

新型 PIS-YOLO 模型下的 X 射线注塑缺陷 检测方法研究^{*}

林子涵 张巧芬 王桂棠

(广东工业大学机电工程学院 广州 510006)

摘要:为了提高深度学习在 X 射线注塑工件缺陷检测中的准确性,实现更高精度的无损检测,提出了一种改进的 YOLOv8-seg 注塑件内部缺陷实例分割模型 PIS-YOLO。首先为了减少参数量并提高特征融合能力,在主干网络中设计了一个多尺度特征融合与通道数减小的 HG-Net 模块,取代传统 C2f 模块;进一步引入倒置残差移动块融合高效多尺度注意力(iRMB_EMA)模块增强了深层传递,经由简化冗余连接的路径聚合网络-特征金字塔网络(PAN-FPN)完成特征融合。同时增加一个额外的输出分割检测头捕捉细小缺陷,提高了模型对小目标缺陷及缺陷边缘的精确识别。在注塑工业零件自制数据集上,主干网络部分提出的 HG-Net 模块对比 C2f 模块实现了同架构下计算量减小 22.03%,在此基础上,结合 iRMB_EMA 注意力融合模块与额外输出检测头设计的模型整体的准确率相较于基准模型分别提升了 2.9% 和 5.7%,且模型更轻量,计算复杂度更低。

关键词:X 射线;注塑工件缺陷检测;YOLOv8-Seg;注意力机制;多尺度特征融合

中图分类号: TP391; TN41 文献标识码: A 国家标准学科分类代码: 520.6040

Research on improved automatic defect detection method of X-ray injection parts under PIS-YOLO model

Lin Zihan Zhang Qiaofen Wang Guitang

(School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: To improve the accuracy of deep learning in X-ray injection molding workpiece defect detection and realize higher precision nondestructive testing, an improved YOLOv8-seg internal defect segmentation model PIS-YOLO was proposed in this paper. Firstly, to reduce the number of parameters and improve the feature fusion capability, a multi-scale feature fusion and channel number reduction HG-Net module is designed in the backbone network to replace the traditional C2f module. The iRMB_EMA attentional fusion module is further introduced to enhance the deep transmission, and the feature fusion is completed by PAN-FPN with simplified redundant connections. Meanwhile, an additional output segmentation detection head is added to capture small defects, which improves the model's accurate recognition of small target defects and defect edges. On the self-made data set of injection molding industrial parts, HG-Net module proposed in the backbone network section achieves a 22.03% reduction in computation under the same architecture compared with C2f module. On this basis, the overall precision of the model combined with the iRMB_EMA attention fusion module and additional output detection head is improved by 2.9% and 5.7%, respectively, compared with the benchmark model, and the model is lighter and less complex.

Keywords: X-rays; injection parts damage detection; YOLOv8-seg; attention mechanism; multi-scale feature fusion

0 引言

在现代工业制造中,X 射线缺陷检测技术已成为保

证产品质量和安全性的重要手段。由于 X 射线能够穿透材料并提供高分辨率的内部结构信息,该技术被广泛应用于金属铸件、电子元件及复合材料的无损检测(NDT),如在自动化集成电路封装缺陷检测系统中用来提高无损

检测技术的检测精度和检测效率^[1-3],在航空航天板壳结构等复合材料无损检测应用场景对其内部损伤与缺陷进行有效的无损检测与评价^[4-5]等。

注塑工业零件作为制造业的关键支撑领域,在全球工业体系中占据着举足轻重的地位,随着材料科学与智能制造技术的不断进步,对注塑工业零件生产的智能化、精细化要求日益提升。注塑零件的质量检测作为确保产品质量与生产效率的重要环节,吸引了众多学者关注。尽管基于机器视觉和深度学习的目标检测技术已被广泛研究和应用,但在注塑产品领域的应用仍相对有限。针对注塑件缺陷特征简单、类型单一的特点,采用基于机器视觉的方法实现对缺陷的自动检测,以满足实际检测需求。王元才等^[6]基于图像处理技术,结合自动控制与视觉算法,首次构建并实现了集成光学组件、机械结构、电子设计与检测算法于一体的注塑产品杂质与焦料缺陷检测系统。为应对注塑制品图像中常见的光照分布不均问题,杜娟等^[7]提出了一种基于行像素灰度校正的图像增强算法,并借助二次定位的阈值分割技术,对柱形工件中的异常区域进行了精准识别与分割。Liu 等^[8]针对注塑产品多样性、小尺寸和复杂背景的挑战,提出了一种注塑产品缺陷检测模型 IMP-DETR,适合在实际场景中应用。Hu 等^[9]创新性地升级了传统的 VGG16 网络,以便对不合格的注塑部件进行进一步分类,改进的 VGG16 网络的分类准确率达到 96.67%,优于传统网络的 53.33%。

近年来,基于深度学习的图像处理技术发展迅速,广泛应用于计算机视觉各个领域。然而,由于注塑件缺陷在类型和形态上的多样性,给自动化检测技术带来了更高的挑战。卷积神经网络(convolutional neural network, CNN)作为深度学习的重要模型,已广泛应用于图像识别、目标检测及分割等任务^[10]。当前主流的目标检测方法主要分为双阶段检测模型和单阶段检测模型两类。前者包括如 SSD^[11-12]、Faster-RCNN^[13-14]等,具备较高的检测精度;而后者以 YOLO 系列算法为代表^[15-17],在实际应用中仍最常用。

