

结合特征复用和廉价操作的高精度光伏玻璃边部缺陷实时检测算法

丁绪星 周学顶 钱强 任悦悦 冯友宏

安徽师范大学物理与电子信息学院 安徽 芜湖 241002

摘要 针对现有的缺陷检测算法计算量大、参数量多、检测速度慢、检测精度低等问题,文中提出了一种基于 YOLOv5 的高精度光伏玻璃边部缺陷实时检测算法。首先使用新设计的一种基于廉价操作和特征复用的稠密连接模块(New_DBlock(C))替代 YOLOv5 特征提取网络的 C3 模块,减少整个算法的计算量和参数量;其次使用融合了通道注意力机制 SE(Squeeze-and-Excitation)的 C2f_SE 模块替换 YOLOv5 特征融合网络的 C3 模块,实现检测速度以及检测精度的提升;最后使用改进的 YOLOv8 解耦检测头取代 YOLOv5 的耦合检测头,提升算法的定位精度和分类精度。实验结果表明,改进后的算法 mAP@0.5 提升了 1.0%,mAP@0.5:0.95 提升了 3.1%,计算量下降了 48.1%,参数量下降了 56.7%,检测速度提升了 18.5%;与其他主流的 YOLO 和 R-CNN 系列算法相比,改进后的算法同样具有较高的检测精度、检测速度以及较低的计算量和参数量,适合光伏玻璃边部缺陷的实时检测。

关键词 光伏玻璃;缺陷检测;YOLOv5s;目标检测;注意力机制

中图分类号 TP391

High-precision and Real-time Detection Algorithm for Photovoltaic Glass Edge Defects Based on Feature Reuse and Cheap Operation

DING Xuxing, ZHOU Xueding, QIAN Qiang, REN Yueyue and FENG Youhong

School of Physics and Electronic Information, Anhui Normal University, Wuhu, Anhui 241002, China

Abstract Aiming at the problems of large amount of calculation, large amount of parameters, slow detection speed and low detection accuracy of existing defect detection algorithms, this paper proposes a high-precision and real-time detection algorithm for photovoltaic glass edge defects based on YOLOv5. Firstly, a newly designed dense connection block(New_DBlock(C)) based on cheap operation and feature reuse is used to replace the C3 block of YOLOv5's feature extraction network, which reduced the calculation amount and parameter amount of the whole algorithm. Secondly, the C2f_SE block fused with the channel attention mechanism SE(Squeeze-and-Excitation) is used to replace the C3 block of the YOLOv5's feature fusion network to improve the detection speed and detection accuracy. Finally, the improved YOLOv8's decoupling detection head is used to replace the coupling detection head of YOLOv5 to improve the positioning accuracy and classification accuracy of the algorithm. The experimental results show that the improved algorithm mAP@0.5 is increased by 1.0%, mAP@0.5:0.95 is increased by 3.1%, the amount of calculation is decreased by 48.1%, the amount of parameters is decreased by 56.7%, and the detection speed is increased by 18.5%. Compared with other mainstream YOLO and R-CNN series algorithms, the improved algorithm also has higher detection accuracy, detection speed, and lower amount of calculation and parameters, which is suitable for the real-time detection of photovoltaic glass edge defects.

Keywords Photovoltaic glass, Defect detection, YOLOv5s, Object detection, Attention mechanism

1 引言

光伏玻璃是太阳能电池重要的光伏组件之一,其质量直接决定了光伏组件的安全性能、使用寿命和发电效率等,因此光伏玻璃在出厂前必须进行缺陷检测。光伏玻璃边部检测是光伏玻璃工业检测中至关重要的一项检测工序,其主要检测方式有人工检测和机器检测两种。人工检测效率低、成本高、检测结果不一致且容易漏检误检,难以满足高质量高效率的检测要求。而基于神经网络的机器检测虽能够精准高效地识别和定位缺陷,但在现有的国内外文献中,基于神经网络的光

伏玻璃边部缺陷检测算法研究成果较少,经典的目标检测算法的检测能力有限,因此需借鉴其他工业制品的神经网络缺陷检测算法作为基础研究对象。

目前,基于神经网络的缺陷检测算法主要可分为两阶段检测算法和单阶段检测算法。两阶段检测算法使用区域提议网络生成一系列候选区域,并对候选区域进行分类和位置回归,代表性的两阶段检测算法主要有 Faster R-CNN^[1], Mask R-CNN^[2]和 Cascade R-CNN^[3]等。该类算法具有较强的鲁棒性和较高的准确率,经常用于自动检测领域,如 Lv 等^[4]将 Faster R-CNN 和 DCGAN^[5]相结合,解决了缺陷样本少的

基金项目:国家自然科学基金(62071005)

This work was supported by the National Natural Science Foundation of China(62071005).

通信作者:丁绪星(dx200@ahnu.edu.cn)

问题;Li 等^[6]使用残差网络 ResNet^[7]作为 Faster R-CNN 骨干特征提取网络,并在其中引入路径增强特征金字塔网络,解决了缺陷检测精度低的问题;Liu 等^[8]结合 Mask R-CNN 和特征金字塔网络,解决了小缺陷检测性能差的问题;Bai 等^[9]在 Cascade R-CNN 的特征提取网络中融合可变性卷积,并引入平衡特征金字塔和 GIoU^[10]回归损失,不仅可以自适应地提取不同形状和尺寸的缺陷特征,还进一步提升了检测精度。然而,两阶段检测算法需要进行两次神经网络计算(一次用于生成候选区域,另一次用于分类和位置回归),导致计算量大、参数量多、运行速度慢,因此只适用于线下检测,无法满足工业实时检测的需求。

相比两阶段检测算法,单阶段检测算法是一种将目标分类和位置回归结合起来的算法,该系列算法具有更简洁的网络结构和更快的检测速度,适合实时的检测领域,代表性的两阶段检测算法有 YOLO 系列^[11-14],SSD^[15],RetinaNet^[16]等。基于单阶段检测算法的缺陷检测虽然检测速度快、检测精度较高,但依旧存在许多不足,如 Hu 等^[17]在 YOLOX^[18]网络模型的基础上,结合 Transformer^[19]结构,实现了缺陷的高精准识别,但 Transformer 结构增加了计算资源的负担,提高了计算硬件设备的要求;Lan 等^[20]通过在 YOLOv8^[21]的骨干特征提取网络和颈部网络之间增加一个跨尺度融合模块,使得网络特征融合和交互的能力增强,提升了检测精度,但同时也增加了整个模型的计算量和参数量;Liu 等^[22]将 SSD 的基础网络替换成 ResNet,并加入预测模块和反卷积模块,实现了缺陷检测精度的提升,但同时牺牲了一定的检测速度;Zhou 等^[23]简化 RetinaNet 网络模型,同时引入难负样本挖掘策略和自适应忽略样本选择策略,提升了缺陷检测的精度,降低了模型过拟合的风险,但难负样本挖掘策略同样会丢失一些损失低但对负样本学习重要的训练样本,使得模型的检测精度

提升有限;Wang 等^[24]通过改进 YOLOX 的 CSPDarknet 模块,并引入通道注意力机制 CA(Coordinate Attention)^[25],实现了缺陷检测精度与检测速度的双项提升,但检测速度提升有限。

