

基于机器视觉的化工多相液态界面智能检测

钱先锋, 薛艺淳, 丁蓓, 杨诚, 吴勇

(中钢集团马鞍山矿山研究院, 中钢矿院(马鞍山)智能应急科技有限公司, 安徽 马鞍山 243000)

摘要: 为解决化工生产过程中多相液体界面接触式检测的难题, 研究一种基于机器视觉的智能界面检测系统, 可实现多相液体界面的外置非接触式检测, 并研究界面检测算法的改进及数据集的构建。结果表明, 系统的检测误差率仅为5.3%, 模型精度达到86.3%, 满足了工业应用的精确度要求和实时检测效果, 且模型权重文件大小仅为13.6 MB, 可在边缘计算设备上部署, 在技术上具备了实时、精确、高效的特点, 成功地化工生产中的多相液体分液问题提供了创新的解决途径。

关键词: 界面检测; 机器视觉; 目标检测; YOLOv5; 非接触式检测

中图分类号: TP29; TP391.7; TQ056.1 **文献标志码:** A **文章编号:** 1671-1807(2024)11-0229-09

在化工生产过程中, 多相液体的分离是一个关键步骤, 这通常涉及油水等不相溶液体的界面检测。在此背景下, 界面检测设备扮演了至关重要的角色。传统上, 人工监测是常见的检测手段。然而, 随着自动化技术的进步, 越来越多的化工企业开始采用各种界面检测设备来执行这一任务。目前市场上普遍应用的界面检测设备包括差压式、伺服式、电导式、超声波式以及辐射式界面仪^[1-4]。这些技术均相对成熟, 可用于替代人工监测。

尽管辐射式界面仪能够实现无接触式测量, 但其高昂的成本和高风险(如使用 γ 射线监测原油液位)限制了其大规模应用^[4]。而在常规的界面检测方法中, 传感器需长时间浸入液体中, 这在强腐蚀性或高黏稠性介质中可能导致传感器发生堵塞和腐蚀, 从而影响设备的性能。

针对这些问题, 已有研究基于机器视觉原理开发出视觉界面检测技术, 这种技术通过摄像头与液体介质之间的玻璃隔离实现非接触式检测, 并通过特定的算法来识别多相界面^[5]。尽管视觉界面仪能有效识别动态多相液态界面, 但其对环境光照条件要求严格, 并且成本较高, 因此并不适合所有化工企业。

深度学习技术在解决复杂的视觉识别难题上显现出了无可比拟的优势, 然而, 这种优势背后是

对存储器容量和计算能力的高度依赖, 尤其是其所涉及的大量模型参数, 对普遍具有有限信息技术(information technology, IT)基础设施的中小企业形成了实质性挑战。另一方面, 工业场景所特有的数据属性, 如数据的高度敏感性和非公开性, 显著增加了获取足够且适宜的数据用于模型训练和验证的难度, 进而成为阻碍工业智能化进程的一大障碍。随着深度学习应用范围的不断拓宽, 学术界对于开发精简高效的实时检测模型以及探索在小样本条件下的深度学习算法愈发重视^[6-11]。本研究响应这一趋势, 采用了基于深度学习框架设计的实时目标检测模型, 目的在于攻克工业环境下多相液体界面实时高精度检测的问题。

本文旨在解决化工分液过程中多相液体界面检测的问题, 具体内容包括: ①设计研发一套基于机器视觉的智能界面检测系统, 令该系统融合硬件层(摄像头和计算设备)和算法层(目标检测算法), 以实现图像的获取、传输和实时识别功能, 实现化工多相液体界面的外接非接触式检测; ②对YOLOv5算法^[12]进行了针对性改进, 引入了SimAM注意力机制^[13]、EIou损失函数^[14]以及IoU-Guided NMS算法^[15], 以期提升模型在界面检测任务上的推理效果并保持实时性; ③通过实景拍摄, 并借助数据增强技术对数据扩充, 构建多相液

收稿日期: 2024-02-28

作者简介: 钱先锋(1981—), 男, 安徽枞阳人, 硕士, 高级工程师, 研究方向为化工安全; 通信作者薛艺淳(1998—), 男, 山西平遥人, 硕士研究生, 研究方向为智慧化工、机器视觉; 丁蓓(1991—), 男, 安徽宣城人, 硕士, 中级工程师, 研究方向为化工领域安全生产智能化; 杨诚(1992—), 男, 安徽滁州人, 硕士, 中级工程师, 研究方向为化工自动化; 吴勇(1979—), 男, 四川武胜人, 高级工程师, 研究方向为化工及冶金工贸自动化控制系统集成。

体界面图像数据集,期望解决该领域缺乏数据集的问题,为界面检测系统的研究开发奠定基础。

1 界面检测系统

界面检测系统的整体架构如图1所示。该系统由硬件层和算法层两大部分构成,它们共同协作以确保检测任务的准确性和高效性。硬件层是该系统的基础,主要包括高精度的摄像头以及一系列计算设备,如主机、服务器和录像机等。这些设备为系统提供了强大的计算和处理能力,保证了图像数据的实时捕捉和处理。算法层作为系统的核心,负责实现目标检测算法。检测算法能够对摄像头捕获的图像帧进行快速而准确的识别检测,并输出检测结果。界面检测系统的硬件层和算法层共同构成了一个高效、稳定的检测框架,满足化工过程中多相液态界面的检测任务。