虽然现有方法在推进注塑工业缺陷检测领域从传统人工转向机器智能方面取得了进展,但大多数方法仍然采用目标检测来捕捉缺陷,且精度一般。为了应对未来的需求,注塑工业缺陷检测任务需要在目标检测基础上引入分割任务,以更全面地分析注塑工业缺陷情况。除了分析缺陷数量外,还需要考虑缺陷面积、占比等更丰富的信息。因此,针对注塑工业缺陷检测的实例分割任务研究成为该领域的趋势方向。

因此,基于实例分割任务的特性,论文创新性地提出一种新型的 PIS-YOLO (precise instance segmentation for YOLO) 的缺陷检测模型,以模型更加轻量化、解决传统模型在缺陷检测任务中精度低为出发点,从而实现更快速、

高效、更适合实例分割检测任务的目标。

1 PIS-YOLO 实例分割模型优化设计

本文在 YOLOv8 模型的基础上进行改进的 PIS-YOLO 实例分割模型优化设计,并将其应用到注塑件内部缺陷检测中。首先通过引入轻量化的 HG-Net 模块替代 YOLOv8 中的 C2f 结构,从而有效降低模型复杂度并提升检测速度。随后,在特征融合阶段引入倒置残差移动块融合高效多尺度注意力(iRMB_EMA)模块,以增强模型对缺陷特征的提取能力,从而提升检测精度。最后,对检测头进行优化设计,采用更少的通道数和卷积层设计解耦头,以进一步提高检测精度并减少模型的计算开销。

1.1 主干网络的 HG-Net 模块设计

本文提出的 PIS-YOLO 模型的 Backbone 主干层 HG-Net 由 Conv、HG-Block 和 SPPF 组成,Conv 在处理过程中逐级减小特征图的空间分辨率,同时提取不同层级的语义信息,有助于模型感知多尺度的目标特征;HG-Block 通过部分跨阶段连接和残差结构,有效减少计算量和参数规模,同时提升特征表达能力;SPPF 通过多尺度池化操作,增强网络的多尺度特征表达能力,更好地应对不同大小的目标物体。

HG-Net 模块的整体结构如图 1(a)所示。HG-Net 是一种结合了多层次、多尺度特征提取与增强的神经网络架构,其核心在于 HG-Block。在主干网络的优化设计中,主要利用通道数减少的思想创新性地对 HG-Net 的 HG-Block 模块进行了改进, HG-Block 模块的整体结构如图 1(b)所示。在模块特征提取过程中,通过堆叠多个 3×3 卷积层来构建丰富的特征图,并利用不同尺度的卷积操作(如 3×3 、 5×5 、 7×7)来捕获图像的多样化特征。这些多尺度特征图随后被融合成一个大尺寸的特征图,并通过 1×1 卷积层进行通道整合或降维。为了进一步提升特征表示能力,引入了轻量级通道注意力模块(enhanced squeeze excitation, ESE)来增强特定通道的特征响应。最终,通过残差连接将 ESE 模块的输出与初始输入特征图相加,以缓解梯度消失问题并促进更好的训练效果,通过部分卷积和跨阶段连接, HG-Block 模块减少冗余计算,同时增强特征表示的能力,这对于轻量化模型特别重要,能够在保持高精度的同时显著降低计算量和参数量。

与 C2f 模块先通过卷积变换通道维度、再通过 Split 方式保留部分特征参与后续处理的机制不同, HG-Block 在中间阶段的通道数量与最终输出一致,无需执行额外的通道调整操作。HG-Block 通过在增加堆叠层数的同时,将大量特征信息压缩至较少的维度,从而在不降低模

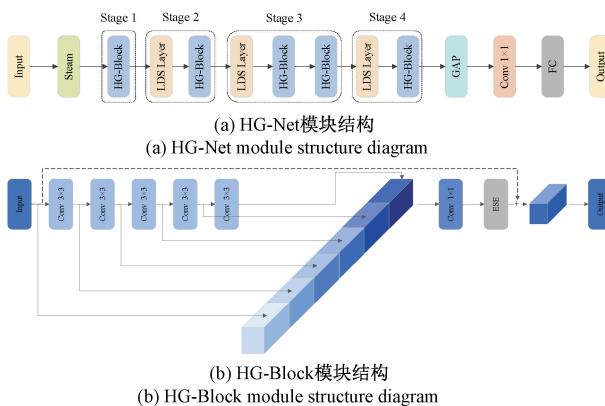


图 1 HG-Net 模块结构和 HG-Block 模块结构

Fig. 1 HG-Net module structure and HG-Block module structure

型性能的前提下有效减轻了计算复杂度。相比于 C2f 模块多次调用普通卷积操作, HG-Block 在整个过程中仅使用一次普通卷积, 并且充分的多维连接使得特征提取能力得以有效维持。因此, HG-Block 不仅在特征提取方面具备更高效率, 还在边缘设备上拥有更好的计算性能。

在 PIS-YOLO 网络结构中, HG-Net 通过逐级堆叠 HG-Block, 结合全局平均池化和全连接层, 实现了对输入数据的精确分类或回归任务, 充分展示了其在处理复杂图像数据时的强大能力, 这对于快速准确的对象识别至

关重要。

1.2 注意力机制 iRMB_EMA 模块设计

凭借结构灵活、易于集成以及良好的性能表现, 注意力机制在众多先进的网络架构中得到了广泛应用。如通道注意力^[18] (squeeze-and-excitation, SE)、轻量级注意力^[19] (spatial attention, SA)、坐标注意力^[20] (coordinate attention, CA), 以及多头自注意力^[21] (multi-head self-attention, MHSA), 同时, 融合空间注意力和通道注意力的复合机制 (convolutional block attention module, CBAM) 等方法也引起了研究者的持续关注与深入探索。