为了解决以上缺陷检测行业出现的问题,本文基于 YOLOv5^[26]提出了一种结合特征复用和廉价操作的高精度光伏玻璃边部缺陷实时检测算法。首先,通过设计一种基于廉价操作和特征复用的稠密连接模块替换 YOLOv5 特征提取网络的 C3 模块,降低模型的参数量和计算量;其次,通过融合了通道注意力机制的 C2f_SE 模块代替 YOLOv5 特征融合网络的 C3 模块,提升模型检测速度和检测精度;最后,通过改进的 YOLOv8 解耦检测头代替 YOLOv5 耦合检测头,实现模型分类精度和定位精度的提升。实验结果表明,针对光伏玻璃的 4 种常见边部缺陷(崩边、亮边、烧边、水滴),本文所设计的算法不仅在计算量、参数量、检测速度以及检测精度方面均取得了大幅度的提升,而且整体检测性能超越了目前主流的目标检测算法,能够实现检测精度、检测速度以及计算资源之间的较好平衡,满足线上光伏玻璃边部缺陷检测的实际要求。

2 研究方法

2.1 YOLOv5 目标检测算法

YOLOv5 是一种典型的单阶段目标检测算法,能够通过神经网络模型学习图像中的物体特征,并使用多尺度预测和网络分割来识别和定位物体。该目标检测算法拥有高效的检测速率以及强大的识别定位能力,经常用于实时的缺陷检测场景。

图 1 为 YOLOv5 的网络结构图,主要包括特征提取网络(Backbone)、特征融合网络(Neck)和检测头(Head)。

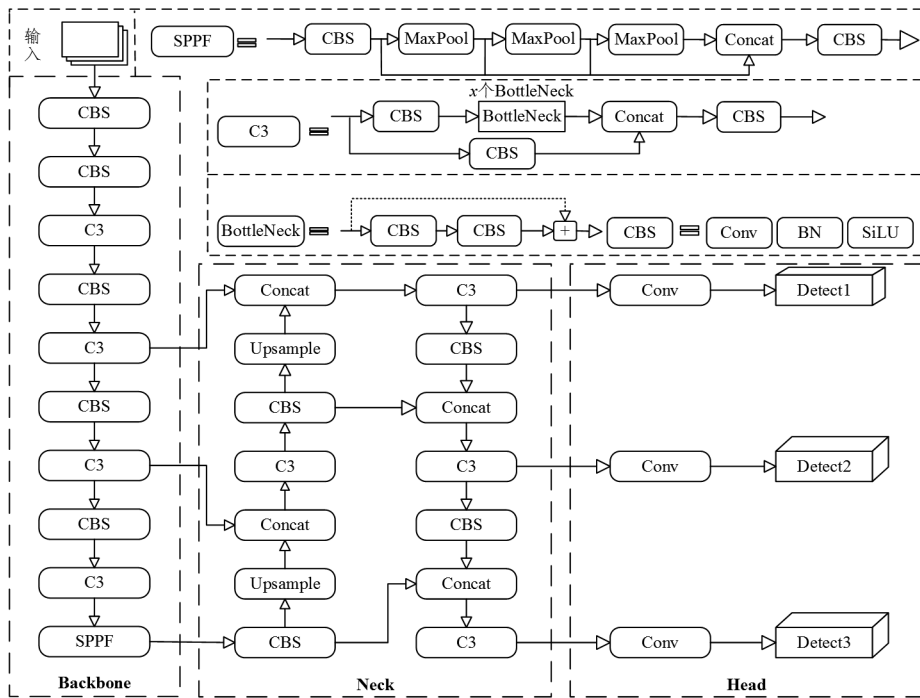


图 1 YOLOv5 结构示意图

Fig. 1 Diagram of the structure of YOLOv5

Backbone 是 YOLOv5 模型的基础部分,负责对输入图像进行特征提取,其主要结构包含 CBS 模块、C3 模块和 SPPF

模块。其中,C3 模块由 3 个 CBS 模块和 1 个连续且重复堆叠的 BottleNeck 组成,用于增加网络深度和感受野,提高特征

提取能力;但该模块在特征提取时,特征图主要由模块的顶层逐步传递至模块的底层,导致整个传递过程中每个网络层的特征图只使用了一次,没有重复利用模块内部卷积层的特征信息,使得模型的检测性能有限。

Neck是YOLOv5模型的颈部结构,负责对来自于特征提取网络的特征图进行多尺度特征融合,并把这些特征传递给预测层。但该网络的特征融合模块C3主要通过并联结构来逐步产生模块的特征融合输出,很大程度上影响了模型的检测速度。

Head是YOLOv5模型的预测层,负责对来自于Neck的特征图进行卷积融合、加权求和等操作,进而预测出目标物体的边界框和类别概率。然而,如图1所示,YOLOv5的检测头所采用的是耦合结构,虽然一定程度上加快了模型的检测速度,但耦合结构使得模型的检测任务和分类任务相互约束,导致模型反馈调参受限。同时,YOLOv5检测头所采用的Anchor-Based方法依靠提前训练好的先验框拟合目标位置,不能通过目标大小灵活调整,导致模型的定位能力有限。

因此,YOLOv5具有较强的目标检测性能,可用于工业缺陷检测类的实时性任务,但在计算量、参数量、检测速度和检测精度等方面仍存在不足。为了解决以上问题,本文针对YOLOv5的Backbone、Neck和Head部分分别进行优化,以胜任高精度的工业缺陷实时检测任务。

2.2 Backbone网络改进

2.2.1 高效的稠密连接模块

在YOLOv5的Backbone网络中,C3模块是特征提取的一个重要组成部分,通过多尺度特征融合技术和跨通道信息传递机制来提高模型的特征表达能力。但C3模块没有充分利用内部卷积层的上下文信息,导致整个主干网络的特征提取能力有限。同时,C3模块结构较为复杂,使得整个主干网络的计算量和参数量也相对较高。为了解决以上问题,本文基于特征复用和廉价操作设计了新的稠密连接模块——New_DBlock(C)代替YOLOv5的C3模块。其中,特征复用指在网络模块中,当前网络层的输入来自于前面所有网络层输出特征图以及模块输入特征图的堆叠,从而将模块的各个网络层的输出特征图传递至底层,使得模型能够充分利用模块内部卷积层的特征信息。同时,每个网络层通过特征图堆叠来收集输入特征图,一定程度上降低了模型的计算量。而廉价操作指在保证输出特征图通道数不变的情况下,使用计算量和参数量更低的模块来产生模块输出特征图的特征冗余部分(相似特征图),从而降低模块输出特征图冗余部分的计算量和参数量,进一步提升模型的检测效率。具体的New_DBlock(C)结构如图2所示。

一方面,整个模块的输入特征图 X_0 传递给DenseLayer_1,得到第一个DenseLayer的输出特征图 Y_1 ,用于调整后面所有DenseLayer的输入和输出通道数,其过程满足式(1):

$$Y_1 = H_1(X_0) \quad (1)$$

其中, $H_1(\cdot)$ 为第一个DenseLayer的运算操作,其结构与CBS模块相似,主要由批量归一化处理BN、激活函数ReLU和卷积Conv依次串联而成,且不改变输出特征图的尺寸。

然后,将整个模块的输入特征图 X_0 和第一个DenseLayer的输出特征图 Y_1 传递给DenseLayer_2,得到第二个DenseLayer的输出特征图 Y_2 ,其过程满足式(2):

$$Y_2 = H_2(X_0, Y_1) \quad (2)$$

其中, $H_2(\cdot)$ 为第二个DenseLayer的运算操作,其结构与第一个DenseLayer相同,且不改变输出特征图的尺寸,同时其输出通道数与第一个DenseLayer的输出通道数保持一致。

