

改进YOLOv8的晶体熔接阶段图像视觉识别检测方法

李振成, 李桐, 胡涛, 张自主, 王海欣

北京印刷学院信息工程学院, 北京

收稿日期: 2025年12月5日; 录用日期: 2026年1月5日; 发布日期: 2026年1月12日

摘要

直拉硅晶体生长熔接流程中生长界面温度检测是保障后续引晶成功的重要任务。现有的目标检测模型对熔接凸点目标检测存在定位不准确, 误检率高等问题, 本文提出了一种改进了基于特征增强的YOLOv8算法。首先, 针对熔接光圈图像中小目标容易出现误检和漏检的常见问题, 引入BiFPN的思想对YOLOv8m中的颈部部分进行改进。为了进一步提升检测精度, 在特征融合网络中采用了更轻量的动态上采样算子DySample, 以提高融合特征的质量和丰富度。在工业提供的数据集上评估了YOLOv8-A模型, 实验结果表明, 与原来算法相比, YOLOv8-A的参数量和计算量分别减少至 2.19×10^7 , 同时实现了98.2%的mAP, 对小目标的检测提升了5.8个百分点。通过与其它主流目标检测算法比较, 验证了该方法的有效性和优越性。

关键词

缺陷检测, YOLOv8, 采样算子DySample, BiFPN模块, 轻量级网络

Improving the Visual Recognition and Detection Method for Crystal Fusion Stage Images in YOLOv8

Zhencheng Li, Tong Li, Tao Hu, Zizhu Zhang, Haixin Wang

School of Information Engineering, Beijing Institute of Graphic Communication, Beijing

Received: December 5, 2025; accepted: January 5, 2026; published: January 12, 2026

Abstract

The detection of the growth interface temperature in the Czochralski process for silicon crystal

文章引用: 李振成, 李桐, 胡涛, 张自主, 王海欣. 改进 YOLOv8 的晶体熔接阶段图像视觉识别检测方法[J]. 计算机科学与应用, 2026, 16(1): 72-86. DOI: 10.12677/csa.2026.161007

growth welding process is considered an important task for ensuring the successful subsequent crystal pulling. Existing target detection models exhibit issues such as inaccurate localization and high false detection rates in the detection of welding bump targets. An improved YOLOv8 algorithm based on feature enhancement is proposed in this paper. Firstly, to address the common problems of false and missed detections of small targets in welding aperture images, the concept of BiFPN is introduced to enhance the neck part of YOLOv8m. To further improve detection accuracy, a more lightweight dynamic upsampling operator, DySample, is utilized in the feature fusion network to enhance the quality and richness of the fused features. The YOLOv8-A model is evaluated on an industrially provided dataset, and experimental results indicate that, compared to the original algorithm, the parameter count and computational load of YOLOv8-A are reduced to 2.19×10^7 , while achieving a 98.2% mAP and improving small target detection by 5.8 percentage points. The effectiveness and superiority of this method are validated through comparisons with other mainstream target detection algorithms.

Keywords

Defect Detection, YOLOv8, Sampling Operator DySample, BiFPN Module, Lightweight Network

Copyright © 2026 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

单晶硅是电子信息产业和新能源光伏发电产业中最基本的原材料。直拉法[1]涉及的生产设备和工艺相对简单，便于实现自动控制，生产效率高，易于制备大直径单晶，是最常用的单晶硅生产技术。直拉法生产经过化料、熔接、引晶、放肩、转肩、等径，收尾 7 个工艺流程[2]后生长成单晶硅棒。

在单晶硅晶体生长阶段，合适的引晶温度是籽晶和熔硅接触后，籽晶周围逐渐出现光圈，最后光圈变圆。如图 1(a)所示若籽晶石正方形，则籽晶和熔硅接触的四条棱变成针状，面上呈圆弧形，圆弧直径略小于或等于籽晶断面的边长。如图 1(b)~图 1(c)所示，双光圈变成单光圈，表示温度达到 1458℃~1460℃ 左右，继续保持等待内点出现。当光圈内点出现后，开始回温(升温)如图 1(d)~图 1(f)所示，随时间推移，在回温过程中外光圈点出现，开始稳温，稳温过程中，根据点实际情况调整温度，至点饱满，进入引晶工序。

引晶过程中的熔接温度调节至关重要，因为平衡熔接温度的控制直接影响晶体的生长质量和形态[3]。现有晶体直径图像测量系统可以采集熔接过程中的籽晶光圈图像，操作工人通过查看采集到的籽晶光圈图像判断炉内单晶硅棒的生长情况是否符合标准，是否可以进入下一步工艺制备阶段。在实际生产中，由于摄像机拍摄角度、热屏形式、坩埚转速及材质等多种因素的影响，标准的“光圈”图像并非一成不变，并且，由于炉内液面的抖动和保护气体等干扰因素对红外传感器的影响，导致其测量结果不够精确，给质量控制带来了重大挑战。传统的熔液温度检测方法，如人工目视检查，效率低下且易受主观因素影响，导致检测准确性不足，即便是经验丰富的操作技师也难以避免温度误判。同时，传统的视觉检测方法依赖于手动设计的特征提取器，在处理复杂场景和多类“光圈”识别时表现不尽如人意，特别是在检测“光圈”图像的微小凸点方面上，凸点初显的图像与凸点饱满的图像无法区分开。

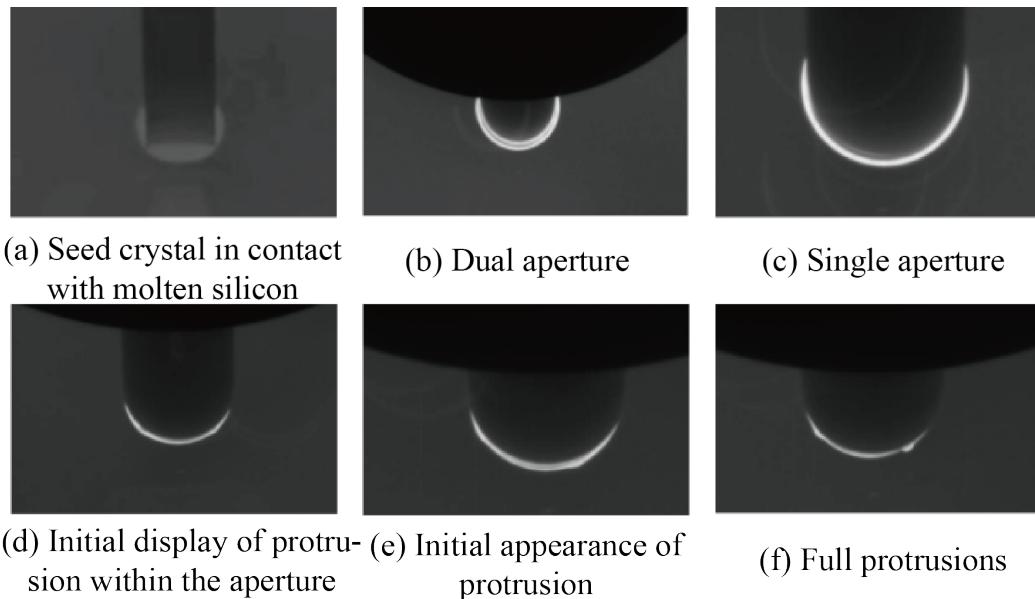


Figure 1. Aperture protrusions diagram during the temperature-raising process
图 1. 回温过程光圈凸点图