虽然上述注意力机制在引导网络关注关键特征方面表现出色, 但 SE 仅针对通道注意力, 未能兼顾空间注意力。而像 BAM 与 CBAM 这类结构虽然融合了通道与空间的感知能力, 但它们未能有效整合多尺度的空间信息, 限制了特征表达的丰富性, 并且在处理远距离依赖关系方面存在不足, 而这对于实例分割任务至关重要。为此, 本文设计了一种 EMA 模块, 在保证较少模型参数量的同时, 增强了模型的高通道分辨率和高空间分辨率。

EMA 模块通过构建空间与通道维度之间的信息交互关系, 有效增强了图像特征的表达能力。其核心在于对整体与局部特征间的关联程度进行精准建模, 从而实现对远距离与近邻区域依赖关系的同时感知, 其结构如图 2 所示。

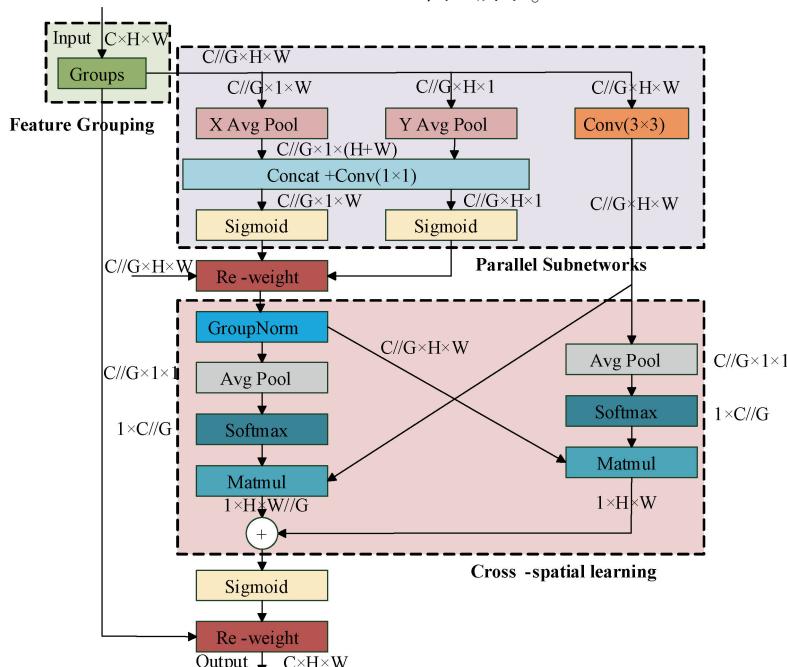


图 2 EMA 模块结构

Fig. 2 EMA module structure diagram

EMA 单元首先将输入张量 $x \in \mathbb{R}^{C \times H \times W}$ 沿通路方向划分为若干子集, 从而便于捕捉多样的语义特征。划分

方式可记为 $x = [x_0, x_1, \dots, x_{G-1}]$, 其中每个子张量 $x_i \in \mathbb{R}^{C \times H \times W}$, 且分组总数 $G \ll C$ 。为提取各子集的关注

权重,EMA 构建了 3 个并行的分支模块,其中两个位于 1×1 卷积路径上,另一个则采用 3×3 卷积进行局部感知建模。在 1×1 路径中,采用了一种一维的全局平均池化方法,分别沿水平与垂直进行投影编码,以增强通道内的信息表达能力,该过程的具体计算如式(1)、(2)所示。

$$Z_c^H(H) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(H, i) \quad (1)$$

$$Z_c^W(H) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, W) \quad (2)$$

式中: C 表示输入通道数; H, W 分别表示输入特征的空间维度; x_c 表示第 c 通道的输入特征。

接着,将上述两个方向上的编码结果在垂直方向上进行拼接(Concat),并应用相同的 1×1 卷积处理,且该卷积操作保持原始通路数量不变。卷积输出随后被拆分为两个独立向量,并分别输入到两个 Sigmoid 激活单元中,用以逼近二维卷积特征中的二项分布特性。每一组内的两个通路注意图通过逐元素乘积进行融合,从而实现两条 1×1 支路间差异化的跨维度信息交互。而后利用二维全局平均操作对该路径输出的整体空间特征进行编码,二维全局平均池化操作如式(3)所示。

$$Z_c = \frac{1}{H \times W} \sum_j^H \sum_i^W x_c(i, j) \quad (3)$$

该计算旨在整合整幅图像的空间上下文信息,同时捕捉长距离区域之间的依赖关系。其中 C 为输入特征图的通道数, H, W 分别表示输入特征的空间维度, $x_c(i, j)$ 表示第 c 通道在位置 (i, j) 处的输入特征。

接下来,最小尺度路径的输出会先进行形状调整,具体为将其转换为一个尺寸为 $R_1^{1 \times C/G}$ 与 $R_1^{C/G \times H \times W}$ 的矩阵相乘操作,以便后续进行特征融合。在此基础上,利用一个基于二维高斯图的 Softmax 非线性函数,来对前面的线性映射结果进行更精细的拟合。同时,对于包含 3×3 感受野的分支,采用带权卷积的方式来捕获各通道间的局部联系,从而提取更丰富的尺度层次信息。