以此类推,直至得到最后一个DenseLayer的输出特征图 Y_n ,其结果满足式(3):

$$Y_n = H_n(X_0, Y_1, \dots, Y_{n-1}) \quad (3)$$

其中, $H_n(\cdot)$ 为最后一个DenseLayer的运算操作, Y_{n-1} 为倒数第二个DenseLayer的输出特征图,每一个DenseLayer的模块结构与第一个DenseLayer相同,同时不改变输出特征图的尺寸,输出通道数同样与第一个DenseLayer的输出通道数保持一致。

另一方面,整个模块的输入特征图 X_0 传递给Cheap模块,得到Cheap模块的输出特征图 Y_c ,其具体过程满足式(4):

$$Y_c = Cheap(X_0) \quad (4)$$

其中, $Cheap(\cdot)$ 为Cheap模块的运算操作,该操作主要由两个并行的 1×1 卷积组成,旨在用于产生每个DenseLayer层的特征冗余部分,具体结构如图3所示。其中,在设计时,Cheap模块的输出通道数与每个DenseLayer的输出通道数比值 r 满足式(5):

$$r = \frac{n(1-\lambda)}{\lambda} \quad (5)$$

其中, n 为DenseLayer的层数,其取值范围为正整数, λ 为DenseLayer层输出通道数的调节因子,其范围为 $0 \sim 1$ 。

接着,将整个模块的输入特征图 X_0 、Cheap模块的输出特征图 Y_c 和每一个DenseLayer的输出特征图 $\{Y_1, Y_2, \dots, Y_n\}$ 进行堆叠,将其作为整个稠密连接层的特征输出,其过程满足式(6):

$$Y_0 = [X_0, Y_1, Y_2, \dots, Y_n] \quad (6)$$

其中, Y_0 为整个稠密连接层的输出。

最后,将整个稠密连接层的输出 Y_0 传递给输出调节层,具体过程满足式(7):

$$Y = CBS(Y_0) \quad (7)$$

其中, $CBS(\cdot)$ 为输出调节层的运算操作, Y 为整个New_DBlock(C)模块的输出。

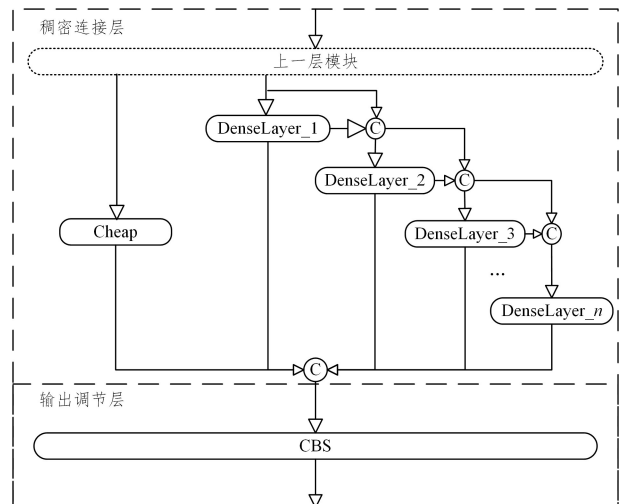


图2 New_DBlock(C)模块结构示意图

Fig. 2 Diagram of the structure of New_DBlock(C) module

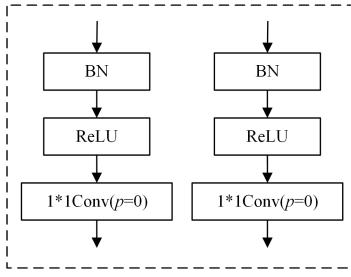


图3 Cheap模块结构示意图

Fig. 3 Diagram of the structure of Cheap module

2.2.2 理论分析

为了进一步证明 New_DBlock(C)模块在计算量和参数量方面的优势,本文接下来将通过理论推导分析在相同输入和输出通道数的情况下,C3模块和New_DBlock(C)模块分别在参数量和计算量上的差值。

如图2所示,假设一个New_DBlock(C)模块有 n 个DenseLayer层,该模块的输入特征图通道数为 C_0 ,每一个DenseLayer层的输出特征图通道数为 λC 。由式(3)可知,DenseLayer层的输入特征图来自于模块的输入特征图和前面所有DenseLayer层的输出特征图叠加,因此第 i 个DenseLayer层的输入特征图通道数 C_i 满足式(8):

$$C_i = C_0 + (i-1)\lambda C \quad (8)$$

New_DBlock(C)模块中,每一个DenseLayer层的卷积操作都是同一种卷积组合,因此,若每一个DenseLayer层输入同样大小的单通道特征图且得到大小不变的单通道输出特征图,其所需的计算量相等(设为 f),参数量相等(设为 p)。由于每一个DenseLayer层的计算量和参数量主要来自于卷积操作,因此可用卷积操作的计算量和参数量代替每一个DenseLayer层的计算量和参数量,则第 i 个DenseLayer层的计算量 f_i 和参数量 p_i 满足式(9)和式(10):

$$f_i = [C_0 + (i-1)C\lambda]C\lambda f \quad (9)$$

$$p_i = [C_0 + (i-1)C\lambda]C\lambda p \quad (10)$$

因此, n 个DenseLayer层的计算量 F_1 和参数量 P_1 分别满足式(11)和式(12):

$$F_1 = \sum_{i=1}^n f_i = \frac{1}{2} n [2C_0 + (n-1)C\lambda] C\lambda f \quad (11)$$

$$P_1 = \sum_{i=1}^n p_i = \frac{1}{2} n [2C_0 + (n-1)C\lambda] C\lambda p \quad (12)$$

由于Cheap模块和DenseLayer的输出通道数需满足式(6),且每一个DenseLayer层的输出特征图通道数为 λC ,则Cheap模块的输出通道数 C_{cheap} 满足式(13):

$$C_{cheap} = n(1-\lambda)C \quad (13)$$

为了能够让层与层之间特征冗余产生的计算量和参数量尽可能少,本文选取了 1×1 卷积作为Cheap模块的核心模块。在相同单通道输入和单通道输出情况下,Cheap模块的计算量和参数量远远小于DenseLayer层的卷积的计算量和参数量,因此Cheap模块产生的计算量 F_2 、参数量 P_2 分别满足式(14)和式(15):

$$F_2 < C_0 C_{cheap} f = C_0 C n (1-\lambda) f \quad (14)$$

$$P_2 < C_0 C_{cheap} p = C_0 C n (1-\lambda) p \quad (15)$$

由此可得,整个稠密连接层的计算量 F' 、参数量 P' 和输出通道数 C_{out} 分别满足式(16)–式(18):

$$F' = F_1 + F_2 < C_0 C n f + \frac{1}{2} C^2 \lambda^2 n (n-1) f \quad (16)$$

$$P' = P_1 + P_2 < C_0 C n p + \frac{1}{2} C^2 \lambda^2 n (n-1) p \quad (17)$$

$$C_{out} = C_0 + n\lambda C \quad (18)$$

设输出调节层的计算量和参数量分别为 F_c 和 P_c ,则New_DBlock(C)模块的计算量和参数量分别满足式(19)和式(20):

$$F = F' + F_c < F_c + C_0 C n f + \frac{1}{2} C^2 \lambda^2 n (n-1) f \quad (19)$$

$$P = P' + P_c < P_c + C_0 C n p + \frac{1}{2} C^2 \lambda^2 n (n-1) p \quad (20)$$