在当今科技飞速发展的背景下，人工智能技术[4]，尤其是深度学习算法，已在工业生产中扮演了关键角色，尤其在产品缺陷检测领域。随着深度学习在计算机视觉领域的成熟应用，特别是目标检测算法如 RCNN [5] (region based convolutional neural network) 和 YOLO [6] (you only look one) 系列的发展，为解决这一挑战提供了新的途径。针对上述问题和 YOLO 算法在目标检测领域的高效性、精确性以及快速响应能力，加之其成熟的部署技术，本文提出了一种基于 YOLOv8 架构改进的目标检测算法 YOLOv8-A。该算法专门针对单晶硅熔接过程，进行精确的光圈凸点质量检测，旨在提升生产过程中的温度监控水平。本文的主要贡献如下：

- (1) 在 YOLOv8 骨干模块中的特征提取结构中使用 GSConv [7] 和深度卷积来减少计算工作量，有效降低了模型的参数量和计算复杂度。
- (2) 为了充分利用多尺度特征信息，本文采用了特征融合策略[8]，在 YOLOv8 的 Neck 网络中使用了动态上采样算子(Dysample)，代替传统的上采样方法，有效融合了不同尺度的特征图，从而提高了模型对缺陷特征的提取能力，并且提升了模型对小目标的检测能力。
- (3) 将 YOLOv8 中的特征金字塔网络(PAN-FPN)替换为多向路径聚合网络 - 特征金字塔网络[9] (BiFPN)，并在 p2 层添加了上采样过程，以加强对小目标特征[10]的关注。通过这种方式，模型能够更全面地融合多尺度特征，实现了多角度、更全面的特征融合。

2. 相关工作

2.1. 传统计算机视觉的晶体熔接阶段图像识别

在传统图像处理领域，赵，程[11]等人提出的图像边缘特征信息检测及依据特征信息进行拟合的图像处理算法初步实现了晶体直径检测与监控，但是对于晶体生长出现的凸点无法准确识别。而赵，王[12]等人却在此基础上利用图像处理与模式识别相结合的方法，采用最小二乘支持向量机分类技术将光圈图像分类成不同的温度模式并进行温度高低的判断，实现了引晶温度的自动检测和辨识，但是在生产晶体制备过程中出现的更多不同种类图像时，错分率变高，模型不能实现学习有效特征。相较与传统图像识别，

在机器学习领域，杨，程[13]提出一种改进的 YOLOv5 网络模型用于直拉法生长单晶硅位错的检测，虽然取得了较高的检测精度，但检测速率有待提升。

2.2. YOLOv8 目标检测算法

目前主流的目标检测算法有两阶段的 RCNN、Fast RCNN、Faster RCNN、Mask RCNN 算法，一阶段的 SSD [14] (single shot multibox detector)、YOLO 算法，以及基于 Transformer 的目标检测算法 DETR [15]。两阶段目标检测算法是先产生候选框，对候选框中的内容进行特征提取，再对区域内容进行目标回归，通常这类算法具有较高的检测精度，但由于要先进行候选框的筛选，损失了一部分的检测速度。一阶段目标检测算法是基于回归的算法，将定位与分类任务合并，获得了更快的检测速度，在实时检测方面有着巨大的优势。基于 Transformer 的目标检测算法 DETR 将目标检测问题转化为一个对象查询问题，可以一次性检测整张图像中的所有目标，避免了传统目标检测方法中需要使用滑动窗口等方式进行多次检测的问题，提高了检测速度，而且 DETR 的可扩展性较强，可以很容易地扩展到新的目标检测任务中。然而，在单晶硅熔接过程“光圈”检测领域，这些基于深度学习的目标检测模型受到终端检测设备计算能力的限制，在模型部署上仍然面临着巨大的挑战，如何对计算量庞大的目标检测模型进行轻量化处理，使其可以部署在计算资源有限的设备上，是当前研究的一个热点问题。标准的 YOLO 模型通常可以分为三个部分：主干网络(特征提取网络 Backbone)、颈部网络(Neck)和检测头(Head)。其中，Backbone 是一个特征提取网络，用于从图像中提取特征信息；颈部网络可以融合从 Backbone 提取的特征，使得网络学习到的特征更加多样化，提高检测网络的性能；检测头可以利用之前的高质量特征工程做出准确的预测。几乎每一代 YOLO 模型都在这三种结构上做出了相应的改进和增强。

2.3. 选择 YOLOv8 作为基线的原因

YOLO 是目前最流行的实时目标检测器，在当前，YOLOv5 [16] 和 YOLOv7 [17] 是最广泛接受的两种算法。YOLOv5 中采用了深度学习技术来实现实时高效的物体检测任务。与前身 YOLOv4 相比，YOLOv5 在模型结构、训练策略、性能等方面都有所改进。YOLOv5 采用 CSP (CrossStage Partial) 网络结构，可以有效减少重复计算，提高计算效率。YOLOv7 提出了一种新颖的训练策略，称为 Trainable Bag of Freebies (TBoF)，用于提高实时目标检测器的性能。TBoF 方法包括一系列可训练的技巧，例如数据增强、MixUp 等，通过将 TBoF 应用于三种不同类型的目标检测器(SSD、RetinaNet 和 YOLOv3)，可以显着提高目标检测器的准确性和泛化能力。然而，不论 YOLOv5 还是 YOLOv7 也有一些缺点。比如在小物体检测方面还存在一些不足，对于密集物体的检测效果也有待提高。

Glenn Jocher [18] 提出 YOLOv8 作为 YOLOv5 的改进。在保留 YOLOv5 原有思想的前提下，参考 YOLOv7 中的 ELAN 结构设计了 C2f 模块，并用其取代了 C3 模块。头部部分也进行了修改，以使用解耦头部技术将分类和检测分开。此外，损失函数利用样本的正负匹配而不是 IOU 匹配。

与之前的 YOLO 算法相比，YOLOv8 的可扩展性[19]非常强。它是一个可以支持 YOLO 之前版本的框架，并且可以在不同版本之间切换，因此很容易比较不同版本的性能。YOLOv8 网络结构精简，导致更快的检测速度和更高的检测精度，因此，选择 YOLOv8m 版本作为实验的基准。