通过矩阵点积计算,对各个并行分支的特征输出执行逐元素乘法操作,得到第 1 个空间感知权重图,该图能够有效整合当前阶段内的多尺度信息。接着,在具备 3×3 感受范围的支路中,借助二维全局平均池化方法,对整体空间范围的特征进行编码处理。同时,在激活融合操作展开前,需将 1×1 卷积分支的输出调整为相应的维度,即形状为 $R_3^{1 \times C/G} \times R_1^{C/G \times H \times W}$ 的矩阵格式。通过上述操作,可以构建出第 2 个空间感知图,其具备对空间信息的完整保留能力。最终,将每个组中的特征响应映射为两个对应的注意权重集合,并通过归一化 Softmax 机制建模像素之间的关系,以增强图像中各像素的全局上下文联系。

倒置残差移动块(inverted residual mobile block,iRMB)将卷积神经网络擅长的局部特征捕获与

Transformer 在长距离依赖建模上的灵活性相结合,相较于传统的卷积结构,该模块能够在较深层中更有效地融合整体语义信息,同时保持模型的高效性与轻量化特征表达能力。改进后的 iRMB 结构如图 3 所示。

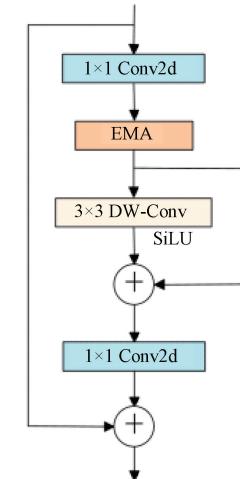


图 3 iRMB_EMA 模块结构

Fig. 3 iRMB_EMA module structure diagram

1.3 轻量级解耦头设计

YOLOv8-seg 模型在解耦头部结构上分别捕捉目标的位置信息与类别特征,由于回归与分类的任务侧重点不同,采用独立分支进行特征学习有助于提升整体检测精度。但与此同时,该方法也带来了一定的额外计算开销。从卷积算子的执行原理来看,较大的卷积窗口虽然有助于平滑特征图,但同时也显著增加了运算量。因此,预测阶段的主要计算压力集中在解耦头中的 3×3 卷积操作上,尤其在特征维度和输入分辨率提升的情况下,推理复杂度显著增加。为了解决这一问题,提出了一种改进后的轻量解耦头设计,采用更小的通道维度和更浅的卷积结构来降低计算成本,如图 4 所示。

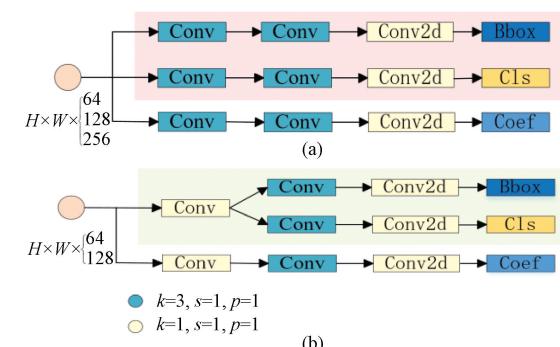


图 4 原始解耦头(a)和改进后的解耦头(b)

Fig. 4 Original decoupling head (a) and improved decoupling head (b)

首先利用一次 1×1 的卷积操作将通道数减少至 64,然后分别传递至独立的分类与回归分支进行特征处理。对于生成掩码权重的分支,其结构设计与分类部分相似,将原先连续两个 3×3 卷积替换为由 1×1 和 3×3 卷积组成的组合方式,最后通过分类结构预测掩码相关系数。该改进方法能够有效缓解分类与回归任务关注点不同所带来的干扰,同时显著降低输出检测头的参数规模与计算开销。

1.4 优化后的 PIS-YOLO 网络模型

综上,优化后的 PIS-YOLO 网络模型如图 5 所示,其主要由主干特征提取网络、特征融合网络和实例分割网络 3 个部分组成。图像输入阶段,首先将注塑件内部缺

陷图调整为 640×640 分辨率,并划分为 3 个通道作为输入。如图 5(a)所示,主干网络部分首先使用一组基础卷积 Conv 模块(卷积核大小为 k 、步长为 s)对输入图像进行一次二倍压缩,以减少计算开销。随后,通过 5 组 HG-Block 模块逐步提取缺陷区域在 4 种不同尺寸下的表征特征;最终,采用 SPPF 模块对最小特征图进行多维度信息整合。在特征融合阶段(图 5(b)),借助类似于 PAN-FPN 的结构,对 3 组中间层特征进行跨尺度整合,同时引入 iRMB_EMA 注意力机制以增强关键缺陷区域的表达能力。如图 5(c)所示,在实例分割部分,采用解耦检测头结构,以减少分类与定位损失之间的相互影响,从而提高分割精度和稳定性。