其中, F 为New_DBlock(C)模块的计算量, P 为New_DBlock(C)模块的参数量。

图4为图1中C3模块的简化结构示意图,为了方便描述,将C3模块的最后一层卷积层命名为C3_down模块,其余部分命名为C3_up模块。

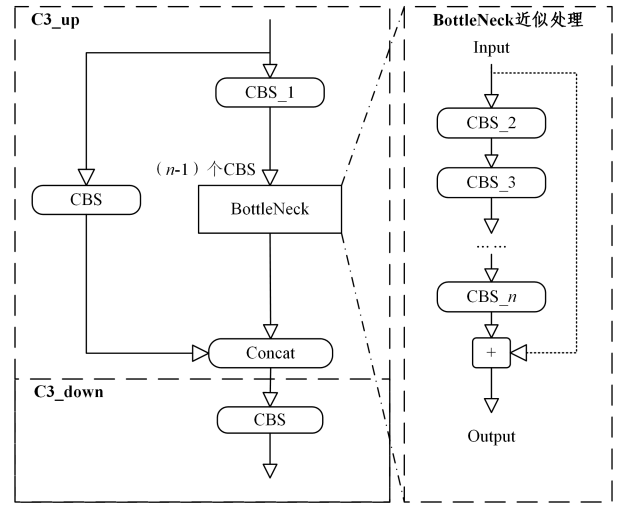


图4 C3模块简化结构示意图

Fig. 4 Diagram of the simplified structure of C3 module

由于C3_down模块与New_DBlock(C)模块的输出调节层是相同的CBS模块,则C3_down模块的计算量 F_{c3} 和参数量 P_{c3} 满足式(21)和式(22):

$$F_{c3} = F_c \quad (21)$$

$$P_{c3} = P_c \quad (22)$$

为了保证New_DBlock(C)模块和C3模块整体输入、输出通道数相等,则C3_up模块的输入通道数和输出通道数必须与New_DBlock(C)模块稠密连接层的输入特征图通道数和输出特征图通道数相同,即分别为 C_0 和 $(C_0 + nC)$ 。

由于YOLOv5在设计时,C3_up模块左右两支的输出通道数相等,且C3_up模块最终的输出通道数由左右两支堆叠而成,因此C3_up模块的左分支输出通道数 C_{l-out} 和右分支输出通道数 C_{r-out} 满足式(23):

$$C_{l-out} = C_{r-out} = \frac{1}{2} (C_0 + nC) \quad (23)$$

而CBS和DenseLayer均为不改变特征图尺寸和通道数的模块,其核心过程均为卷积,因此可近似看作是相同操作,则C3_up模块的左分支的计算量 F_{lc3} 和参数量 P_{lc3} 分别如式(24)和式(25)所示:

$$F_{F_{c3}} = C_0 C_{f-out} f = \frac{1}{2} (C_0 + nC) C_0 f \quad (24)$$

$$P_{P_{c3}} = C_0 C_{p-out} p = \frac{1}{2} (C_0 + nC) C_0 p \quad (25)$$

而 C3_up 模块右分支的第一个 CBS 层的输入和输出通道数分别为 C_0 和 C_{r-out} , 其余 $(n-1)$ 层卷积的输入和输出通道数均为 C_{r-out} , 则 C3_up 模块的右分支第一个 CBS 层的计算量 f_{r1} 和参数量 p_{r1} 分别满足式(26)和式(27):

$$f_{r1} = C_0 C_{r-out} f = \frac{1}{2} (C_0 + nC) C_0 f \quad (26)$$

$$p_{r1} = C_0 C_{r-out} p = \frac{1}{2} (C_0 + nC) C_0 p \quad (27)$$

其余 $(n-1)$ 个 CBS 的计算量 f_r 、参数量 p_r 分别满足式(28)和式(29):

$$f_r = C_{r-out} C_{r-out} f = \frac{1}{4} (C_0 + nC)^2 f \quad (28)$$

$$p_r = C_{r-out} C_{r-out} p = \frac{1}{4} (C_0 + nC)^2 p \quad (29)$$

则 C3_up 模块右分支的计算量 F_{r-c3} 和参数量 P_{r-c3} 分别如式(30)和式(31)所示:

$$\begin{aligned} F_{r-c3} &= (n-1)f_r + f_{r1} \\ &= \frac{1}{4} [n(n-1)C + (n+1)C_0] (C_0 + nC) f \end{aligned} \quad (30)$$

$$\begin{aligned} P_{r-c3} &= (n-1)p_r + p_{r1} \\ &= \frac{1}{4} [n(n-1)C + (n+1)C_0] (C_0 + nC) p \end{aligned} \quad (31)$$

则 C3 模块的整体计算量 F'' 和参数量 P'' 分别满足式(32)和式(33):

$$F'' = F_{f-c3} + F_{r-c3} + F_{c3}$$

$$= F_c + C_0 C n f + C_0^2 f + \frac{1}{4} (C_0 + nC)^2 (n-1) f \quad (32)$$

$$\begin{aligned} P'' &= P_{f-c3} + P_{r-c3} + P_{c3} \\ &= P_c + C_0 C n p + C_0^2 p + \frac{1}{4} (C_0 + nC)^2 (n-1) p \end{aligned} \quad (33)$$

将 C3 模块和 New_DBlock(C) 模块两者的计算量和参数量分别作差, 得到式(34)和式(35)的结果:

$$\begin{aligned} \Delta F &= F'' - F \\ &> \frac{1}{4} [(n-2\lambda^2)nC^2 + C_0^2 + 2C_0 C n] (n-1) f + C_0^2 f \end{aligned} \quad (34)$$

$$\begin{aligned} \Delta P &= P'' - P \\ &> \frac{1}{4} [(n-2\lambda^2)nC^2 + C_0^2 + 2C_0 C n] (n-1) p + C_0^2 p \end{aligned} \quad (35)$$

其中, ΔF 为两者计算量的差, ΔP 为两者参数量的差。

在 New_DBlock(C) 模块中, n 为 DenseLayer 的层数, 其取值范围为正自然数, λ 为 DenseLayer 输出通道数的调节因子, 其取值范围为 $0 \sim 1$, 则 ΔF 和 ΔP 大于 0 恒成立。因此, 在相同输入和输出通道数的情况下, New_DBlock(C) 模块的计算量和参数量始终少于 C3 模块。

2.3 Neck 网络的改进

YOLOv5 的 Neck 网络通过上下语义层的特征融合, 可以一定程度上提高模型的检测性能, 但该网络同样使用了结构复杂且参数量众多的 C3 模块, 导致模型的训练时间和实际的检测时间增加。为了提升模型的检测速度, 本文设计了一种新的梯度融合模块——C2f_SE 来代替 Neck 网络中的 C3 模块, 具体结构如图 5 所示。

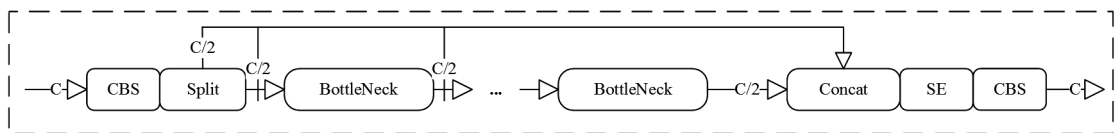


图 5 C2f_SE 模块结构示意图

Fig. 5 Diagram of the structure of C2f_SE module

首先将输入特征图平均分成两部分, 一部分直接传递至底层的堆叠层, 旨在保留输入特征的原始信息; 另一部分通过一系列模块中的卷积操作, 进行不同梯度的特征提取, 旨在保留不同网络层次的特征信息。接下来将不同网络层次的特征图和保留了原始信息的特征图进行堆叠拼接, 得到前面所有层次的特征图信息。