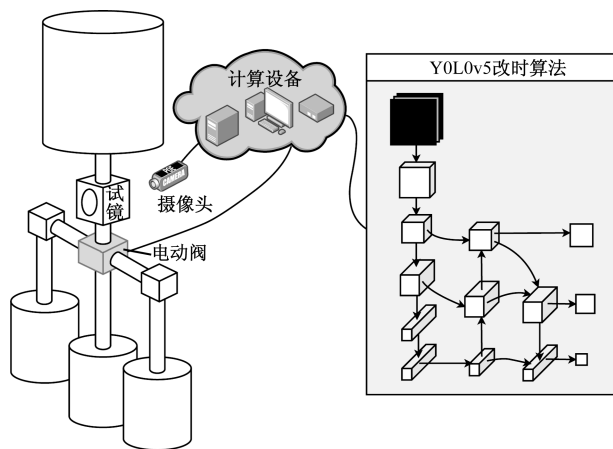


图1 界面检测系统架构

1.1 硬件层

界面检测系统的硬件层主要由摄像头和计算设备组成,这些设备紧密配合,以确保系统能够稳定运行并精确地完成检测任务。摄像头被部署在化工分离装置的分离前管道视镜外侧,其镜头方向对准管道视镜,专门用于捕捉液体的界面图像。这些图像对于监测和分析化工过程中的物质分离至关重要。

计算设备与摄像头通过有线或无线方式连接,为摄像头提供必要的计算资源。这些设备上运行的目标检测算法程序能够对摄像头捕获的图像进行高效识别和检测,并将检测结果输出。此外,系统还提供了输出信号的接口,这使得系统能够将检测结果传递给其他系统或设备,以便于实现实时的反馈控制,如控制电动阀门的开关等。

在工业场景中,计算设备通常包括工业主机、

工控服务器和微型计算机等。这些设备可以使用中央处理机(central processing unit,CPU)或图形处理器(graphics processing unit,GPU)来进行模型的推理计算,确保算法的高效执行。在选择计算设备时,企业可根据自身的实际需求和预算,选择性价比最高的设备,以实现最佳的性能和效益平衡。

总体而言,界面检测系统的硬件层设计考虑了化工现场的实际需求,通过精确的摄像头部署和强大的计算支持,确保了系统在各种复杂环境下的稳定性和准确性,为化工行业提供了可靠的监测和控制解决方案。

1.2 算法层

在算法层,界面的视觉检测功能通过部署于计算设备上的目标检测模型实现,该模型采用YOLOv5算法。YOLO系列算法基于深度学习,适用于实时目标检测领域,并在工业界得到广泛应用^[8-11]。YOLOv5作为该系列的第5个版本,以其快速的推理速度、高精度度及稳定性而备受推崇。尽管YOLOv7^[16]、YOLOv8^[17]以及基于Transformer的实时检测模型RT-DETR^[18]等SOTA(state-of-the-art,最新的)算法在性能上更为出色,但它们模型的复杂性较高,并且在CPU上的推理速度通常不如YOLOv5。鉴于多数中小企业在计算资源方面的限制,YOLOv5被选为界面视觉检测任务的基准算法模型。这一选择确保了在计算资源有限的条件下,界面的实时监控和精准检测仍能得以维持。

2 视觉检测算法

选择YOLOv5算法模型作为研究基准,针对其结构、损失函数及后处理算法进行深入优化。具体改进措施如下:首先,在模型Neck部分嵌入注意力机制SimAM模块,以增强模型对目标特征的关注度;其次,引入EIou,对原模型中的定位损失函数进行改进,更准确地量化预测边界与真实边界的差异;最后,将原模型的后处理算法NMS(非极大值抑制)替换为IoU Guided-NMS^[15],以优化目标检测的精度和速度平衡。通过这些方法,旨在提升YOLOv5算法在目标检测任务中的性能表现。

2.1 YOLOv5 算法

YOLOv5算法的核心理念在于将待检测的图像细分为 $S \times S$ 的网格单元,并对每个网格单元内的内容进行分类判断。如果内容不属于背景类别,该单元将生成一个锚点,即预测的目标框中心。每个锚点负责预测多个不同尺度的候选框,这些候选

框即为潜在的目标框。随后,通过非极大值抑制(NMS)算法,从这些候选框中筛选出最有可能的边界框,作为最终的检测结果。YOLOv5 算法的核心理念如图 2 所示。

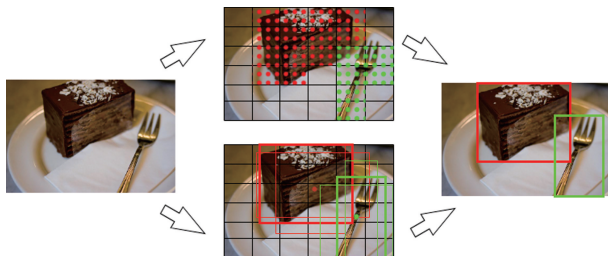


图 2 YOLOv5 算法思想

YOLOv5 的网络架构由 3 个主要部分构成:主干网络(Backbone)、颈部网络(Neck)和检测头部(Head)。Backbone 的主要作用是提取输入图像的特征,并为 Neck 部分提供多层特征图。Backbone 包括 5 个 CBS 层、4 个 C3 模块和一个 SPPF 模块。每个 CBS 层由一个卷积层、一个批量归一化层(BatchNorm)和一个 SiLU 激活层组成,以扩大模型的感受野并增强模子的稳定性。C3 模块主要由 1×1 卷积层组成,同时融合了残差网络的概念,以增加网络的深度和特征提取能力。SPPF 模块是对传统 SPP 模块的改进,它能够避免图像失真等问题,同时解决了对特征重复提取的问题,且相较于 SPP 模块,结构更简单,速度更快。