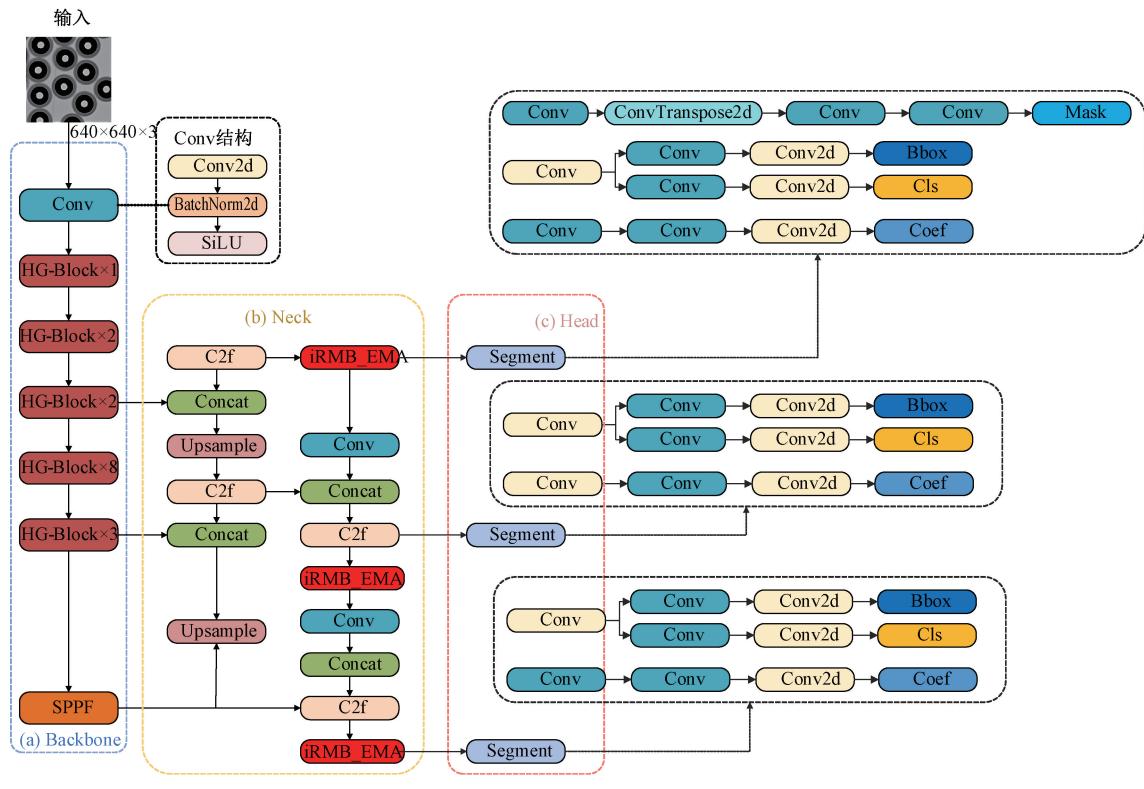


图 5 PIS-YOLO 结构

Fig. 5 PIS-YOLO structure diagram

2 数据集准备与试验评价

为了验证优化后的 PIS-YOLO 网络模型的有效性,论文进一步将其应用在 X 射线注塑件缺陷检测中。论文数据集由项目合作单位广东欧派克家居智能科技有限公司提供,通过实际 DR 成像技术共收集注塑工件 X 射线图像 550 张作为注塑件内部缺陷原始数据集,再通过 3 种数据增强方法随机缩放(random resize)^[22]、图像融合(Mixup)^[23]、随机拼图(Mosaic)^[24]获得图像共 1 770 张,包含缺陷样例 3 665 个。其中含缺陷数据集含 I 级气

孔缺陷 752 张,II 级气孔缺陷 1 541 张,III 级气孔缺陷 874 张,IV 气孔级缺陷 497 张。缺陷采用开源软件 labelImg 标记,获得 txt 格式标记文件。对 1 770 张缺陷数据集按照 7:3 比例,随机分训练集、验证集各 1 239 张、531 张。

为评估本文所提出模型的性能表现,需搭建相应的实验环境,该环境由硬件系统与软件系统两部分组成。其中,硬件部分依托于实验室配备的深度学习工作站,具备良好的并行处理能力。相关的软硬件配置信息详见表 1 所示。

表1 实验配置信息

Table 1 Experimental configuration information

实验环境	版本型号
操作系统	Ubuntu18.04(64 bits)
编程语言	Python 3.9
CUDA	12.1
Pytorch	2.1.0
编程软件	Pycharm Community
显卡	GeForce GTX 3080Ti
显存	16 G
处理器	(CPU) Intel Core i7-9700K
内存条	32G DDR4 3200 MHz

论文采用召回率(recall, R)、精确率(precision, P)浮点运算量(FLOPS)、帧数每秒(FPS)以及Segment平均精度均值(mAP)来评价模型的识别精度和计算复杂度。

Segment mAP@0.5表示当预测掩码和真实标注掩码之间的交集超过0.5时,每个类别的平均精度(AP),其计算公式如式(4)所示。

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (4)$$

精确率是指正确识别图像的比例,其计算公式如式(5)所示。

$$P = \frac{TP}{TP + FP} \quad (5)$$

召回率表示正确识别阳性物体的比例,其计算公式

表2 5种注意力机制对比效果

Table 2 Comparative effects of five attention mechanisms

算法	Seg mAP@0.5/%	Params/(×10 ⁶)	计算量/GFLOPs	帧率/fps
YOLOv8-seg	82.4	3.51	12.0	80.7
YOLOv8-seg+EMA	82.6	2.69	10.5	66.1
YOLOv8-seg+CBAM	82.9	2.58	10.3	42.0
YOLOv8-seg+SE	82.0	2.52	10.3	68.8
YOLOv8-seg+CA	83.1	3.52	12.1	52.7
YOLOv8-seg+iRMB_EMA	83.5	2.70	10.5	71.4

从实验结果来看,iRMB_EMA相比于CBAM,在增加少量参数量的条件下,Seg mAP@0.5提高0.7%,检测速度提高29 fps。相比于EMA,在增加少量参数量的代价下,Seg mAP@0.5提升1.1%,检测速度提高5 fps。表明构建的注意力机制iRMB_EMA有更好的检测精度和实时检测性能,证明了iRMB_EMA模块的有效性。