然而, 收集前面所有模块的输出特征图, 虽然能够使得整个模块获得丰富的特征信息, 但大量无差别地收集特征图必定也会收集到特征信息相对稀疏的特征图, 影响模型的检测性能。为了解决这个问题, 本文在堆叠层之后添加了通道注意力机制 SE (Squeeze-and-Excitation)^[27], 旨在引入权重参数区分不同特征通道的重要程度, 进而实现检测精度的提升。具体而言, 如图 6 所示, 通道注意力机制 SE 主要包括两个步骤: Squeeze 和 Excitation。其中, Squeeze 步骤用于捕捉输入特征图中每个通道的全局信息并输出一个通道维度的向量, 该步骤通过全局平均池化操作完成; Excitation 步骤能够将通道维度上的向量转换为通道维度上的权重, 用于与输入特征图相乘, 该步骤主要由全连接层和非线性激活函数组成。

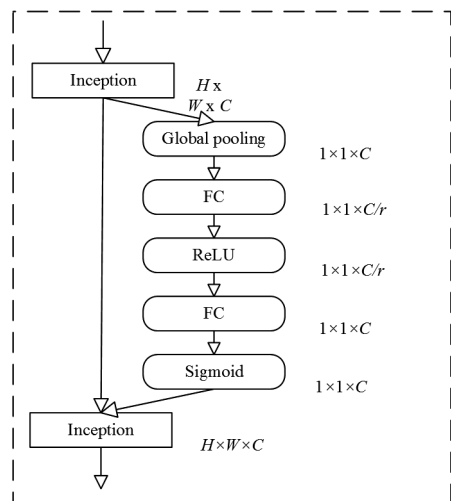


图 6 SE 模块结构示意图

Fig. 6 Diagram of the structure of SE module

最后, 为了进一步让不同通道的特征信息相互联系, 经过 SE 模块输出的不同通道特征图经过 CBS 模块进行特征提取, 实现了不同语义层的特征融合。

基于以上过程可知, C2f_SE 模块不仅能够同时利用低层细节和高层语义信息, 提高目标检测的准确性和鲁棒性, 还能同时使用特征堆叠, 加快特征图的收集, 提升整个模块的推理速度。此外, C2f_SE 模块能够自适应捕捉每个含有重要特征的通道, 关注输入数据的重要部分, 忽略或减少不重要的特征信息, 最终提高模型的表现能力。相对于 C3 模块, C2f_SE 模块通过结合多梯度的特征提取和注意力机制, 有效地提升了目标检测任务的性能和准确率, 同时利用特征堆叠保留了高效的计算能力和推理速度。因此, C2f_SE 模块的性能更加优越, 可替代 YOLOv5 特征融合网络的 C3 模块。

2.4 Head 网络改进

YOLOv5 的 Head 网络采用了耦合结构, 该结构在训练过程中共享分类和定位任务的权重信息, 导致分类和定位任务相互影响, 模型反馈调参相互牵制, 影响模型的检测能力。同时, YOLOv5 检测头采用了 Anchor-Based 方法, 该方法通过预定义的锚框来迅速生成候选目标框, 虽然提高了检测速度, 但由于锚框的尺寸和宽高比是固定的, 其对目标的定位能力有限。为了解决以上 YOLOv5 的 Head 网络存在的问题, 本文通过改进后的 YOLOv8 检测头来代替 YOLOv5 检测头, 具体 YOLOv8 检测头和改进后的 YOLOv8 检测头的结构如图 7 所示。

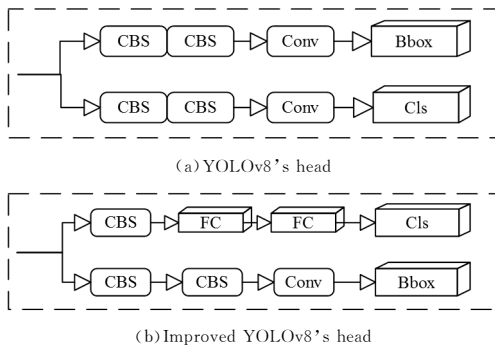


图 7 YOLOv8 检测头和改进后的 YOLOv8 检测头结构示意图

Fig. 7 Diagram of the structure of YOLOv8's head and improved YOLOv8's head

如图 7(a) 所示, YOLOv8 的检测头采用目前主流的解耦结构, 该结构通过两条不同的卷积分支将检测头的定位任务和分类任务分开进行, 虽不能同时优化分类和定位任务, 但定位和分类任务在反馈过程中不会相互牵制, 使得模型可以得到充分的训练。YOLOv8 检测头同时采用了 Anchor-Free 方法, 该方法不依赖于预设的锚框来定位正样本的位置, 而是基于物体自身的形状和大小来确定正样本的位置, 使得网络训练更加简便, 目标定位更加精准。其中采用的 Task-Aligned Assigner 正负样本匹配方式以分类和定位的分数加权作为选择正样本的分数, 能够有效地解决正负样本不均衡的问题, 提高模型定位的性能。

然而, Wu 等^[28]通过大量基础实验发现, 基于卷积层的检测头更适合做定位任务, 基于全连接层的检测头更适合做分类任务。鉴于以上结论, 本文对 YOLOv8 的检测头进行进一步改进, 改进后的结果如图 7(b) 所示。改进后的 YOLOv8 检测头依旧保留主流的解耦结构, 不同的是用于做分类任务的

3 个卷积层, 设计成了由一个卷积层和两个全连接层组成的串联分支。其中, 该分支的第一个卷积层用于调整输入特征图的尺寸和通道数, 第二、三个全连接层用于适配检测头的分类任务以提高分类性能。

因此, 相对于 YOLOv5 的检测头, 改进后的 YOLOv8 的检测头采用了目前主流的解耦结构, 解决了 YOLOv5 检测头模型反馈调参过程中分类和定位任务受限的问题; 采用的 Anchor-Free 方法, 使得网络能够根据目标自身的形状和大小自动定位标记, 解决了 YOLOv5 检测头定位不够精准的问题; 替换的全连接层分支适配于检测头的分类任务, 进一步提升了网络的分类效果。因此, 改进后的 YOLOv8 检测头整体性能更优越, 能够替代 YOLOv5 的检测头。

3 实验

3.1 实验数据集

本实验使用的边部缺陷图像来自于某光伏玻璃工厂采集所得, 经过裁剪、缩放、尺寸变化等预处理后得到最终实验所需数据集。该数据集包含 4 种光伏玻璃边部缺陷图像, 分别是崩边、亮边、烧边和水滴, 如图 8 所示。每个数据经过 Make Sense 工具标注并得到 .txt 格式的标签文件, 具体信息如表 1 所列。

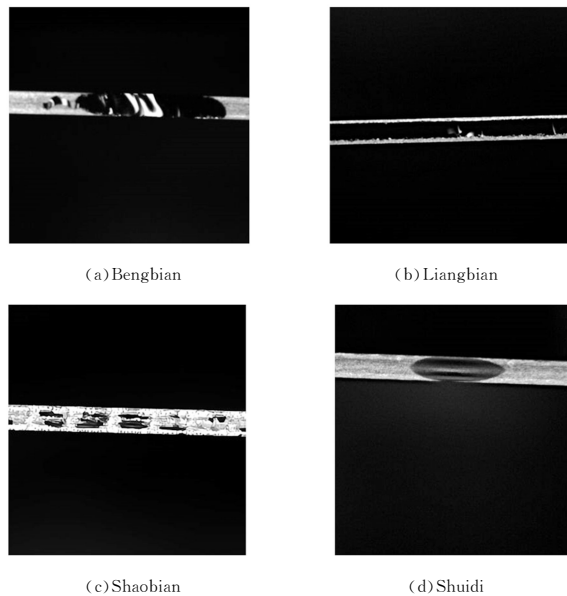


图 8 光伏玻璃边部缺陷图片

Fig. 8 Image of edge defect of photovoltaic glass

表 1 实验数据集信息

Table 1 Experimental dataset information

Parameters	Parameter values
Number of images	3 164
Average resolution	368 × 368
Average memory	7.02 kB
Categories	Bengbian, Liangbian, Shaobian, Shuidi
Number of defects of each category	Bengbian(1898), Liangbian(1 424), Shaobian(1 043), Shuidi(1 456)

