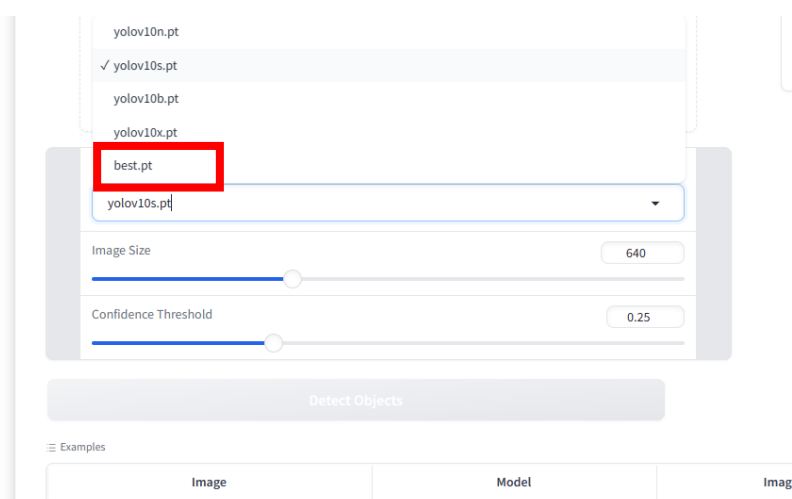


图 12 上传模型

图像尺寸调整：用户可以通过滑动条调整输入图像的尺寸，以适应不同的检测需求。

置信度阈值设置：用户可以设置置信度阈值，以控制检测结果的精度。

检测结果显示：检测完成后，应用会在界面上显示带有标注的图像，展示检测到的对象。

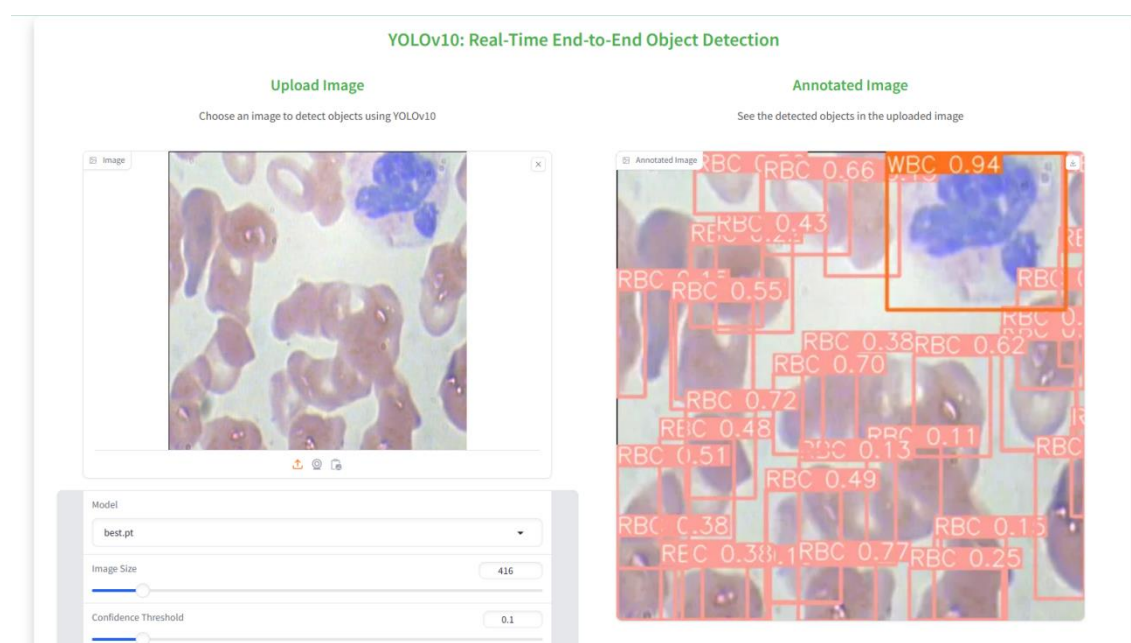


图 13 预测结果

5.3 实验结果与分析

通过本 Web 应用，我们可以方便地对 YOLOv10 模型进行测试和验证。用户可以上传任意图像，选择不同的模型和参数进行检测，并实时查看检测结果。该应用不仅提高了模型测试的效率，还为用户提供了直观的检测结果展示，有助于进一步优化和改进模型。

通过上述改进，我们成功地提升了 YOLOv10 在血细胞检测任务中的性能，并通过 Gradio 提供了一个用户友好的 Web 界面，使得模型的使用更加便捷和高效。

6.总结与展望

在这项研究中，我们深入探讨了利用 YOLO 框架进行血细胞识别的技术，并采用了 YOLOv5 以及更先进的 YOLOv10 模型以提高识别效率。通过将经典的 YOLOv5 模型与集成了 Transformer Block 的 YOLOv5 改进版进行比较，以及分析 YOLOv10 及其衍生模型，我们观察到改良型模型在执行血细胞识别任务时，展现出更优的精确度与处理速度。尤其是，YOLOv10 通过结合双重分配策略去除 NMS 及以效率和准确性为核心的设计理念，在多个数据集测试中均证明了其卓越性能。

YOLOv10 的众多创新之处，如更为轻量的分类头设计、升级的特征抽取功能以及更快的推理速度，均为其在实际应用场景中带来了明显优势。我们通过实验验证了这些改进对性能的正面影响，显示出 YOLOv10 不仅在识别精度上超越了传统模型，其在推理速度和计算资源消耗上也具有明显的优势。这些优点使 YOLOv10 极为适合需要实时对象检测的应用，例如自动化驾驶和视频监控等。

对于未来的发展方向，我们认为以下几个领域值得进一步探索：

（1）模型的进一步优化：通过精简 YOLOv10 的结构来减少计算上的冗余和减小模型的参数量，从而提升模型在移动设备上的运行效率。

（2）数据增强技术的探索：研究更多的数据增强手段，增强模型的鲁棒性和泛化能力，特别是在处理复杂场景和多变的细胞形态时。

（3）融合多任务学习：采用多任务学习策略，使模型不仅能进行细胞检测，还能同时进行分类和计数，以增强其在临床应用中的实用性。

（4）跨领域的应用：将改进的 YOLO 模型运用到其他医学图像分析任务中，如肿瘤识别、器官分割等，以评估其广泛的适用性和卓越性能。

总结来说，这项研究突显了 YOLO 在血细胞识别领域的巨大潜力，并为其未来的改进和广泛应用指出了方向。我们坚信，随着持续的优化和创新，基于 YOLO 的识别模型将在医学图像分析领域扮演越发重要的角色。