Neck 部分采用特征金字塔网络结构,对 Backbone 提取的特征图进行上采样和下采样,以融合不同尺度的特征,进一步提升网络的特征提取能力。Head 部分负责预测目标物体的类别和位置,由 3 个检测头组成,分别针对不同尺度的目标进行检测,以提高模型的检测精度和减少漏检率。

2.2 SimAM 注意力机制

注意力机制在模仿人类视觉和认知系统的基础上,有效地引导神经网络聚焦于输入数据中与任务最为相关的部分,从而精准提取关键信息,显著增强模型的性能和泛化能力。注意力机制的典型类型包括通道注意力机制、空间注意力机制以及自注意力机制。

针对界面检测任务,该任务在化工领域具有重要应用,因此算法的改进必须考虑到在低算力设备上的适配问题。SimAM 模块融合通道注意力和空间注意力,实现对特征的全方位、立体化提取。与传统的注意力模块不同,SimAM 模块直接计算三

维权重,省去了繁琐的参数推导过程,简洁而高效。SimAM 模块基于神经科学理论,设计了一种能量函数,通过优化这一能量函数来发现每个神经元的重要性,并导出了能量函数的解析解以加快计算速度。SimAM 模块的计算公式为

$$\tilde{\mathbf{X}} = \text{sigmoid}\left(\frac{1}{\mathbf{E}}\right) \times \mathbf{X} \quad (1)$$

式中: $\mathbf{X}$ 为特征输入,它连接到前一层神经网络的输出; $\mathbf{E}$ 为各通道能量函数的矩阵,其中特征 t 的能量函数 e_t 为

$$e_t = \frac{4(\hat{\sigma}^2 - \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (2)$$

式中: $\hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i$ 和 $\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2$ 分别为每个通道上的均值和方差, M 为每个通道中特征的数量; λ 为一个常数,通常设置为 10^{-4} 。

通过将 SimAM 模块应用于 C3 模块进行改进,可以形成 BottleNeck-SimAM 模块,具体方法是在 BottleNeck 模块中引入 SimAM 模块。同样地,在 C3 模块的最后,即 CBS 层之后添加 SimAM 模块,形成 C3-SimAM 模块。这样的改进结构如图 3 所示,不仅增强了模型的处理能力,也提高了模型在特定领域的适应性。

2.3 ElIoU 损失函数

在 YOLOv5 模型中,损失函数的设计旨在综合分类准确性与定位精确度,以达到最优的检测性能。该损失函数由两个主要部分组成:二元交叉熵(BCE)损失与完全交并比(CIoU)损失,其中 BCE 损失又包括分类 BCE 损失(BCE_{cls})和目标 BCE 损失(BCE_{obj})。损失函数 loss 的计算公式为

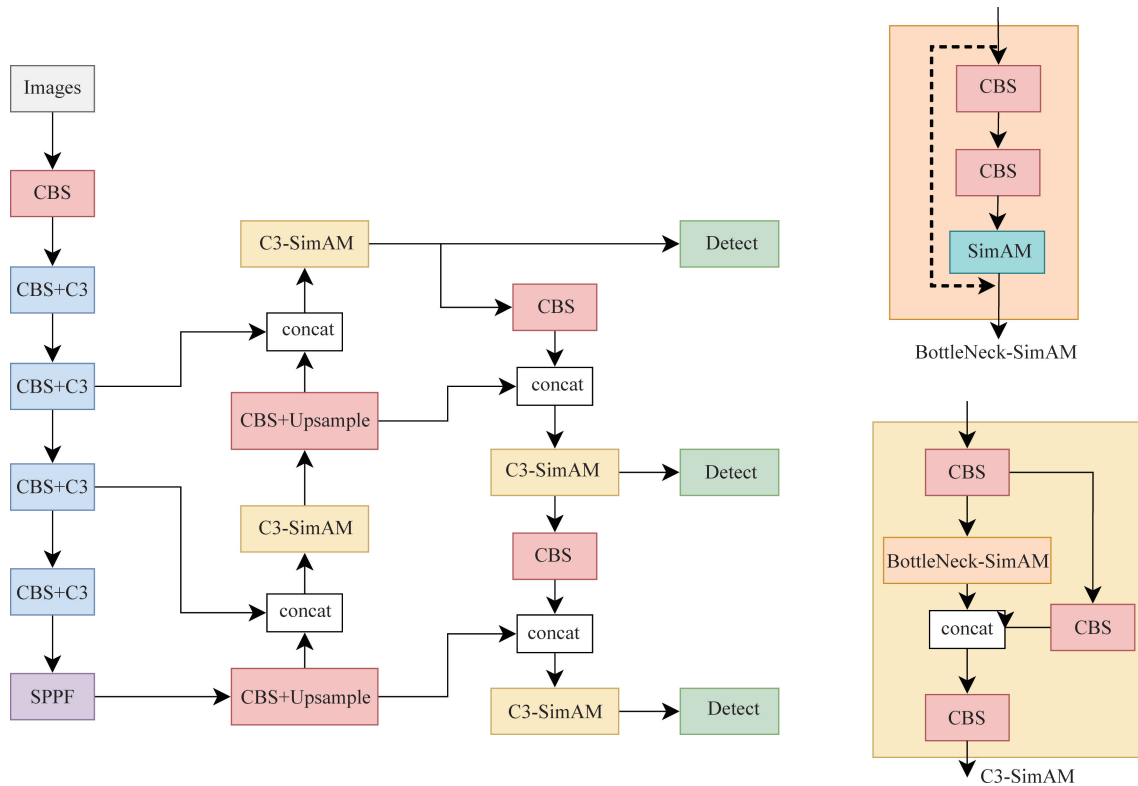
$$\text{loss} = \text{loss}_{\text{cls}} + \text{loss}_{\text{obj}} + \text{loss}_{\text{det}} = \text{BCE}_{\text{cls}} + \text{BCE}_{\text{obj}} + \text{CIoU} \quad (3)$$