3. 改进 YOLOv8m 目标检测算法

3.1. YOLOv8 改进模型

由于本文光圈凸点的缺陷检测需要工业化应用，须同时考虑检测精度和检测速度，采用 YOLOv8m 网络模型，结构如图 2 所示：

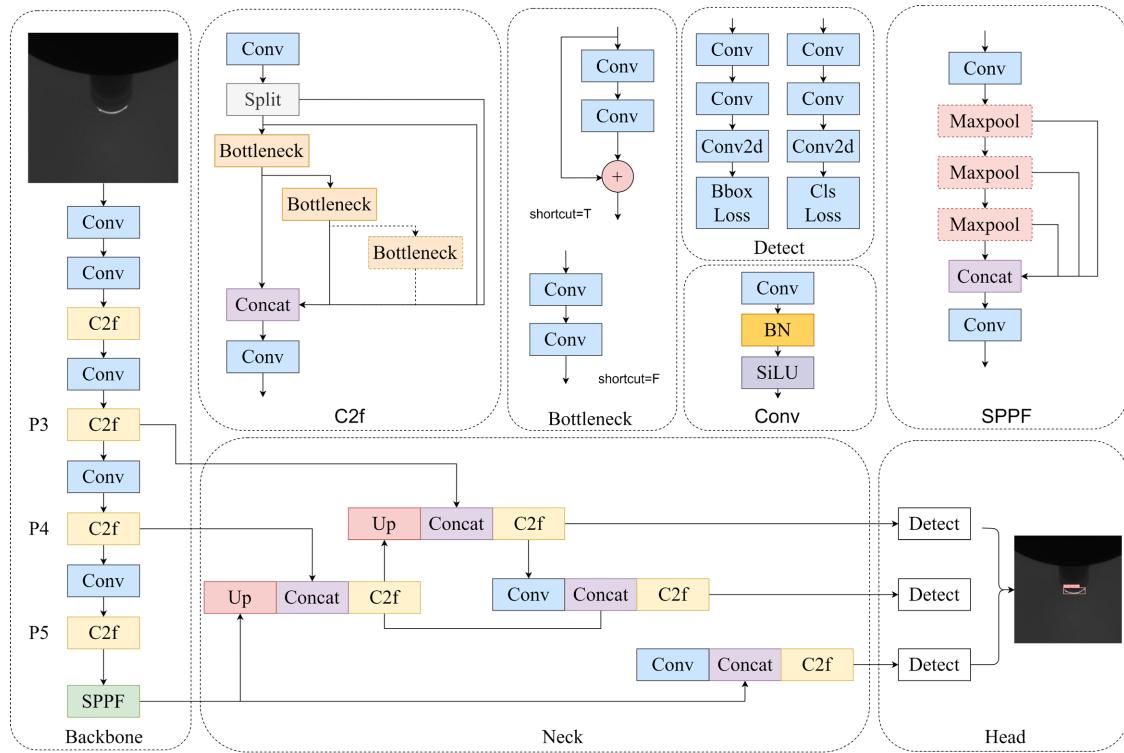


Figure 2. YOLOv8m structure

图 2. YOLOv8m 结构

YOLOv8m 模型整体性能值得称赞，但在识别小密集缺陷方面表现出一定的局限性。为了克服这些挑战并增强模型在检测光圈凸点微小面积方面的功效，本文主要增强了 YOLOv8 架构的三项功能。图 3 显示了增强型网络架构。

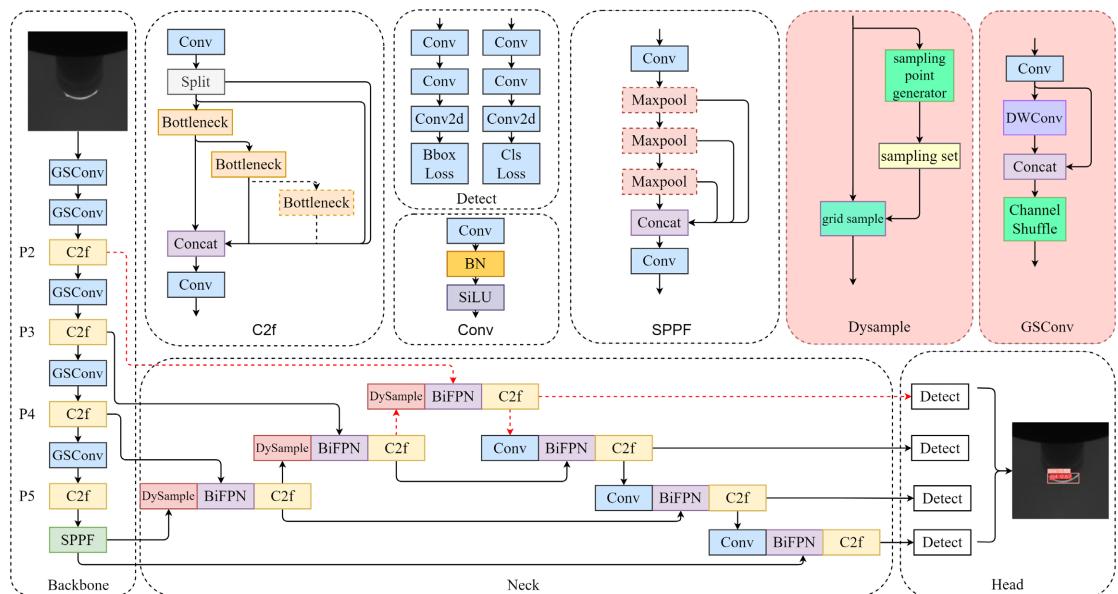


Figure 3. Improved YOLOv8-A structural

图 3. 改进的 YOLOv8-A 结构图

首先，本文将一种称为 GSConv 模块的轻量级卷积机制集成到 YOLOv8 框架中。这种集成的目的是增加模型对较小目标的关注，以提高其在晶体工艺熔接阶段检查期间识别和分类关键小凸点缺陷的能力。其次，通过 Dysample 上采样操作结合 p2 层改进了模型的多尺度信息融合，这也增加了小目标检测的准确性。最后，采用 BiFPN 轻量级下采样模块通过模拟人类视觉系统依赖上下文信息理解场景取代了 yolov8 传统的 C2f 模块，实现减少参数数量的同时有效提升熔接任务中的目标检测的精度。此修改的目的是将模型的注意力集中在更高质量、有代表性的示例上。

3.2. GSConv 模块

常见的 CNN 设计体系中有设计缺陷，即使用卷积步长或者池化层，这会导致细小目标的信息丢失或者低效的特征学习。当前的模型设计在处理小目标和低分辨率的图像时会遭受细小目标信息的丢失，无法充分学习小目标的特征。引入来替代 YOLOv8 模型中步长为 2 的卷积下采样模块，以很小的代价生成许多能从原始特征发掘所需信息的“Ghost”特征图[20]，达到减少网络计算量的同时，防止小特征丢失。