3.2 实验环境

本实验使用的开发软件为 PyCharm, 开发框架为 Pytorch, 编程语言为 Python, 其他实验环境参数配置如表 2 所列。

表2 实验环境配置

Table 2 Experimental environment configuration

Parameters	Parameter values
CPU	Intel(R) Core(TM) i7-12700H(20 CPUs) @2.3GHz
GPU	NVIDIA GeForce RTX3060
RAM	32 GB
ROM	1 TB
Operating system	Windows 11

3.3 评价指标

本实验选用平均精度(Average Precision, AP)、平均精度均值(Mean Average Precision, mAP)、P-R曲线(Precision-recall curve)、参数量(Parameters)、计算量(Computation)以及检测一张缺陷图片的运行时间(Time)作为评价模型性能指标。其中 mAP 值为所有目标检测类别 AP 的均值, AP 值为 P-R 曲线的面积, P-R 曲线由召回率(Recall, R)和准确率(Precision, P)组成,具体表达式如式(36)~式(39)所示:

$$P = \frac{TP}{TP + FP} \quad (36)$$

$$R = \frac{TP}{TP + FN} \quad (37)$$

$$AP = \int_0^1 P(R) dR \quad (38)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i) \quad (39)$$

其中, TP 为模型预测为正例的正样本数; FP 为模型预测为负例的正样本数; FN 为模型预测为负例的正样本数; n 为缺陷类别数量; $AP(i)$ 为第 i 类缺陷的 AP 值。

3.4 实验结果与分析

3.4.1 模型改进前后的 P-R 曲线对比

图 9 给出了 YOLOv5s 模型改进前后 P-R 曲线结果,该 P-R 曲线结果由测试集数据产生。由图可知,对于崩边和水滴类别,改进后模型的 P-R 曲线面积比改进前模型的 P-R 曲线面积明显更大,说明了改进后的模型对崩边和水滴的识别率有了明显提升。整体上看,针对所有分类(all classes 曲线),改进前模型的 P-R 曲线平衡点为(0.930, 0.930),改进后模型的 P-R 曲线平衡点为(0.971, 0.971),即改进后模型的 P-R 曲线平衡点显著高于改进前模型的 P-R 曲线平衡点,说明改进后的模型在分类任务上性能更好。

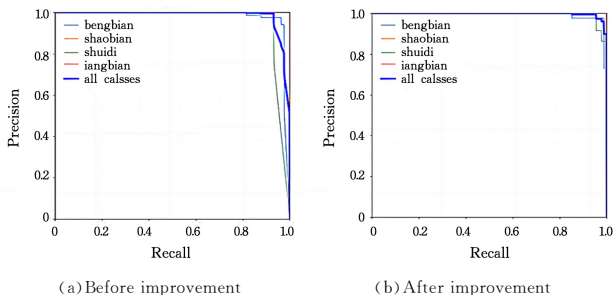


图9 改进前后 P-R 曲线结果图

Fig. 9 Results of P-R curves before and after improvement

3.4.2 模型改进前后的检测精度对比

为了检验改进 YOLOv5s 后的模型在检测精度上的性能,本文通过表 1 中的数据集进行仿真测试实验,实验结果如表 3 所列。其中,各缺陷类别的识别率是 IoU 为 0.5 时所得。 $mAP@0.5$ 表示在 IoU 为 0.5 时所有类别的 AP 值; $mAP@0.5:0.95$ 表示 IoU 分别为 0.5, 0.55, 0.6, 0.65, 0.7, 0.75,

0.8, 0.85, 0.9, 0.95 时,所有阈值下的 mAP 平均值。

表3 改进前后缺陷识别率的对比

Table 3 Comparison of defect recognition rate before and after improvement

Model	AP				mAP@	
	Bengbian	Liangbian	Shaobian	Shuidi	0.5	0.5:0.95
YOLOv5s ^[26]	0.978	0.995	0.995	0.958	0.982	0.882
Ours	0.989	0.995	0.995	0.991	0.992	0.913

注:表中各类别的 AP 值是在 IoU 为 0.5 时所得,加粗字体为每列最优值。

由表 3 的实验结果可知,当 IoU 设置为 0.5 时,改进后的模型对亮边和烧边的识别率与改进前的模型识别率一样,但改进后的模型对崩边和水滴的识别率分别为 98.9% 和 99.1%,相较于 YOLOv5s 模型,分别提升了 1.1% 和 3.3%;整体上看,改进后模型的 $mAP@0.5$ 为 99.2%, $mAP@0.5:0.95$ 为 91.3%,相较于改进前的模型,分别提升了 1.0% 和 3.1%。这主要由于改进后的 New_DBlock(C) 模块通过特征复用机制,将模块各网络层中的特征图传递至模块底层,增强了缺陷信息的有效传递,特别是对于存在小型缺陷的崩边缺陷和水滴缺陷,缺陷特征信息不会在特征提取过程中丢失,从而提升整个检测模型的检测精度。同时, C2f_SE 模块通过通道注意力机制 SE 将重要通道的缺陷特征图进行加权,增强了含有缺陷信息特征图的重要程度,使得模型的检测精度进一步提升。因此,在检测精度方面,改进后的模型识别率更高。

3.4.3 模型改进前后的检测速度对比

为了能够提升整个模型的检测效率,本文通过减少模型的计算量和参数量来轻量化模型,进而缩短模型的运行时间。表 4 列出了改进前后的 YOLOv5s 模型在计算量、参数量和运行时间上对比结果。其中, Δ_G , Δ_M , Δ_T 分别表示模型改进前后,在计算量、参数量和运行时间上的差值。

表4 改进前后计算量、参数量和运行时间上的对比

Table 4 Comparison of computation, parameters and running time before and after improvement

Model	Computation		Parameters		Running time	
	GFLOPs	Δ_G	Params	Δ_M	Time/ms	Δ_T
YOLOv5s ^[26]	16.0	7.7	7.03×10^6	3.99×10^6	15.7	2.9
Ours	8.3		3.04×10^6		12.8	

注:加粗字体为各列最优值。

由表 4 可知,改进后模型特征提取网络的核心特征提取模块,使用了计算量和参数量更低的 New_DBlock(C) 模块,降低了整个模型的计算量和参数量。其中,模型的计算量为 8.3 GFLOPs,参数量为 3.04×10^6 ,相较于改进前的 YOLOv5s 模型,分别下降了 48.1% 和 56.7%。同时,模型参数量和计算量的减少,提升了模型的检测速度; C2f_SE 的特征堆叠加快了特征融合网络的特征图收集,同样促进了模型检测速度的提升。由表 4 的实验结果可知,改进后的模型运行时间为 12.8 ms,相较于改进前的 YOLOv5s 模型,运行时间提升了 18.5%。因此,在计算量、参数量和检测速度方面,改进后的模型性能更优。