式中: loss_{cls} 为分类损失; loss_{obj} 为目标损失; loss_{det} 为定位损失。

BCE 损失函数负责评估模型的分类性能。它衡量的是模型预测与真实标签之间的差异。具体地,BCE 通过计算预测类别概率与实际类别之间的误差来完成这一任务。BCE 损失的计算公式为

$$\text{BCE} = \mu[y \log_2 p(y') + (1 - y) \log_2 (1 - p(y'))] \quad (4)$$

式中: μ 为超参数, y 为二元标注向量,表示每个像素是否属于目标物体; y' 为预测分类向量, $p(y')$ 为模型输出的概率分布,表示每个像素属于目标物体的概率。



CBS、C3 为原 YOLOv5 模块;concat 层为拼接层,将两层网络输出进行拼接;Upsample 层为上采样层,使特征图的尺寸扩大;Detect 为输出层,是由一层 1×1 卷积构成

图3 YOLOv5-SimAM 网络结构

CIoU 损失函数则专注于提升定位的准确性。它通过考虑预测边界框与真实边界框之间的几何关系,以及它们之间的交并比(IoU)来计算定位误差。CIoU 损失式中:包含了 IoU(预测框和标注框的面积之比)、预测形心 b 与标注形心 b^{gt} 、欧氏距离 r^2 ,以及对角线长度 c 和用于计算宽高比一致性的权重函数 α 和一致性度量函数 ν , α 和 ν 的计算公式如式(6)和式(7)所示,其中 w 、 h 分别为边界框的宽和高,上标 gt 表示其为标注值。

$$CIoU = IoU - \left(\frac{r^2(b, b^{gt})}{c^2} + \alpha \right) \quad (5)$$

$$\alpha = \frac{\nu}{1 - IoU + \nu} \quad (6)$$

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (7)$$

尽管 CIoU 在提升定位准确性方面表现出色,但其对宽高比差异的过度关注可能会忽略其他重要的形状差异。因此,为了更全面地优化模型的形状,本文提出使用 EIou 损失函数,如式(8)所示, C 为常值系数。EIou 通过将 CIoU 中的单一宽高比损失分解为宽损失和高损失,来更细致地调整模型,从而加速模型的收敛,并提升回归的准确性。

$$EIou = IoU - \left[\frac{r^2(b, b^{gt})}{c^2} + \frac{r^2(w, w^{gt})}{C_w^2} + \frac{r^2(h, h^{gt})}{C_h^2} \right] \quad (8)$$

综上所述,通过结合 BCE 和 EIou 损失函数,YOLOv5 模型能够在保证分类精度的同时,显著提高定位的准确性,进而提升整体的目标检测性能。

2.4 后处理算法

在 YOLOv5 模型进行目标检测的过程中,产生多个边界框是为了准确捕捉到目标的位置和属性。然而,每个目标只需要一个最精确的边界框来定义其位置。为此,YOLOv5 采用了非极大值抑制(NMS)算法来筛选出最佳的边界框。NMS 通过评估每个边界框的分类置信度,并根据交并比(IoU)来去除多余的边界框。只有那些置信度和 IoU 都较高的边界框才会被保留。

在标准的 NMS 算法中,每个边界框都会被单独考虑,这可能会导致一些问题,如对于同一个目标产生多个检测框等。为了解决这些问题,研究者们提出了 IoU-Guided NMS 算法^[15]。这种方法通过引入与特定位置的置信度信息,来改进边界框的筛选过程。具体而言,IoU-Guided NMS 使用位置

置信度作为过滤标准,选取那些位置置信度高且分类得分最高的边界框,从而解决了传统 NMS 中存在的一些不足。

算法 1(NMS 算法)

输入:边界框集合 B,边界框打分集合 S,NMS 阈值 N_t
 输出:抑制后的边界框集合 D 及打分集合 S

```

1  D = {}
2  if B ≠ ∅ loop:
3      m = index[max(S)] # 取 S 中得分最高元素的下标
4      M = bm # 得分最高边界框
5      D = D ∪ M
6      B = B - M
7      for bi ∈ B loop:
8          if IoU(M, bi) ≥ Nt :
9              B = B - bi
10             S = S - si
11  return D, S
  