对于细小目标检测来说，速度和准确性同样重要。先前的轻量级工作，如 MobileNets [21] 和 ShuffleNets [22]，通过 DSC (深度可分离卷积) 操作极大地提高了检测器的速度。但是，当这些模型应用于细小目标时，这些模型的较低准确性令人担忧。实际上，这些工作提出了一些缓解 DSC 固有缺陷(也是其特殊性)的方法：MobileNets 使用大量的 1*1 密集卷积来融合独立计算的通道信息；ShuffleNets 使用“通道混洗”来实现通道信息的交互，主要解决深度可分离的卷积输入特征图中的通道信息在计算过程中被分离的问题；GhostNet [23] 使用“半分”SC (普通卷积) 操作保留通道之间的交互信息，解决标准卷积的输出通常有许多相似的特征图的问题。然而，1*1 密集卷积占用更多计算资源，使用“通道混洗”的效果仍然无法接触 SC 的结果，而 GhostNet 或多或少又回到了 SC 的路上，受影响的因素可能来自多个方面。许多轻量级模型使用类似的思路设计基本架构：从深度神经网络的一开始到结束只使用 DSC。但是，DSC 的缺陷在骨干网络中直接放大，无论是用于图像分类还是检测。本文认为 SC 和 DSC 可以合作。可以注意到，仅通过混洗 DSC 的输出通道生成的特征图仍然是“深度分离的”。为了使 DSC 的输出尽可能接近 SC，本文引入了一种新方法 - 使用 SC、DSC 和混洗的混合卷积，称为 GSConv。如图 4 所示，在 GSConv 的结构中融合了 GhostNet 和 ShuffleNetv2 的轻量级思想。相较于 GhostNet 卷积网络，保留了其减少相似的特征图问题，并且没有繁琐的步骤，从而提升了模型的准确性。

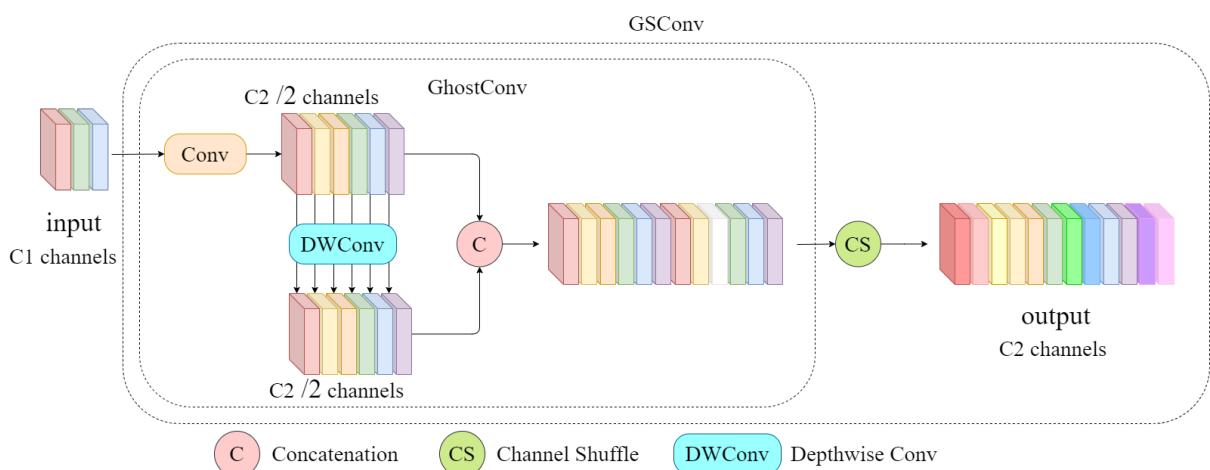


Figure 4. Structural of GSConv

图 4. GSConv 结构图

图 5 可视化了添加 GSConv 卷积前后的中间特征图。第 1 列显示输入图像以及目标检测效果，第 2 列表示输入特征图，第 3、4、5 列分别对应经过第二、第三、第四个 C2f 模块后获得的特征图。行(a)显示没有添加 GSConv 卷积的模型的图，而行(b)显示应用所提出的 GSConv 卷积方法之后的模型的图。

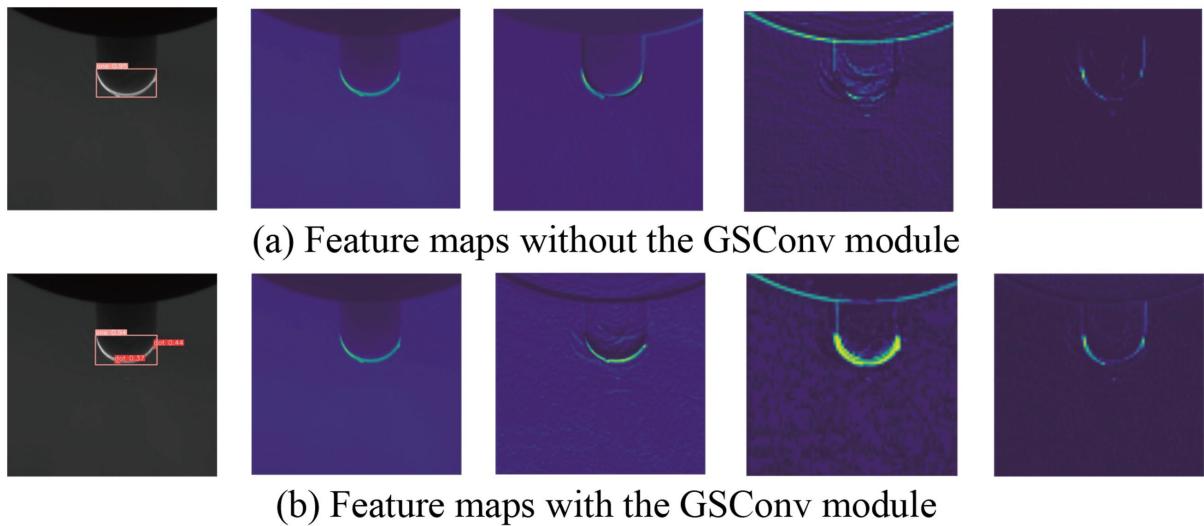


Figure 5. Visualization of feature maps

图 5. 特征图的可视化

3.3. Dysample 模块

上采样是关于对几何信息进行建模。工作流还通过动态采样图像或特征图来建模几何信息，作为标准网格采样的替代。可变形卷积网络其中标准卷积中的矩形窗口采样被移位点采样取代。Deformable DETR [24] 遵循这种方式，对某个查询相关的关键点进行采样，以进行可变形注意力。当图像被下采样为低分辨率图像以调整内容感知图像大小时，也会发生类似的做法。因此，根据上采样的本质，如图 6 所示，本文使用基于采样点的动态上采样的方案，而采样点的具体生成方式如图 7 所示，采样集(sampling set)是生成的偏移量和原始网格位置的总和。上面的框显示了带有“静态范围因子”的版本，其中偏移是用线性层生成的。底部描述了具有“动态范围因子”的版本。