3.4.4 消融实验

为了探究本文所提出的每个改进重点对网络模型的优化效果,本文基于 YOLOv5s 模型进行了消融实验,实验结果如表 5 所列,实验结论如下。

1) New_DBlock(C)模块替换 Backbone 部分的 C3 模块,通过特征复用机制,模块中的特征信息得到有效传递,提升了模型对崩边微小缺陷和水滴小型缺陷的精准识别。其中,当 IoU 设置为 0.5 时,崩边缺陷和水滴缺陷分别提升了 0.68% 和 2.83%,整体上看,模型的 mAP@0.5:0.95 提升了 1.5%。同时,引入的廉价操作以及特征复用机制大幅度降低了模型的计算量和参数量,其中计算量降低了 70.6%,参数量降低了 56.9%,运行速度提升了 4.5%。

2) C2f_SE 模块替换 Neck 部分的 C3 模块,计算量增加了 0.6GFLOPs,参数量增加了 0.28×10^6 。但该模块通过多尺度融合和注意力机制,增强了模型的检测精度,其中,当 IoU 设置为 0.5 时,崩边、亮边和水滴的缺陷识别率分别提升了 0.35%,0.01% 和 0.12%,从整体上看,模型的 mAP@0.5:

0.95 提升了 1.1%。同时,由于 C2f_SE 模块通过堆叠融合来加快模块特征图的收集,运行速度提升了 12.7%。

3) 改进后的解耦检测头替换 YOLOv5 的耦合检测头,计算量增加了 3GFLOPs,参数量增加了 0.73×10^6 。但 Anchor Free 方法以及全连接分支分别一定程度上增强了模型的定位效果和分类精度,提升了模型的检测精度。因此,整体上看,模型的 mAP@0.5:0.95 提升了 0.5%。

因此,通过引入 New_DBlock(C)模块、C2f_SE 模块和解耦检测头,模型的检测精度均得到了提升,但 New_DBlock(C)模块同时大幅度降低了模型的参数量、计算量以及运行时间,C2f_SE 模块和改进的解耦检测头同时提升了模型的检测速度。因此改进后的模型性能显著优于改进前的模型性能,可代替 YOLOv5s 模型用于光伏玻璃的边部缺陷检测。

表 5 消融实验结果

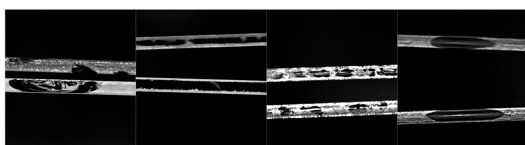
Table 5 Results of ablation experiments

①	②	③	AP/%				mAP@0.5: 0.95/%	GFLOPs	Params	Time/ms
			Bengbian	Liangbian	Shaobian	Shuidi				
—	—	—	97.80	99.48	99.51	95.79	88.2	16.0	7.03×10^6	15.7
✓	—	—	98.48	99.49	99.51	98.62	89.7	4.7	2.033×10^6	15.0
✓	✓	—	98.83	99.50	99.48	98.74	90.8	5.3	2.313×10^6	13.1
✓	✓	✓	98.91	99.52	99.53	99.13	91.3	8.3	3.043×10^6	12.8

注:“①”代表 New_DBlock(C)模块,“②”代表 C2f_SE 模块,“③”代表检测头改进,“✓”表示当前模型采用此模块,“—”表示当前模型未采用此模块,表中各类别的 AP 值是在 IoU 为 0.5 时所得,加粗字体为每列最优值。

3.4.5 模型改进前后的检测效果对比

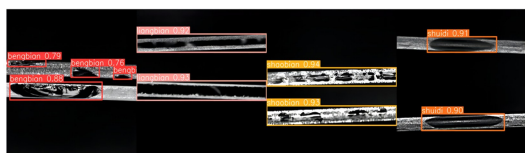
为了能够直观地感受 YOLOv5s 模型改进前后的检测效果,本章针对表 1 所列的部分测试集图片进行检测效果可视化处理,具体结果如图 10 所示。



(a) Original images



(b) Detection results before model improvement



(c) Detection results after model improvement

图 10 模型改进前后的可视化检测效果

Fig. 10 Visual detection effect before and after improvement

从识别效果上看,改进前的 YOLOv5s 模型对于大型的边部缺陷展现出了良好的识别能力,但对于小型的崩边缺陷存在明显的漏检现象。相比之下,改进后的模型通过特征复用机制和通道权重划分,不仅能够有效地识别出大型的边部缺陷,同时能够精准无误地识别出小型的崩边缺陷。从定位效果来看,改进前的 YOLOv5s 模型只能对水滴缺陷具有较好的定位效果,但对于大型或连通的崩边、烧边以及亮边缺陷,其定位效果较差,如将大型的崩边缺陷定位成两个重叠的

崩边缺陷,将大型的亮边缺陷定位成两个重叠的亮边缺陷,将连通的烧边缺陷标记成两个重叠的烧边缺陷。相对之下,改进后的模型通过 Anchor-Free 方法,不仅能够准确地找出所有的缺陷位置,同时可以根据缺陷的自身尺寸自动拟合,精准地标记出大型和连通的缺陷,使得缺陷定位更加准确灵活。因此,在检测效果方面,改进后的模型检测效果明显优于改进前的模型检测效果。

3.5 不同模型下的性能对比

为了进一步验证改进后模型性能的优越性,本文在相同数据集(见表 1)和相同实验环境(见表 2)下,将改进后的模型与现有经典的目标检测算法(Faster R-CNN, R-FCN^[29], Cascade R-CNN, SSD, YOLOv5, YOLOX, YOLOv7^[30], YOLOv8)以及缺陷检测算法(MFAM-Net^[31], FINet^[32], DefectNet^[33])进行对比,具体的实验结果如表 6 所列。

由表 6 的实验结果可知,相对于经典的二阶段目标检测模型,改进后的模型检测精度相差不大,但计算量、参数量以及运行时间得到了大幅度提升,其中,改进后的模型计算量减少了 93.2%~98.1%,参数量减少了 89.7%~97.2%,运行时间缩短了 31.9%~64.1%。相对于经典的单阶段目标检测模型,改进后的模型在检测精度、运行速度、计算量以及参数量上均展现出了优越的性能。其中,mAP@0.5 提高了 1.0%~5.8%,mAP@0.5:0.95 提高了 0.3%~7.6%,计算量减少了 37.1%~83.0%,参数量减少了 49.5%~87.5%,运行时间缩短了 6.6%~35.7%。相对于已有的缺陷检测算法 FMAM_Net, FINet 和 DefectNet,改进后的算法模型检测精度最高、运行速度最快,其中,mAP@0.5 提高了 0.2%~1.9%,运行速度提升了 16.9%~48.2%。在计算量和参数量明显优于 FMAM_Net 和 DefectNet,其中计算量分别相对减少了 48.4%和 67.7%,参数量分别相对减少了 73.6%和 76.5%。因此,综合考虑检测精度、检测速度以及资源占比等

各方面因素,改进后的模型检测精度高,检测速度快,计算资源占用少,更适合光伏玻璃边部缺陷的实时检测。

表6 不同模型的性能对比

Table 6 Performance comparison of different models

	Model	mAP@0.5	mAP@0.5:0.95	GFLOPs	Params	Time/ms
Two-Stage models	Faster R-CNN ^[1]	0.990	0.907	164.1	39.72×10^6	35.7
	R-FCN ^[29]	0.988	0.912	121.2	29.40×10^6	18.8
	Cascade R-CNN ^[3]	0.994	0.916	433.8	107.11×10^6	27.3
One-Stage models	SSD ^[15]	0.934	0.837	48.7	24.4×10^6	15.9
	YOLOv5s ^[26]	0.982	0.882	16.0	7.03×10^6	15.7
	YOLOv7_tiny ^[30]	0.979	0.874	13.2	6.02×10^6	18.4
	YOLOX_s ^[18]	0.975	0.888	26.9	8.94×10^6	19.9
	YOLOv8s ^[21]	0.982	0.910	28.4	11.13×10^6	13.7
Defect detection algorithms	FMAM-Net ^[31]	0.973	0.907	16.1	11.50×10^6	17.4
	FINet ^[32]	0.985	0.891	—	—	15.4
	DefectNet ^[33]	0.990	0.904	25.7	12.93×10^6	24.7
	Ours	0.992	0.913	8.3	3.04×10^6	12.8