```

算法 2(IoU-Guided NMS 算法)

输入:边界框集合 B,分类置信度打分集合 S^{cls},IoU 预测置信度打分集合 S^{bbox},NMS 阈值 N_t
 输出:检测框集合 D 及分类打分集合 S

```

1  D, S = {}, {}
2  if B ≠ ∅ loop:
3      m = index(max(Sbbox)) # 取 IoU 得分集中得分最高元素的下标
4      M = bm # 得分最高边界框
5      B = B - M
6      s = Scls(M)
7      for bi ∈ B loop:
8          if IoU(M, bi) ≥ Nt :
9              s = max(s, Scls(bi)) # 保留最高分类得分
10             B = B - bi
11             S = S ∪ {s}
12             D = D ∪ {M}
13  return D, S
  
```

在针对化工动态液体多相界面检测的实际应用中,考虑到每个镜头画面中通常只包含一个视镜,且每个视镜中不超过一个界面,本文将目标类别细分为视镜和界面两类。为了提高模型的推理效率,对每个类别的最大目标数设置为 1,对总目标数的限制设置为 2,这样可以确保模型专注于检测每个场景中的关键目标,同时减少不必要的计算负担。通过这样的优化,模型的推理性能得到了显著提升,同时保持了高水平的检测精度。

3 数据及参数准备

为了进一步提高深度学习模型的性能,同时鉴于化工多相界面图像不易获取,遂在数据准备和增强方面进行了相应的工作。首先,围绕液体界面这一特定主题,通过网络搜集和实景拍摄相结合的方法

式创建了一个专门的数据集。该数据集共收集了 1 000 张图像。其中,网络图片 700 张,包括有界面图像 350 张和无界面图像 50 张;实景拍摄 300 张,包括有界面图像 150 张和无界面图像 150 张。这些图像全部涉及视镜/容器,并已按类别进行了标注,涵盖了有界面和无界面两种状态,为模型的学习提供了丰富的视角和场景。

考虑到原始数据集的规模和深度学习模型对数据量及质量的需求,使用数据增强技术,特别是采用了 Copy-Paste 图像增强技术^[19],将液体界面图像与从 MS COCO 数据集中选取的背景图像相结合,生成了额外的 4 000 张含目标的训练样本和 1 000 张纯背景样本,大幅提高了数据集的多样性。数据集样例如图 4 所示。

在模型架构方面,由于 YOLOv5 依赖于锚框进行目标检测,锚框的初始化对模型的训练结果具有重要影响。针对本数据集的独特性质,采用 K-means 聚类算法对数据进行精确的锚框尺寸划分,得到了 3 组预定义的锚框大小:[62, 17, 97, 19, 166, 40], [321, 36, 153, 76, 258, 54], [280, 108, 166, 210, 389, 604]。

迁移学习是一种有效的方法,它允许我们将在其他任务上训练好的模型所学到的特征迁移到新的任务中,以此初始化模型,从而提高训练效果。本文应用了迁移学习技术,通过将 YOLOv5s 模型在其他任务上学到的知识应用到当前任务中,不仅减少了所需样本量,也加快了训练速度,提高了训练效率。这种方法能够在保证模型性能的同时,显著降低时间成本。

4 实验分析

4.1 实验准备

实验硬件环境采用了海康威视的 DS-7104N-F1/4P 摄像头及硬盘录像机,搭载了 Intel © Core™ i9-13900K CPU,该 CPU 拥有较高的处理速度和多任务处理能力,确保了数据处理的效率和实时性。同时,NVIDIA RTX 3090 GPU 为实验提供了强大的图形处理能力,特别在进行深度学习和图像识别等计算密集型任务时表现卓越。

在软件环境方面,实验选用了稳定性好、安全性高的 Kylin V10 SP1 操作系统,该系统基于 Linux 内核,优化了资源管理和进程调度,为实验提供了一个稳定可靠的平台。同时,实验配置了 Python 3.9.16 解释器,并集成了 PyTorch 1.13.1 深度学习框架和 CUDA 11.6 编译环境,这使得实验能够充分利用



图4 数据集样例

GPU 资源,加速模型的训练和推理过程。

在度量指标方面,包括准确率 P 、召回率 R 、平均精度 AP 、平均精度均值 mAP ,计算公式如下:

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

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

$$AP = \int_0^1 P(R) dR \quad (11)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (12)$$

式中:准确率 P 表示模型正确预测的正样本数占总正样本数的比例;召回率 R 表示模型正确预测的正样本数占实际正样本数的比例;平均精度 AP 是针对每个类别而言,反映了模型对正样本的识别能力;平均精度均值 mAP 则是所有类别 AP 值的平均,是衡量模型总体性能的一个综合性指标;TP 为被模型正确识别为正类的真实样本数量,即被模型预测到的实际存在的目标实例;FP 为模型错误地将非目标样本或背景误判为了目标类别;FN 为那些