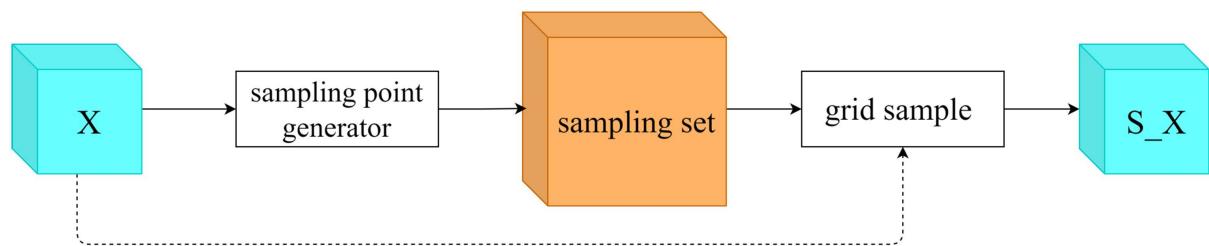


Figure 6. Dynamic up-sampling based on sampling points
图 6. 基于采样点的动态上采样

网格采样(Grid Sampling): 给定大小为 $C \times H_1 \times W_1$ 的特征图 X 和大小为 $2 \times H_2 \times W_2$ 的采样集 S ，其中第一维的 2 表示 x 和 y 坐标，网格采样函数使用 S 中的位置来重新将假设的双线性插值 X 采样为大小为 $C \times H_2 \times W_2$ 的 X' 。这个过程定义为：

$$X' = \text{grid sample}(X, S) \quad (1)$$

朴素实现(Naive Implementation): 给定上采样比例因子 s 和大小为 $C \times H \times W$ 的特征图 X , 使用输入和输出通道数为 C 和 $2s^2$ 的线性层来生成大小为 $2s^2 \times H \times W$ 的偏移量 O , 然后通过像素洗牌将其重塑为 $2 \times sH \times sW$ 。那么采样集合 S 就是偏移量 O 与原始采样网格 G 之和, 即:

$$O = \text{Linear}(X), S = G + O \quad (2)$$

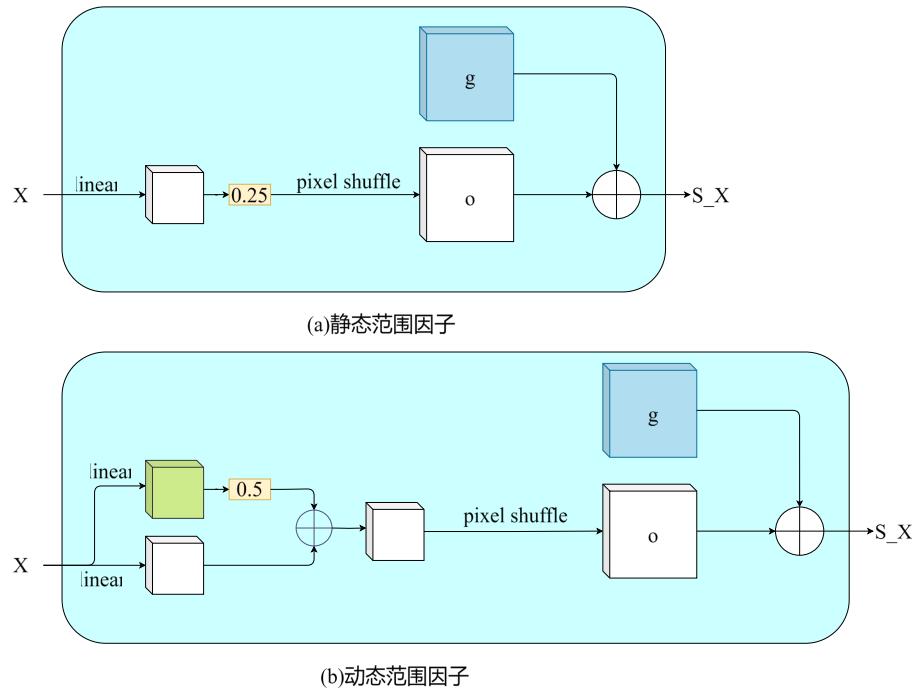


Figure 7. Sampling point generator in DySample
图 7. DySample 中的采样点生成器

3.4. BiFPN 模块

YOLOv8 的颈部网络采用路径聚合网络(PAN)和特征金字塔网络(FPN)架构的协同组合[25], 如图 8 所示。FPN 框架巧妙地将深层特征信息传递到较浅层, 从而通过关键的高级见解丰富它们。相反, PAN 架构有利于将精确位置数据从表层向上流动到更深层、特征丰富的底层。这种融合创造了 PANet 结构, 巧妙地融合了浅层和深层特征, 显著增强了模型识别最细微特征的能力。

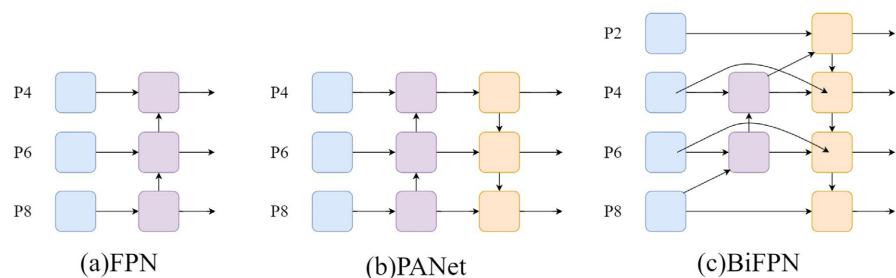


Figure 8. Neck feature network design
图 8. 颈部特征网络设计

然而 PaNet 的构造具有明显的不足。进入 PAN 的通路，之前由 FPN 处理，无意中被过滤掉一些最初从 YOLOv8 主干收集的典型特征信息。为了解决这个问题，本文创新性地集成了双向特征金字塔网络(BiFPN)进入 YOLOV8 模型，BiFPN 结构如图 8(c)所示。BiFPN 架构通过引入现有 FPN + PAN 框架的两个额外横向连接路径。这些新颖的路径巧妙地保存和合并直接从数据中提取的原始特征主干网络进入检测特征图[26]。

此外，在改进的模型中，灵活性地将 P2 层合并到模型的颈部网络中。这 P2 层以其扩展的特征图大小和最小的卷积操作为特征，并且伴有额外的检测头。这些增强功能有双重目的：它们不仅加强了模型内位置和特征信息的融合而且还显着提高了检测微小目标的精度。在图 8 中生动地展示了这些增强的结构。

4. 实验结果与分析

4.1. 数据集及缺陷种类

研究的晶体工艺加工熔接阶段图片数据集由成都中光睿华科技有限公司提供。如图 9 所示，搭建熔接工艺缺陷自动检测平台来采集缺陷图片，共采集有效图片 12000 张，原始图像分辨率为 4000×3000 的图片，将图片划分为单光圈、双光圈、凸点 3 种类型，每类图片有 4000 张，将图片分别以 7:2:1 的划分比例随机划分为训练集、验证集、测试集。



Figure 9. Neck feature network design
图 9. 晶体工艺设备

熔接阶段的细节图如图 10 所示，依次是单光圈、双光圈、有凸点(凸点数量为 1, 2)3 种类型的检测任务。

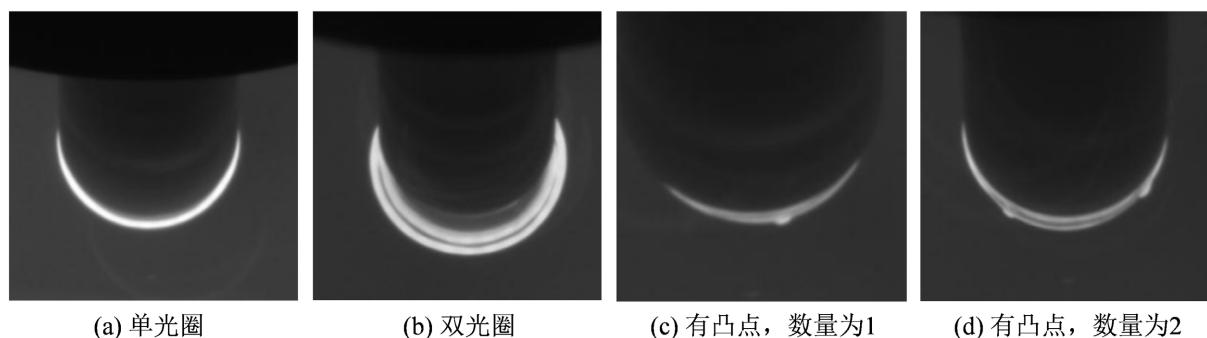


Figure 10. Type of defect
图 10. 缺陷种类

4.2. 实验环境及训练参数

1. 计算平台:

本文中的所有实验都使用同一台机器，实验中使用的硬件是 Intel(R) Xeon(R) Gold 5320 CPU @ 2.20 GHz 和 NVIDIA A100-PCIE-40GB GPU 来训练和测试模块。如表 1 所示：

Table 1. Experimental environment

表 1. 设备信息

Configuration	Environment
Operating System	Ubuntu 20.04.4
Accelerated environment	CUDA 11.4
Deep learning framework	PyTorch 1.13.1
Programming Language	Python 3.8.10

2. 参数设置:

本实验中的模型均使用相同的训练参数进行训练，如表 2 所示：

Table 2. Training parameters

表 2. 训练参数

Parameter	Value
Epochs	300
Batch	16
Image size	640 × 640
Workers	8
Learning rate	0.001
Optimizer	AdamW
NMS IoU	0.7
Weight-Decay	0.0005

4.3. 实施细节

4.3.1. 训练设置

训练过程中，各实验组采用一致的训练参数，保证实验的精度。输入分辨率配置为长边设置为 640 像素，保留图像的原始长宽比，批量大小固定为 16。在训练过程中，提前停止耐心设置为 50，即如果模型在 50 个 epoch 内没有表现出改进，则训练提前终止，epoch 设置为 300。损失函数的优化是通过利用 AdamW 下降算法实现的，其中动量值为 0.937，权重衰减系数为 5e-4。初始学习率设置为 0.001，置信度阈值定义为 0.4。采用马赛克数据增强，而所有其他参数与 YOLOv8 中的参数保持一致。

在推理过程中，使用了长边设置为 640 像素的标准化输入分辨率，同时保留了图像的原始长宽比。置信度精确定义为 0.001，交并集(IOU)阈值设定为 0.7。在速度测试中，实现了单一 GPU 利用率，并且批量大小被专门设置为 1，表示对单个图像进行顺序处理。

4.3.2. 评价指标

为了定量说明本文开发的算法的有效性，使用当前目标检测模型中常用的精度(P)、召回率(R)、平均精度(mAP)、参数量(Paremeters)评估指标来评估模型性能。比较了许多具有相同超参数的常用模型，以评估本文方法的性能。

精确度(P): 精确度衡量的是模型预测为正样本中有多少是真正的正样本。在目标检测中，精确度可以理解为模型预测的边界框中有多少是准确匹配到真实目标的。定义如下：

$$P = \frac{TP}{TP + FN} \quad (3)$$

召回率(R): 召回率衡量的是实际为正样本中有多少被模型正确预测为正样本。在目标检测中，召回率可以理解为所有真实目标中有多少被模型正确检测到。定义如下：

$$R = \frac{TP}{TP + FN} \quad (4)$$

其中 TP 对应于真阳性， FP 对应于假阳性， FN 对应于假阴性。

mAP 指标基于精确召回指标，该指标处理多个对象类别，并使用 IoU (交并集)定义正预测。它选择给定的 IoU 阈值，并计算该阈值在不同召回水平下获得的精度值的平均值。 IoU 是两个集合相似度的度量，是计算机视觉和图像处理中常用的一种度量，用于数值评估两个边界框(或两个区域)的重叠程度：

$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (5)$$

特定类别的 AP (平均精确度)是通过对模型的预测值和召回值进行排序，并计算出一条直线所围成的区域，该直线在直角坐标系中的纵轴代表精确度，横轴代表召回率。

$$AP = \sum_n (R_{n+1} - R_n) \cdot P_n \quad (6)$$

$mAP@0.5$ 是 IoU 临界值为 0.5 时的 AP 值。计算从 0.5 到 0.95 的每个 IoU 临界值(增量步长为 0.05)的 AP 值，并取平均值，得出 $mAP@0.5:0.95$ 。

$$mAP@0.5:0.95 = \frac{AP_{IoU=0.5} + AP_{IoU=0.55} + \dots + AP_{IoU=0.95}}{n} \quad (7)$$

n 的值为 10。通过使用 $mAP@0.5$ 和 $mAP@0.5:0.95$ ，本文评估了模型在各种 IoU 阈值下准确检测船舶的能力。此外，所提出模型的性能还通过其参数数量和 FLOPs 来描述，它们代表模型所需的计算量并衡量模型的复杂性。模型大小可以反映模型包含的参数数量。

4.4. 实验对比

4.4.1. 消融实验对比

Table 3. Training results of ablation experiments
表 3. 消融实验对比

Model	GSConv	DySample	P2	BiFPN	mAP@0.5	mAP@0.5:0.95
YOLOv8m					0.941	0.779
YOLOv8-A	√				0.951	0.78

续表

YOLOv8-A	√	√		0.95	0.781
YOLOv8-A	√	√	√	0.975	0.794
YOLOv8-A	√	√	√	√	0.982

为了更深入地研究增强四个不同组件对网络模型性能的影响，在此次研究中进行了五项实验。每个实验都涉及添加不同的模块，同时利用评价指标进行对比分析，结果如表 3 所示。从表的数据可以看出，集成 GSConv 模块后模型的精确率有了显着的提高。这一改进意味着 GSConv 模块在特征处理过程中优化了凸点特征的处理。Dysample 上采样层模块的引入增强了对熔接过程中产生的小凸点的识别，最大限度地减少漏检，从而提高召回率。加入 P2 检测头后显著提升了准确率，表明检测头的增加在光圈凸点识别中可以有效利用检测框的几何特性，从而提高目标检测的精度。从 YOLOv8-A 不同模块的效果对比可以看出集成了所有模块后的模型的 mAP 值在训练中具有更稳定的提升，并且对于精确率来说，YOLOv8-A 模型的目标检测效果也是最好。一般来说，不同模块的添加对模型性能的影响不同，但综合考虑，YOLOv8-A 模型取得了最佳的整体检测结果。