注:加粗字体为每列最优值。

结束语 为了能够实时精准地检测出光伏玻璃的边部缺陷,本文提出了一种基于YOLOv5模型改进后的缺陷检测算法。该算法首先通过新设计的稠密连接模块New_DBlock(C)代替YOLOv5的特征提取网络的C3模块,实现模型计算量和参数量的大幅降低;其次使用融合了通道注意力机制SE的C2f_SE模块代替YOLOv5的特征融合网络的C3模块,实现检测精度和检测速度的双项提升,最后使用改进后的YOLOv8的解耦检测头代替YOLOv5的耦合检测头实现分类精度和定位精度的提升。实验结果表明,改进后的算法在检测精度、检测速度以及计算资源占比方面均取得了显著的提升,并且针对光伏玻璃的边部缺陷,其整体检测性能已超越目前主流的YOLO系列算法、R-CNN系列算法以及现有3种典型的缺陷检测算法(FMAM_Net, FINet和DefectNet)等,适合光伏玻璃边部缺陷的实时检测。但考虑到实际缺陷样本收集不易,下一步将重点研究如何通过少量的缺陷样本来生成大样本缺陷数据以训练神经网络算法模型。

参考文献

- [1] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [2] HE K, GKIOXARI G, PIOTR D, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision, Venice, Italy, IEEE, 2017: 2961-2969.
- [3] CAI Z W, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, IEEE, 2018: 6154-6162.
- [4] LV Y F L, MA L, JIANG H Q. A Mobile Phone Screen Cover Glass Defect Detection MODEL Based on Small Samples Learning[C]//2019 IEEE 4th International Conference on Signal and Image Processing, Wuxi, China, IEEE, 2019: 1055-1059.
- [5] CAO Y J, JIA L L, CHEN Y X, et al. Recent advances of generative adversarial networks in computer vision[J]. IEEE Access, 2018, 7: 14985-15006.
- [6] LI L, Z. JIANG Z J, LI Y N. Surface Defect Detection Algorithm of Aluminum Based on Improved Faster RCNN[C]//2021 IEEE

- 9th International Conference on Information, Communication and Networks, Xi'an, China IEEE Press, 2021: 527-531.
- [7] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA; IEEE Press, 2016: 770-778.
- [8] LIU G, WANG F, HUI H. Application of Mask R-CNN in defect detection of industrial filter cloth[C]//2022 Global Conference on Robotics, Artificial Intelligence and Information Technology, Chicago, IL, USA; IEEE, 2022: 794-798.
- [9] BAI M L, WANG M W. Fabric defect detection algorithm based on improved Cascade R-CNN [J]. Computer Science, 2023, 50(6A): 220300224-6.
- [10] REZATOFI G H, TSOI N, GWAK J A, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA; IEEE, 2019: 658-666.
- [11] JIANG P, ERGU D, LIU F, et al. A review of yolo algorithm developments [J]. Procedia Computer Science, 2022, 199: 1066-1073.
- [12] TERVEN J, CORDOVA-ESPARZA D M, ROMERO-GONZALEZ J A. 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.
- [13] KANG C H, KIM S Y. Real-time object detection and segmentation technology: an analysis of the YOLO algorithm [J]. JMST Advances, 2023, 5(2): 69-76.
- [14] HASAN R H, HASSOO R M, ABOUD I S. Yolo versions architecture: review [J]. International Journal of Advances in Scientific Research and Engineering, 2023, 9(11): 73-92.
- [15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector [C]//European Conference on Computer Vision, Cham; Springer International Publishing, 2016: 21-37.
- [16] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [17] HU H T, DU H C, WANG S Q, et al. Improved YOLOX method for detecting surface defects of drug blister aluminum foil [J]. Journal of Graphics, 2022, 43(5): 803-814.
- [18] GE Z, LIU S, WANG F, et al. Yolox: exceeding yolo series in

- 2021[J]. arXiv:2107.08430,2021.
- [19] HAN D,PAN X,HAN Y,et al. Flatten transformer:vision transformer using focused linear attention[C] // 2023 IEEE/CVF International Conference on Computer Vision. Paris, France:IEEE,2023:5961-5971.
- [20] LAN H,ZHU Z H,LUO R,et al. PCB defect detection algorithm of improved YOLOv8[C] // 2023 8th International Conference on Image, Vision and Computing. Dalian, China: IEEE, 2023:178-183.
- [21] SOHAN M,SAI-RAM T,REDDY R,et al. A review on YOLOv8 and its advancements[C] // International Conference on Data Intelligence and Cognitive Informatics. Springer, Singapore:Springer,2024:529-545.
- [22] LIU Z,ZHOU B. Research and application on the improved SSD chip defect inspection algorithm[C] // 2021 IEEE International Conference on Consumer Electronics and Computer Engineering. Guangzhou, China:IEEE,2021:551-555.
- [23] ZHOU B,JIANGP F,DUAN C,et al. Study on single background object detection oriented Improved-RetinaNet model and its application [J]. Computer Science,2023,50(7):137-142.
- [24] WANG X,GAO J S,HOU B J,et al. A Lightweight Modified YOLOX Network Using Coordinate Attention Mechanism for PCB Surface Defect Detection[J]. IEEE Sensors Journal,2022, 22(21):20910-20920.
- [25] HOU Q,ZHOU D,FENG J. Coordinate attention for efficient mobile network design[C] // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN, USA: IEEE,2021:13713-13722.
- [26] OLORUNSHOLA O E,IRHEBHUDE M E,EVWIEKPAEFE A E. A comparative study of YOLOv5 and YOLOv7 object detection algorithms[J]. Journal of Computing and Social Informatics,2023,2(1):1-12.
- [27] HU J,SHEN L,SUN G. Squeeze-and-excitation networks[C] // 2018 IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA:IEEE,2018:7132-7141.
- [28] WU Y,CHEN Y,YUAN L,et al. Rethinking classification and localization for object detection[C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA:IEEE,2020:10186-10195.
- [29] DAI J,LI Y,HE K,et al. R-FCN:object detection via region-based fully convolutional networks[C] // 30th Conference on Neural Information Processing Systems. Barcelona, Spain: Curran Associates, Inc,2016,29:379-387.
- [30] WANGCY,BOCHKOVSKIYA,LIAOHYM. YOLOv7:Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada:IEEE,2023:7464-7475.
- [31] ZENG Q,WEI D,ZHANG X,et al. MFAM-Net:a surface defect detection network for strip steel via multiscale feature fusion and attention mechanism[C] // 2023 International Conference on New Trends in Computational Intelligence (NTCI). Qingdao, China:IEEE,2023:117-121.
- [32] ZHANG Z D,ZHANG B,LAN Z C,et al. FINet: an insulator dataset and detection benchmark based on synthetic fog and improved YOLOv5[J]. IEEE Transactions on Instrumentation and Measurement. 2022,71:1-8.
- [33] LI F,XI Q G. DefectNet:toward fast and effective defect detection[J] IEEE Transactions on Instrumentation and Measurement. 2021,70:1-9.



DING Xuxing, born in 1971, professor, master's supervisor. His main research interests include digital image processing, computer vision, and object detection.