实际属于目标类别的样本,但在模型预测过程中未被正确识别,即真实存在但被忽视或否定预测的目标实例数目; N 为目标类别数。除了上述这些衡量模型性能的经典指标外,本实验还特别关注了 CPU 和 GPU 的推理速度,以及模型参数量。这些指标能够全面反映模型的性能,并且更符合工业实际应用,为后续的模型优化和选择提供了重要依据。

通过这些详细的配置和全面的度量指标,本实验旨在全面评估模型在处理目标分类任务时的性能,并为进一步的研究和优化打下坚实的基础。

4.2 模型改进及消融实验

首先,验证了所提出的 EIou 损失函数的有效性,并与 CIou、DIou 及 GIou 进行了详细的对比实验。实验结果见表 1,从中可以观察到,采用 EIou 损失函数的模型在 mAP 指标上取得了最高值。同时,各组模型在 GPU 与 CPU 上的推理时间及权重文件大小均无显著差异。特别是在 CPU 上,单张图像的平均推理时间为 23 ms,充分满足了实时检测的应用需求。

此外,对注意力机制 SimAM 模块安放位置进行实验。根据表 2 的结果可以看出,当 SimAM 模块添加于 Neck 部位时,模型的 mAP 值最高,比未添加 SimAM 的模型精度提高了 1.4%。而在 Backbone 部分添加 SimAM 则导致模型性能下降,精度比未添加 SimAM 组低 3.7%。在 Neck 部分和 Backbone 部分都添加 SimAM 后,mAP 比仅在 Backbone 部分添加 SimAM 的组别提高了 2.8%,进一步证明了 SimAM 添加于 Neck 部位的有效性。4 组实验的推理时间和权重文件大小基本相近,推理耗时最多相差 1~2 ms,都达到了实时检测的要求。此外,权重文件大小均未超过 14 MB,可方便地部署于边缘端设备上。

进一步地,对模型中的各项改进进行了详细的消融实验,以评估各部分改进对模型性能的贡献。消融实验包括了 4 组对照实验:原 YOLOv5 模型、用 EIoU 替换原定位损失、在 Neck 部分引入了注意

力机制的 SimAM 模块,以及用 IoU-Guided NMS 替代传统的 NMS。表 3 展示了消融实验的结果,通过对比可以发现,综合改进后的模型精度相较于原 YOLOv5 模型提高了 3%。尽管 CPU 上的推理耗时增加了 2 ms,但整体性能提升使得该改进模型依然能够保持在实时检测的性能要求之内。

为了全面评估本文改进模型的性能,选择了经典的实时检测模型以及当前状态下的 SOTA 模型进行了同条件下的对比实验,结果见表 4。实验数据显示,相较于经典轻量级模型,本文的改进模型不仅在检测精度上有所提升,同时也保持了实时推理的性能。特别是在 CPU 上的推理速度,仅稍逊于 YOLOv3-tiny 与 YOLOv6n,这使得该模型非常适合工业界中小型企业 的实际部署需求。此外,改进后模型的权重文件大小适中,便于在工业场景中广泛应用。

表 1 损失函数改进对比结果

实验组别	mAP@0.5			单图推理时间 (GPU)/ms	单图推理时间 (CPU)/ms	权重文件 大小/MB
	视镜	界面	整体			
CIoU	0.995	0.592	0.793	6.9	25.0	13.6
DIoU	0.995	0.694	0.844	8.3	25.1	13.6
GIoU	0.995	0.691	0.843	8.2	24.8	13.6
EIoU	0.995	0.700	0.847	6.7	25.4	13.6

表 2 网络结构对比结果

实验组别	mAP@0.5			单图推理时间 (GPU)/ms	单图推理时间 (CPU)/ms	权重文件 大小/MB
	视镜	界面	整体			
YOLOv5	0.995	0.700	0.847	6.7	22.1	13.6
YOLOv5+SimAM-Backbone	0.995	0.625	0.810	7.7	23.7	13.6
YOLOv5+SimAM-Neck	0.995	0.728	0.861	7.2	22.9	13.6
YOLOv5+SimAM	0.995	0.671	0.833	8.3	24.1	13.6

表 3 消融实验结果

实验组别	EIoU	C3-SimAM	IoU-Guided NMS	mAP@0.5	检测速度(CPU)/ms	检测速度(GPU)/ms	权重文件大小/MB
原生 YOLOv5	—	—	—	0.833	25.0	6.9	13.6
改进 1	✓	—	—	0.847	22.1	6.7	13.6
改进 2	✓	✓	—	0.861	22.9	7.2	13.6
改进 3	✓	✓	✓	0.863	22.9	7.3	13.6

注:—为未使用该项改进;✓为使用了该项改进。

表 4 网络对比结果

模型	mAP@0.5	检测速度 (CPU)/ms	检测速度 (GPU)/ms	权重文件 大小/MB
YOLOv3-tiny	0.693	21.8	7.9	17.0
YOLOv6n	0.823	19.5	7.0	10.0
YOLOv7-tiny	0.801	23.4	6.8	46.3
YOLOv8n	0.845	23.6	3.2	6.2
RT-DETR	0.812	135.8	10.6	77.1
YOLOv5s	0.833	25.0	6.9	13.6
本文改进	0.863	22.9	7.3	13.6