4.4.2. 与其他检测模型的对比实验

为了评估提出的方法在晶体加工熔接阶段目标检测上的性能，将本文方法和 Faster R-CNN、YOLOv3 [27]、YOLOv5、YOLOv6 [28] 等目前主流的目标检测算法进行比较，在凸点缺陷数据集上开展实验，将改进算法的检测结果通过 2 次实验取平均值，避免偶然性误差，增强实验结果的稳定性，通过精确率、召回率、mAP@0.5 等评价指标来判定改进的有效性。熔接阶段缺陷检测结果如表 4 所示。从不同的模型对比来看，在检测单光圈和双光圈方面，经典检测算法的效果与 YOLO 系列目标检测算法的效果大相径庭。在基于 YOLOv8 上提出了 YOLOv8-A。所提出的 YOLOv8-A 的 FPS 低于原始 YOLOv8m，但仍然比其他算法更快。更重要的是，YOLOv8-A 在类别为 dot 的凸点缺陷中 P、mAP 等参数方面优于其他方法，这证明本文的模型取得了令人满意的结果。特别是在凸点缺陷的准确率(P)上优于其他方法，意味着 YOLOv8-A 在熔接任务图像中的小目标检测方面具有优势。总体而言，这次比较再次证明对原始 YOLOv8 模型所做的修改是成功的，所提出的模型在检测精度和部署难度方面满足了实际生产场景的需求，具有相当的鲁棒性和实用性。

此外，如图 11 所示的不同模型训练过程中 mAP@0.5: 0.95 曲线图，其中不同颜色的曲线代表不同的模型。与其他模型相比，YOLOv8-A 模型在整个训练过程中在 mAP@0.5:0.95 方面始终保持优势。这表明 YOLOv8-A 模型在光圈凸点识别方面表现出改进的目标检测能力，从而获得更高的准确率。此外，YOLOv8-A 模型不仅表现出优越的性能，而且在训练过程中收敛速度更快。

Table 4. Effect of running each YOLO version on industrial datasets
表 4. 在工业数据集上各个 YOLO 版本的运行效果

Model	P			AP	R	mAP@0.5	mAP@0.5:0.95	Parameters	time(ms)
	dot	Single aperture	Double aperture						
Faster R-CNN	0.633	0.815	0.856	0.768	0.722	0.813	0.756	139.24M	16.8
YOLOv3m	0.364	0.753	0.99	0.702	0.713	0.706	0.6	103.69M	6.7
YOLOv5m	0.819	0.989	0.989	0.933	0.923	0.942	0.775	25.06M	12.9

续表

YOLOv6m	0.813	0.988	0.99	0.93	0.912	0.938	0.776	51.99M	4.5
YOLOv8s	0.786	0.985	0.9	0.89	0.766	0.763	0.754	11.13M	1.9
YOLOv8m	0.862	0.992	0.988	0.947	0.889	0.931	0.774	25.86M	3.8
YOLOv8l	0.822	0.993	0.991	0.935	0.915	0.933	0.779	43.61M	13.5
YOLOv8x	0.835	0.996	0.99	0.94	0.918	0.942	0.782	68.15M	23.2
YOLOv8-A	0.92	0.998	0.996	0.971	0.97	0.982	0.806	23.67M	3.7
YOLOv9	0.859	0.94	0.996	0.932	0.902	0.949	0.775	25.53M	9.5

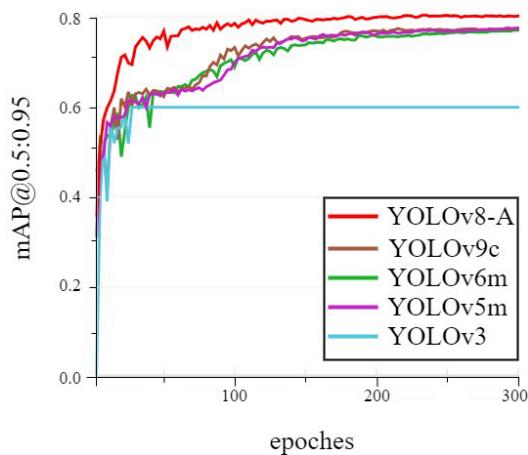


Figure 11. Training effects of different YOLO versions mAP@0.5:0.95 curve chart
图 11. 不同 YOLO 版本训练效果 mAP@0.5:0.95 曲线图

4.4.3. 模型改进效果对比

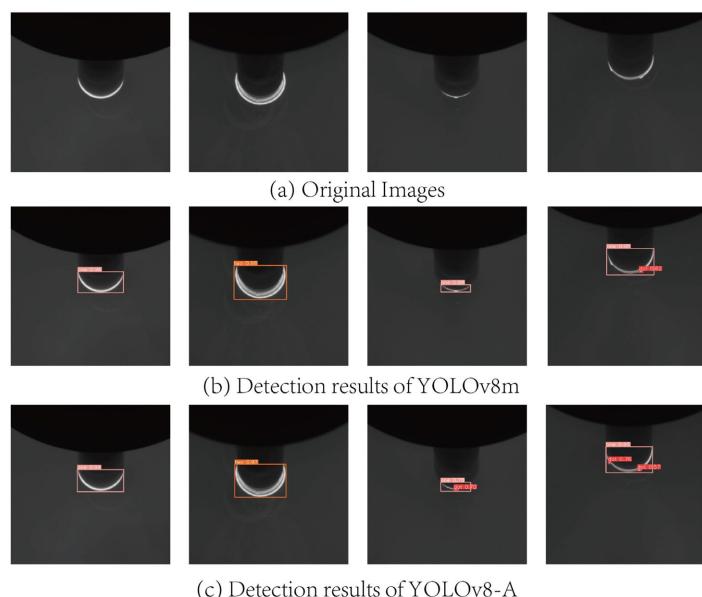


Figure 12. Comparison of model detection results
图 12. 工业数据集上 YOLOv8 和 YOLOv8-A 检测效果

为了直观地展示 YOLOv8-A 的检测性能, YOLOv8-A 和 YOLOv8m 在熔接任务光圈凸点(如图 12(a)所示)数据集上的部分检测结果如图 12(b)~图 12(c)所示。可以看出, YOLOv8m 表现出不同程度的漏检,然而, YOLOv8-A 弥补了这个问题。特别是,当出现多个小目标以及小目标凸点的起伏弧度很低时, YOLOv8-A 的凸点小目标预测更加准确。这进一步表明本文的修改可以增强原始 YOLO 的性能,这有助于提高引晶工艺熔接阶段中出现凸点目标提取的准确性。而在工业部署测试阶段,验证其有效降低漏检率和误检率的能力,已经满足工业生产要求。

5. 结论

本文提出了一种新的 YOLOv8 检测模型来支持光伏晶体生产引晶加工时对晶体工艺加工过程进行实时监控和分析,可以有效提高判断熔接程度的效率。

通过在主干网络中添加 BiFPN 模块,可以在利用全局信息学习的能力捕获复杂光圈图像中最具辨别力的区域的同时巧妙地保存和合并直接从数据中提取的原始特征主干网络进入检测特征图,从而得到上下文信息理解,使模型聚焦于目标而不是背景,提高熔接阶段情况下光圈凸点目标检测模型的可靠性。YOLOv8-A 重新设计了颈部网络,添加了 P2 检测头使其拥有更高的分辨率,使得模型能够更好地捕捉到小尺寸目标的细节,增强模型准确获取凸点小目标特征的能力。采用 GSConv 以及使用轻量级上采样模块 Dysample,从而显著降低了模型大小和计算复杂度为了使网络轻量化并保证检测有效性。改进后的模型与原始模型相比将 mAP@0.5 和 mAP@0.5:0.95 分别提高了 5.48% 和 4.1%,参数量减少了 8.47%,实现了检测性能的提高和计算成本的降低。测试结果表明,该方法在检测精度方面优于几种最先进的目标检测模型,能够满足引晶熔接阶段的光圈图像凸点识别任务中可靠、准确、快速的目标检测要求,并且能够部署在嵌入式终端。