表3 改进模型消融实验结果

Table 3 Ablation results of the improved model

模型	P/%	R/%	Seg mAP@0.5/%	帧率/fps	计算量/GFLOPs	Params/(×10 ⁶)
YOLOv8-seg	86.3	76.7	82.4	80.7	12.0	3.51
YOLOv8-seg+HG-Net	84.3	74.8	81.0	71.9	12.2	3.26
YOLOv8-seg+iRMB_EMA	85.0	77.4	83.5	71.4	10.5	2.70
YOLOv8-seg+HG-Net+解耦检测头	87.3	78.7	85.6	78.0	9.8	2.54
本文	89.4	78.9	86.1	88.3	9.4	2.40

如(6)所示。

$$R = \frac{TP}{TP + FN} \quad (6)$$

式中:TP为真正例,代表模型成功判断为目标类别的样本数;FP为假正例,表示模型误将非目标类别样本判定为目标类别的数量;FN为假负例,则是指模型未能识别出的实际属于目标类别的样本数量。AP是根据精确率-召回率曲线下的积分值所得,其值越高,说明检测效果越好。mAP是所有类别的AP的均值,用于综合评估模型在整个数据集上的检测能力。

浮点计算总量是衡量网络计算开销和复杂度的重要指标。

3 实验分析

3.1 多种注意力机制对比实验

为验证构建iRMB_EMA注意力机制的改进算法对注塑工件内部气孔缺陷的检测效果,本文选取了对于缺陷检测有较好效果的5种注意力机制进行比较,包括EMA、CBAM、SE和CA,使用这5种注意力机制在自制数据集上进行性能对比实验,实验环境和参数设置均保持一致,对比结果如表2所示。

3.2 改进点消融实验

为验证本文所提出的每个改进模块的有效性,设计了5组消融实验,使用相同的实验环境和数据集进行训练和测试,以确保可比性。现对比每个添加模块对YOLOv8-seg性能的影响,实验结果如表3所示。

由表 3 可知,原始 YOLOv8-seg 模型的准确率、召回率分别为 86.3% 和 76.7%,体现实例分割精度的 Segment mAP@0.5 可达到 82.4%,帧数为 80.7 fps,浮点运算量为 12.0 GFLOPs ,模型参数量为 3.51×10^6 ;添加轻量级主干网络 HG-Net,虽然模型的准确率、召回率、Segment mAP@0.5 等指标都有所下降,但可显著降低模型的权重,降低了模型 7.2% 的参数量,其原因在于 HG-Net 可能减少了深层网络结构的通道数或者使用了更少的卷积核,导致模型捕获的特征信息量下降,使检测能力有所削弱,但是参数量的减少意味着模型在嵌入式设备或计算资源有限的环境下更具优势,能够显著降低计算资源消耗。此外,在模型的特征融合部分加入 iRMB_EMA 注意力机制后,模型的识别准确率提升 0.8%、召回率提升了 3.5%、Segment mAP@0.5 提升 3.1%、帧数每秒降低了 0.5 fps,计算量降低了 13.9%,表明在模型中引入 iRMB_EMA 注意力机制后,在保证低参数量的同时保证模型的高通道分辨率和高空间分辨率,进而增强模型对关键特征的提取能力,使模型在对重要信息、多尺度目标的检测能力上得到大幅提升。最后,加上对解耦头的改进,结果表明模型的识别准确率提升 2.7%、召回率提升了 1.7%、Segment mAP@0.5 提升 2.5%,帧率提升 9.3%、计算量降低 6.7%、参数量降低 5.9%,可见对模型解耦头的改进,可有效降低模型头部的参数量,减少参数冗余,提升类别和边界框回归的协同能力,从而降低模型浮点运算量,图片处理速度大幅提升。

最后,相对于原始 YOLOv8-seg 模型,PIS-YOLO 在各项指标上的提升分别为准确率提升 3.6%、召回率提升 2.9%、Segment mAP@0.5 提升 4.5%、帧率提升了 9.4%、浮点运算量降低 21.7%、参数量降低 31.6%。由此可得出,通过对 YOLOv8-seg 模型的改进,各个改进模块对模型各项评价指标的提升都起到了一定的作用,改进之后的 YOLOv8-seg 不仅可实时完成对目标的分割,在多尺度目标识别上效果显著,达到了轻量级网络设计要求。

为了验证改进算法 PIS-YOLO 对于不同大小目标气孔缺陷的检测效果,对比原始 YOLOv8-seg 和 PIS-YOLO 在注塑工件内部的各类型缺陷检测上的实例分割任务平均精度,对比结果如图 6 所示。

由图 6 可知,PIS-YOLO 在各个不同大小目标缺陷中均有良好的检测精度,其中 I 级气孔的检测平均精度提高 2.6%,II 级气孔的检测平均精度提高 1.2%,III 级气孔的检测平均精度提高 8.3%,IV 级气孔的检测平均精度提高 3.1%。

3.3 对比实验

依据实现流程的差异,缺陷检测算法可被划分为单阶段目标检测算法(仅需单次特征提取便能完成目标检

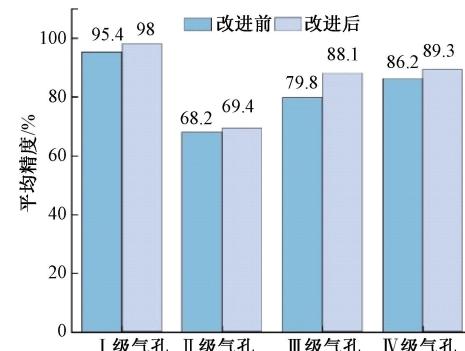


图 6 改进前后不同大小目标缺陷检测的平均精度对比

Fig. 6 Comparison of average accuracy of defect detection of different size objects before and after improvement