4.3 检测效果分析

为了检验系统的可用性,对整个检测系统的主要指标进行评估,相关指标值见表 5。

为了更直观地展示本文改进模型在化工液体多相界面检测任务上的效果,选取不同模型在测试集下的推理结果进行对比,如图 5 所示,选取了 4 张具有不同光照条件和界面样式的图像作为样例。图 5(a)是使用原始 YOLOv5 算法得到的检测结果,

图 5(b)是使用改进的 YOLOv5 算法产生的结果。可以看出,原 YOLOv5 模型存在误检情况,会导致产生多个界面检测框,相比之下,改进后的模型只检测到实际的界面,检测效果更好。

表 5 系统主要指标

指标	取值
单图推理时长(CPU)/ms	22.9
单图推理时长(GPU)/ms	6.9
图像捕获帧率/fps	25
误差率	0.053

5 结论

(1)设计并实施了一套基于机器视觉的智能界面检测系统,该系统实现完全非接触式检测,首创利用外置摄像头实现化工多相液体界面检测。该系统由硬件层和算法层两大部分组成。硬件层主要包括摄像头和计算设备,而算法层则通过界面检测算法实现界面的视觉检测。此外,算法层与硬件层通过通信接口紧密相连,以确保摄像头的图像捕获与相关计算控制设备的高效配合。同时,系统的检测误差率仅为

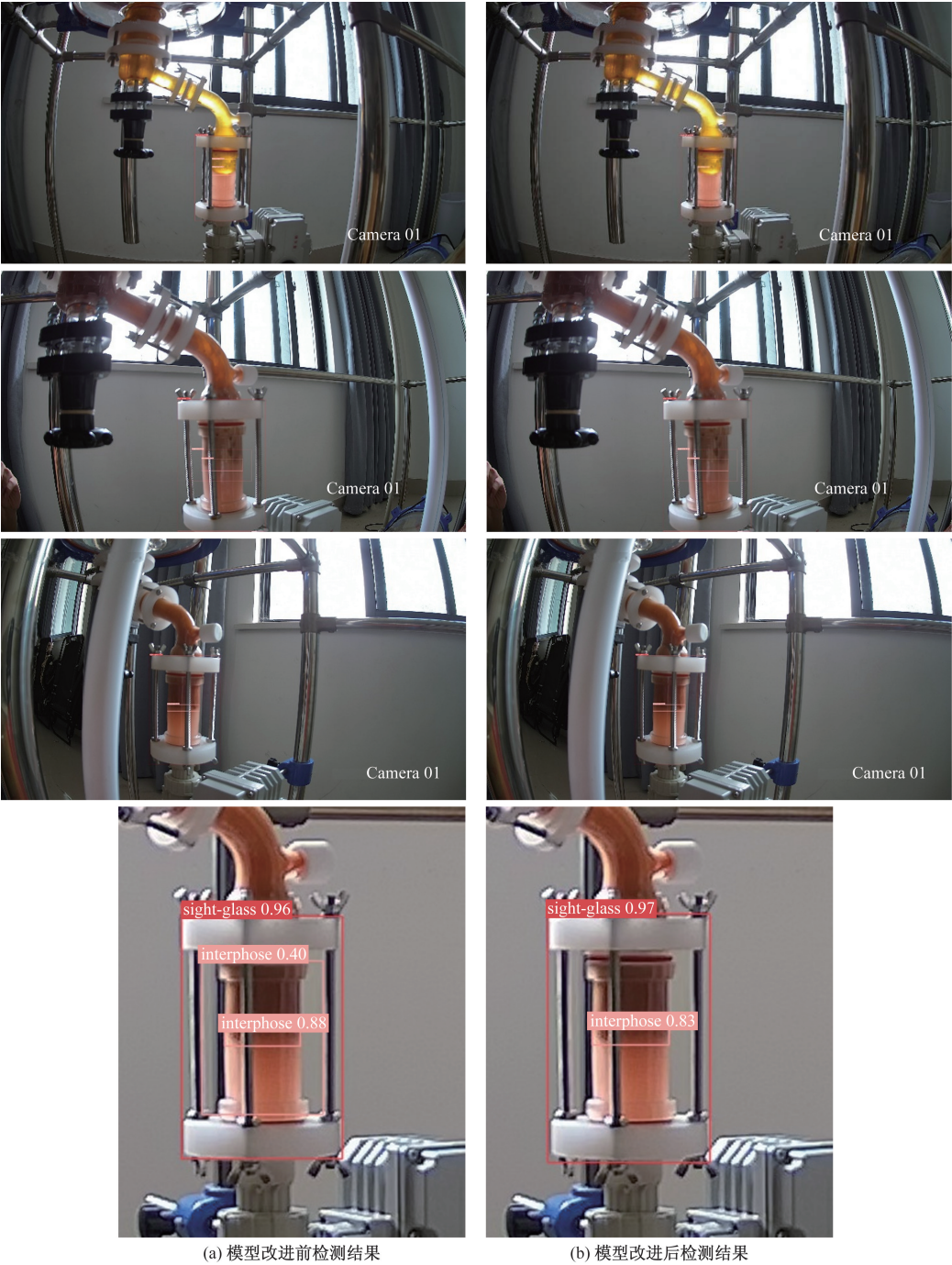


图 5 模型检测效果示例

5.3%,满足了工业应用的精确度要求。

(2)基于YOLOv5框架,设计并应用了一个专门针对界面检测任务的目标检测算法模型。为了提升模型的特征提取能力,同时保持推理速度,引入了Si-mAM注意力机制。在损失函数方面,采用了EIoU损失来提高模型训练效果。此外,引入了IoU-Guided NMS算法来改进模型的后处理过程,有效提升了检测精度。经过这些改进,模型的精度达到了86.3%,无论是在CPU还是GPU上,单张图片的推理时间均能达到实时效果所需的水平。同时,模型权重文件大小仅为13.6 MB,非常适合在边缘计算设备上部署。