参考文献

- [1] 王正省,任永生,马文会,等.直拉法单晶硅生长原理、工艺及展望[J].材料导报,2024,38(9): 5-17.
- [2] 马亚萍.直拉法单晶硅生长原理及工艺分析[J].中国粉体工业,2023(1): 29-32.
- [3] 刘赟,薛忠营,魏星,李炜.直拉单晶硅晶体缺陷研究进展[J].微纳电子与智能制造,2022,4(1): 64-74.
- [4] Soori, M., Arezoo, B. and Dastres, R. (2023) Artificial Intelligence, Machine Learning and Deep Learning in Advanced Robotics, a Review. *Cognitive Robotics*, **3**, 54-70. <https://doi.org/10.1016/j.cogr.2023.04.001>
- [5] Yang, L. et al. (2023) An Improving Faster-RCNN with Multi-Attention ResNet for Small Target Detection in Intelligent Autonomous Transport with 6G. *IEEE Transactions on Intelligent Transportation Systems*, **24**, 7717-7725. <https://doi.org/10.1109/TITS.2022.3193909>
- [6] 邵延华,张铎,楚红雨,等.基于深度学习的 YOLO 目标检测综述[J].电子与信息学报,2022,44(10): 3697-3708.
- [7] Li, H., Li, J., Wei, H., Liu, Z., Zhan, Z. and Ren, Q. (2022) Slim-Neck by GSConv: A Better Design Paradigm of Detector Architectures for Autonomous Vehicles. arXiv: 2206.02424.
- [8] Zhang, Q., Zhang, H. and Lu, X. (2022) Adaptive Feature Fusion for Small Object Detection. *Applied Sciences*, **12**, Article 11854. <https://doi.org/10.3390/app122211854>
- [9] 张姝瑾,许兴时,邓洪兴,等.基于 YOLOv8n-seg-FCA-BiFPN 的奶牛身体分割方法[J].农业机械学报,2024,55(3): 282-289+391.
- [10] Zhang, M., Wang, Z., Song, W., Zhao, D. and Zhao, H. (2024) Efficient Small-Object Detection in Underwater Images Using the Enhanced YOLOv8 Network. *Applied Sciences*, **14**, Article 1095. <https://doi.org/10.3390/app14031095>
- [11] 赵其杰,程德富,卢建霞,等.基于视觉的单晶生长直径检测与监控[J].制造业自动化,2011,33(7): 22-25.
- [12] 赵跃,王欣.基于光圈模式识别的直拉晶体生长温度测量技术研究[J].传感技术学报,2018,31(4): 573-578.
- [13] 杨舟,程莹,张诗婧,等.基于改进 YOLOv5 算法的直拉法单晶硅位错检测模型研究[J].应用光学,2023,44(5): 1022-1029.
- [14] Wang, H., Qian, H., Feng, S. and Wang, W. (2024) L-SSD: Lightweight SSD Target Detection Based on Depth-

- Separable Convolution. *Journal of Real-Time Image Processing*, **21**, Article No. 33. <https://doi.org/10.1007/s11554-024-01413-z>
- [15] Luo, F., Dai, Y., Fuentes, J., Ding, W. and Zhang, X. (2024) M-DETR: Multi-Scale DETR for Optical Music Recognition. *Expert Systems with Applications*, **249**, Article 123664. <https://doi.org/10.1016/j.eswa.2024.123664>
- [16] Liu, B., Wang, H., Wang, Y., Zhou, C. and Cai, L. (2023) Lane Line Type Recognition Based on Improved YOLOv5. *Applied Sciences*, **13**, Article 10537. <https://doi.org/10.3390/app131810537>
- [17] Yang, Z., Xie, R., Liu, L. and Li, N. (2024) Dense-YOLOv7: Improved Real-Time Insulator Detection Framework Based on YOLOv7. *International Journal of Low-Carbon Technologies*, **19**, 157-170. <https://doi.org/10.1093/ijlct/ctad122>
- [18] Wang, X. and Liu, J. (2024) Vegetable Disease Detection Using an Improved YOLOv8 Algorithm in the Greenhouse Plant Environment. *Scientific Reports*, **14**, Article No. 4261. <https://doi.org/10.1038/s41598-024-54540-9>
- [19] Yu, X., Yu, Q., Mu, Q., Hu, Z. and Xie, J. (2023) MCAW-YOLO: An Efficient Detection Model for Ceramic Tile Surface Defects. *Applied Sciences*, **13**, Article 12057. <https://doi.org/10.3390/app132112057>
- [20] Zhao, X. and Song, Y. (2023) Improved Ship Detection with YOLOv8 Enhanced with MobileViT and GSConv. *Electronics*, **12**, Article 4666. <https://doi.org/10.3390/electronics12224666>
- [21] Rybczak, M. and Kozakiewicz, K. (2024) Deep Machine Learning of MobileNet, Efficient, and Inception Models. *Algorithms*, **17**, Article 96. <https://doi.org/10.3390/a17030096>
- [22] Ma, N., Zhang, X., Zheng, H. and Sun, J. (2018) Shufflenet V2: Practical Guidelines for Efficient CNN Architecture Design. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 122-138. https://doi.org/10.1007/978-3-030-01264-9_8
- [23] Han, K., Wang, Y., Tian, Q., Guo, J., Xu, C. and Xu, C. (2020) GhostNet: More Features from Cheap Operations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 1580-1589.
- [24] Han, W., He, N., Wang, X., Sun, F. and Liu, S. (2023) IDPD: Improved Deformable-DETR for Crowd Pedestrian Detection. *Signal, Image and Video Processing*, **18**, 2243-2253. <https://doi.org/10.1007/s11760-023-02896-2>
- [25] Li, N., Ye, T., Zhou, Z., Gao, C. and Zhang, P. (2024) Enhanced YOLOv8 with BiFPN-SimAM for Precise Defect Detection in Miniature Capacitors. *Applied Sciences*, **14**, Article 429. <https://doi.org/10.3390/app14010429>
- [26] Chiley, V., Thangarasa, V., Gupta, A., Samar, A., Hestness, J. and DeCoste, D. (2023) RevBiFPN: The Fully Reversible Bidirectional Feature Pyramid Network. arXiv:2206.14098.
- [27] Chen, W.-Y., Zhao, H.-C., Liu, P., Fang, J. and Sun, H. (2024) Vehicle Detection Algorithm Based on Improved YOLOv3. *Control and Decision*, **39**, 1151-1159.
- [28] Wang, J., Li, Q., Fang, Z., Zhou, X., Tang, Z., Han, Y., et al. (2023) YOLOv6-ESG: A Lightweight Seafood Detection Method. *Journal of Marine Science and Engineering*, **11**, Article 1623. <https://doi.org/10.3390/jmse11081623>