测任务,例如 SSD、YOLO 系列等)和双阶段目标检测算法(首先生成一系列候选区域框,再判定其中是否含有目标,比如 Faster R-CNN、Mask R-CNN 等)。本文采用 PIS-YOLO 与其他主流的单阶段和双阶段目标检测算法进行对比,以此来验证其模型效能。在 X 射线注塑工件缺陷检测任务中,典型实例分割模型的性能结果如表 4 所示。从表 4 可以观察到,YOLO 系列算法的表现优于 Mask R-CNN,而在 YOLO 系列中,YOLOv8n 模型的性能最为出色。经过进一步优化后,PIS-YOLO 的性能得到了显著提升,在准确性和实时性方面均达到了最优水平。

表 4 不同模型的对比实验

Table 4 Comparative experiments of different models

模型	Segment mAP@0.5/%	帧率/fps	Params/($\times 10^6$)
YOLOv7-seg	81.4	23.6	11.93
YOLOv8n-seg	82.4	80.7	3.51
本文	86.1	88.3	2.40
Mask R-CNN	80.8	16.4	13.90
YOLOACT	82.2	61.2	2.79

为了评估模型的实际应用能力,本文使用上述的 3 种网络与改进网络在检测结果上进行对比,如图 7 所示。

在对气孔缺陷进行识别时,Mask R-CNN 在预测置信度方面整体偏低;YOLOv7 模型存在漏检现象;而 YOLOv8 原始模型则容易出现重复检测及错误识别的问题。在实例分割任务中,原始模型在处理小尺寸目标时表现欠佳,整体分割效果不及改进后的模型。而针对实例分割任务的 YOLOv8-seg,对小目标的检测效果要优于前两者,但从缺陷边缘部位的检测效果和分割精度上看,拥有高分辨率分支的 PIS-YOLO 因其更丰富的细节信息以及多尺度特征的更好融合,使得最终的实际检测结果要优于以上模型。因此,改进网络 PIS-YOLO 具有较好的泛化性。

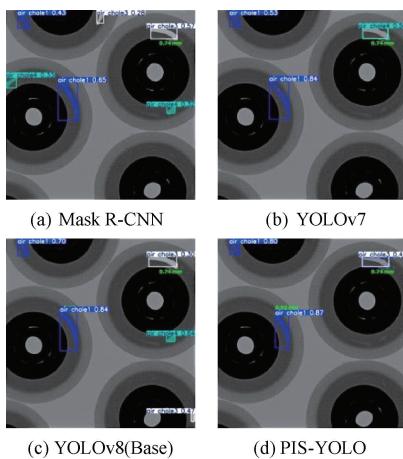


图 7 模型效果对比

Fig. 7 Comparison of model effects

4 结 论

为了提高注塑工件内部气孔缺陷的检测精度和速度,本文提出了一种针对小目标工业场景 X 射线注塑工件缺陷检测的轻量化缺陷检测网络 PIS-YOLO, 主要通过 3 个创新模块来提升性能与效率。为了有效减少模型的参数量和计算量, 增强模型的特征提取和表达能力, 结合部分卷积操作设计了 HG-Net 模块。其次, 采用共享卷积层并结合 iRMB_EMA 注意力机制, 优化特征的上下文处理, 提升了模型的目标定位和分割性能。最后, 采用特征通道数减少的解耦检测头, 有效减少了模型的参数量和计算复杂度, 提高了模型对尺度变化的特征性。实验结果表明, PIS-YOLO 在自制数据集上的表现优越, 模型泛化能力强, 参数量仅为 2.40×10^6 , 计算量为 9.4 GFLOPs, 权重大小为 4.9 MB, 准确率为 89.4%, 体现实例分割精度的 Segment mAP@0.5 可达到 86.1%, 显著优于同类算法, PIS-YOLO 模型不仅提升了检测精度, 也在保持高效性的同时, 大幅度降低了资源消耗, 展示了在工业应用中的广泛潜力和实际应用价值。