(3)拍摄并构建了一个多相液体界面图像数据集,并应用了Copy-Paste数据增强技术对数据集进行扩充。这一举措为液体界面的智能化视觉检测任务提供了必要的数据支持。

参考文献

- [1] 叶尔兰·卡那提百克. 便携式原油岸罐明水测量装置的设计与使用[J]. 质量与认证, 2023(7): 60-62.
- [2] 刘闯, 张照, 彭正辉, 等. 电容式油水界面仪的设计与应用[J]. 石油化工自动化, 2021, 57(5): 63-65.
- [3] 刘畅. 精细化工企业实现中和、水洗过程自动化的对策探讨[J]. 山东化工, 2023, 52(8): 217-219.
- [4] 胡少聪. 原油储罐多相界面动态检测仪器的研究与开发[D]. 济南: 济南大学, 2022.
- [5] 黄俊杰, 沈鸿飞, 陈栋飞, 等. 一种用于化工界面检测的视觉界面仪及其检测方法: CN202210522272.2[P]. 2022-08-16.
- [6] 崔兆韵, 崔焕森, 宋德娟, 等. 改进卷积神经网络的冬小麦提取方法[J]. 科技和产业, 2024, 24(1): 281-286.
- [7] 任燕龙, 刘英宪, 侯亚伟, 等. 基于机器学习的海上多层砂岩油藏精细注水开发实践[J]. 科学技术与工程, 2023, 23(33): 14183-14191.
- [8] 詹云峰, 黄志斌, 付波, 等. 基于视觉的水面背景下目标检测与跟踪算法[J]. 科学技术与工程, 2022, 22(33): 14809-14819.
- [9] 张世玉, 高德欣. 电动汽车充电站火灾多目标实时检测与预警方法[J]. 科学技术与工程, 2023, 23(28): 12136-12144.
- [10] 张世强, 史卫亚, 张绍文, 等. 基于改进YOLOv5算法的钢铁表面缺陷检测[J]. 科学技术与工程, 2023, 23(35): 15148-15157.
- [11] 蒲玲玲, 杨柳. 改进YOLOv5的多车辆目标实时检测及跟踪算法[J]. 科学技术与工程, 2023, 23(28): 12159-12167.
- [12] OCHER G, STOKEN A, BOROVEC J, et al. ultralytics/yolov5: v5.0-YOLOv5-P6 1280 models, AWS, Supervisely and YouTube integrations[EB/OL]. (2023-03-29) [2024-03-01]. <https://docs.ultralytics.com/yolo-v5/>.
- [13] YANG L, ZHANG R, LI L, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks[C]//Proceedings of the 38th International Conference on Machine Learning. San Diego: PMRL, 2021: 11863-11874.
- [14] ZHANG Y, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [15] JIANG B, LUO R, MAO J, et al. Acquisition of localization confidence for accurate object detection[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018: 784-799.
- [16] WANG C, BOCHKOVSKIY A, LIAO H. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 7464-7475.
- [17] TERVEN J, CORDOVA D. A Comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS[J]. Machine Learning and Knowledge Extraction, 2023, 5(4): 1680-1716.
- [18] LV W, XU S, ZHAO Y, et al. DETRs Beat YOLOs on Real-time Object Detection[EB/OL]. arXiv preprint arXiv: 2304.08069, 2023. [2024-03-01]. <https://doi.org/10.48550/arXiv.2304.08069>.
- [19] GHIASI G, CUI Y, SRINIVAS A, et al. Simple copy-paste is a strong data augmentation method for instance segmentation[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 2918-2928.

Intelligent Detection of Chemical Multiphase Liquid Interface Based on Machine Vision

QIAN Xianfeng, XUE Yichun, DING Bei, YANG Cheng, WU Yong

(Institute for Intelligence & Emergency, Sinosteel Maanshan General Institute of Mining Research Co., Ltd., Sinosteel MIMR (Maanshan) Intelligency & Emergency Technology Co., Ltd., Ma'anshan 243000, Anhui, China)

Abstract: In order to address the challenge of contact-based detection of multiphase liquid interfaces in chemical production processes, an intelligent interface detection system based on machine vision was investigated. The system enables external non-contact monitoring of multiphase liquid interfaces and investigates improvements to interface detection algorithms along with the construction of a dedicated dataset. The results demonstrate that the system boasts a detection error rate of only 5.3%, with model accuracy reaching 86.3%, satisfying the precision and real-time detection requirements for industrial applications. Additionally, the model's weight file is a mere 13.6 MB in size, allowing for deployment on edge computing devices. It is concluded that the system exhibits real-time, precise, and efficient characteristics, providing an innovative solution to the problem of liquid separation in multiphase systems within the chemical production industry.

Keywords: interface detection; machine vision; objection detection; YOLOv5; non-contact detection