参考文献

- [1] LIU M, CHEN Y, XIE J, et al. LF-YOLO: A lighter and faster yolo for weld defect detection of X-ray image [J]. IEEE Sensors Journal, 2023, 23(7): 7430-7439.
- [2] ADIBHATLA V A, CHIH H C, HSU C C, et al. Applying deep learning to defect detection in printed circuit boards via a newest model of you-only-look-once [J]. Mathematical Biosciences & Engineering, 2021, 18(4): 4411.
- [3] 周钦沅. 基于 X 射线成像的集成电路封装缺陷检测技术 [J]. 中国战略新兴产业, 2024 (36): 129-131.
- [4] 傅健, 张昌盛, 朱国港, 等. X 射线分层层析成像技术及在航空航天领域的应用 [J]. 航空制造技术, 2019, 62(14): 49-54.
- [5] FU J, ZHANG CH SH, ZHU G G, et al. X-ray points various imaging techniques and the application in the field of aerospace [J]. Aviation Manufacturing Technology, 2019, 62(14): 49-54.
- [6] 张瑾, 李洁, 魏子璇, 等. GFRP 损伤 X 射线和超声无损检测的融合方法 [J]. 电子测量与仪器学报, 2024, 38(8): 169-177.
- [7] ZHANG J, LI J, WEI Z X, et al. GFRP damage of the fusion of X ray and ultrasonic nondestructive testing method [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(8): 169-177.
- [8] 王元才, 冯小辉, 史卜凡, 等. 基于机器视觉的注塑件缺陷检测系统研究 [J]. 化工自动化及仪表, 2024, 51(1): 113-119.
- [9] WANG Q C, FENG X H, SHI B F, et al. Defect-detecting system for injection molding parts based on machine vision [J]. Control and Instruments in Chemical Industry, 2024, 51(1): 113-119.
- [10] 杜娟, 谭笑宇, 黄涛, 等. 针对不均匀光照高反光注塑工件图像的缺陷定位与分割 [J]. 制造业自动化, 2021, 43(2): 31-38.
- [11] DU J, TAN X Y, HUANG T, et al. Defects location and segmentation for images of high reflection molding products with uneven illumination [J]. Manufacturing Automation, 2021, 43(2): 31-38.
- [12] LIU A, HAN L. IMP-DETR: Optimization model for defect detection of injection-moulded products [J]. The Journal of Engineering, 2024, 2024(12): e70032.
- [13] HU Z, YIN Z, QIN L, et al. A novel method of fault diagnosis for injection molding systems based on improved Vgg16 and machine vision [J]. Sustainability, 2022, 14(21): 14280.
- [14] 苏佳, 秦一畅, 贾泽, 等. 基于 ATO-YOLO 的小目标检测算法 [J]. 计算机工程与应用, 2024, 60(6): 68-77.
- [15] SU J, QIN Y CH, JIA Z, et al. Small object detection algorithm based on ATO-YOLO [J]. Computer Engineering and Applications, 2024, 60(6): 68-77.
- [16] GARFO S, MUKTADIR M A, YI S. Defect detection on 3D print products and in concrete structures using image processing and convolution neural network [J]. Journal of Mechatronics and Robotics, 2020, 4(1): 74-84.
- [17] ZHOU Q Y. Integrated circuit package defect detection technology based on X-ray imaging [J]. China Strategic Emerging Industries, 2024 (36): 129-131.

- [12] LI Y D, HAN D, LI H G, et al. Multi-block SSD based on small object detection for UAV railway scene surveillance[J]. Chinese Journal of Aeronautics, 2020, 33(6): 1747-1755.
- [13] 罗仁泽, 唐祥, 余泓, 等. 基于改进 Faster RCNN 的射线图像焊缝缺陷检测方法[J]. 电子测量技术, 2023, 46(22): 160-168.
- LUO R Z, TANG X, YU H, et al. Weld defect detection method of ray image based on improved Faster RCNN[J]. Electronic Measurement Technology, 2023, 46 (22): 160-168.
- [14] JIANG D, LI G, TAN C, et al. Semantic segmentation for multiscale target based on object recognition using the improved Faster-RCNN model [J]. Future Generation Computer Systems, 2021, 123: 94-104.
- [15] 高嘉, 刘涛, 王显峰, 等. TWRD-Net: 一种用于曳引钢丝绳缺陷的实时检测网络算法[J]. 仪器仪表学报, 2023, 44(6): 223-235.
- GAO J, LIU T, WANG X F, et al. TWRD-Net: A real-time detection network algorithm for traction wire rope defects[J]. Chinese Journal of Scientific Instrument, 2023, 44(6): 223-235.
- [16] 赵佰亭, 张晨, 贾晓芬. ECC-YOLO: 一种改进的钢材表面缺陷检测方法[J]. 电子测量与仪器学报, 2024, 38(04): 108-116.
- ZHAO B T, ZHANG CH, JIA X F, et al. ECC-YOLO: An improved method for detecting surface defects in steel[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4): 108-116.
- [17] 王子龙, 荣杰. 基于 YOLOv7-tiny 改进的遥感小目标检测算法[J]. 国外电子测量技术, 2024, 43(7): 65-74.
- WANG Z L, RONG J, et al. Improved remote sensing tiny object detection algorithm based on YOLOv7-tiny [J]. Foreign Electronic Measurement Technology, 2024, 43(7): 65-74.
- [18] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [19] ZHANG Q L, YANG Y B. SA-Net: Shuffle attention for deep convolutional neural networks [C]. ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021: 2235-2239.
- [20] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [21] XIE W, ZHANG Y, SUN H, et al. Hyperspectral image classification Via 3D-CNN MHSA fusion transformer[C]. IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2023: 7633-7636.
- [22] TAKAHASHI R, MATSUBARA T, UEHARA K. Data augmentation using random image cropping and patching for deep CNNs[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30 (9): 2917-2931.
- [23] ZHANG H, CISSE M, DAUPHIN Y N, et al. mixup: Beyond empirical risk minimization [J]. ArXiv preprint arXiv:1710.09412, 2017.
- [24] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. ArXiv preprint arXiv:2004.10934, 2020.

作者简介



林子涵, 2022 年于南昌航空大学获得学士学位, 广东工业大学在读研究生, 主要研究方向为图像处理、机器视觉。

E-mail: lzh15216099662@163.com

Lin Zihan received her B. Sc. degree from Nanchang University of Aeronautics and Astronautics (NCAA) in 2022. She is now a M. Sc. candidate at Guangdong University of Technology (GDUT). Her main research interests include image processing and machine vision.



张巧芬(通信作者), 2014 年于广东工业大学获得博士学位, 现为广东工业大学副教授, 主要研究方向为机器视觉、光纤通信。

E-mail: zhqf@gdut.edu.cn

Zhang Qiaofen (Corresponding author) received her Ph. D. degree from Guangdong University of Technology (GDUT) in 2014. She is now an associate professor at GDUT. Her main research interests include machine vision and fiber optic communication.