

DOI:10.19651/j.cnki.emt.2417091

基于改进 YOLOv8 的布匹缝头检测算法^{*}

刘瑞明 陈伦奥 徐春融 周韬

(江苏海洋大学电子工程学院 连云港 222005)

摘要: 工业场景下的布匹缝头检测在纺织应用领域越来越重要。然而,缝头检测面临着小目标尺寸、可利用特征少、复杂多变的环境因素等挑战,难以保证稳定且实时的检测的效果。针对这一系列问题提出一种基于改进 YOLOv8 的布匹缝头检测算法 YOLOv8-DVB。根据 Deformable Convolutional Networks v4 的特点优化 C2f 模块,提出一种多尺寸特征采样的 C2f-DCNv4 模块,强化网络对不同尺寸特征信息的提取。在颈部网络,采用 BiFPN 结构作为特征融合的方式,通过引入自上而下和自下而上的双向通路,使得不同尺度的特征可以在多层次上进行更充分的融合。其次,引入更高效的 VoV-GSCSP 模块轻量化特征融合网络,帮助颈部网络降低计算量和参数量,减少计算负担。最后,设计一个专门的小目标检测层,优化小目标的特征提取。通过实验对 YOLOv8-DVB 模型与原模型以及 YOLOv5、YOLOv7、和 Faster R-CNN 等进行比较,验证模型的检测准确率和检测精度。实验结果表明,该方法在自建数据集上获得 84.7% 的检测准确率,相比于原模型和其他网络模型都有着更高的准确率,能够快速有效的在复杂的工业生产环境中准确的检测到目标类别和位置。

关键词: YOLOv8; 目标检测; 可变形卷积; BiFPN; 轻量化

中图分类号: TN911.73 文献标识码: A 国家标准学科分类代码: 520.6040

Fabric seams detection algorithms based on improved YOLOv8

Liu Ruiming Chen Lun'ao Xu Chunrong Zhou Tao

(School of Electronic Engineering, Jiangsu Ocean University, Lianyungang 222005, China)

Abstract: Fabric seams detection in industrial setting is becoming increasingly important in textile applications. However, seam detection faces challenges such as small target size, few available features, and complex environmental factors, which make it difficult to ensure stable and real-time detection results. A fabric seam detection algorithm YOLOv8-DVB based on improved YOLOv8 is proposed to address this series of problems. The C2f module is optimized based on the characteristics of Deformable Convolutional Networks v4, a C2f-DCN module with multi-size feature sampling is proposed to strengthen the network's extraction of feature information of different sizes. In the neck, the BiFPN structure is used as a feature fusion approach, which allows features of different scales to be more fully fused at multiple levels by introducing top-down and bottom-up bidirectional pathways. Additionally, a more efficient VoV-GSCSP module is introduced to lightweight the feature fusion network, which helps the neck network to reduce the computational load and parameter count. Finally, a dedicated small target detection layer is designed to optimize the feature extraction of small targets. YOLOv8-DVB is compared with the original model as well as YOLOv5, YOLOv7 and Faster R-CNN through experiments to verify the detection accuracy and detection precision. The experimental results show that the method obtains 84.7% detection accuracy on the self-constructed dataset, which is higher than the original model and other network models, and is able to quickly and effectively accurately detect the target categories and locations in complex industrial environments.

Keywords: YOLOv8; object detection; deformable convolution; BiFPN; lightweighting

0 引言

纺织行业作为规模庞大的传统企业,同时也是我国消

费品三大支柱产业之一,在我国经济发展中起着举足轻重的作用。就目前情况而言,我国的纺织工业已经形成了世界上规模最大、最完善的工业生产体系。生产制造加工能

收稿日期:2024-10-12

*基金项目:国家自然科学基金(62271236)项目资助

力以及国际贸易体量长期居于世界首位。2022 年,中国规模以上纺织企业营业收入 52 564 亿元,同比增长 0.9%,利润总额 2 067 亿元^[1]。

在布匹的生产过程中,缝头位置对生产质量有着重要的影响,工业上存在通过检测缝头的位置来进行躲刀的应用,在毛毯的生产过程中,需要将毛毯上的长短不一的绒毛过一次滚刀,从而使绒毛的长度保持一致,直至运行到缝头位置时滚刀抬起,目前这一环节大部分是靠人手工控制,效率低下且可能存在误检或者错检,很难保证在复杂的工业环境下稳定且实时的进行布匹缝头检测的效果。除此之外还有通过检测布匹缝头来监控布匹数量的应用^[2]。这些都只是纺织工业中众多需要进行缝头检测的场景的一部分。所以对于缝头检测的研究能够帮助纺织工业实现进一步的自动化,提高生产效率,促进整个纺织工业体系更加系统化,对于推动纺织工业发展有积极意义。但是想要实现自动化的缝头检测依然存在许多研究难点,首先是工业环境复杂,各式各样的灯光以及设备上的玻璃等反光材料都会加深场景的复杂性,对数据的采集以及后续的识别都会造成干扰;其次是缝头检测与常见的检测目标不同,布匹缝头一般是非常细长的目标,长宽比较大,没有固定的形状。且存在布匹拼接时缝合过紧导致缝头位置较深的情况,往往只有一小部分漏出来,这一类的缝头目标小特征少,进一步增加了识别的难度。虽然近些年以来,随着图像处理技术的不断进步,基于图像处理的布匹瑕疵识别技术^[3]得到了快速发展。但是依然存在着在复杂环境下效果不理想的情况。传统的缝头检测对于不同的布料,因其纹理不同可能需要更改参数,或者对光线变化的适应能力弱难以保证检测效果。

目前基于深度学习的目标检测主要分为两大类:单阶段算法^[4]和两阶段算法^[5]。单阶段检测器目前比较流行的网络包括单步多框目标检测^[6](single shot multiBox detector, SSD)算法和 YOLO^[7](you only look once)系列。SSD 的出现因其检测速度快且又能保持很高的准确率而备受关注,它的结构相比于其他网络的优势在于它能够在单个网络中直接预测出目标的预测框和类别概率,不需要提前生成候选区域。相比于 SSD 倾重于平衡速度和准确率,YOLO 则侧重于极致的速度。YOLO 把目标检测作为一个回归问题,通过图像像素来预测边界框和类别概率,这样的设计能够实现非常高的检测速度,更适合实时应用^[8];两阶段代表算法包括基于区域的卷积神经网络^[9](region-based convolutional neural network, R-CNN)以及在此基础上改进得到的 Fast R-CNN^[10]、Faster R-CNN^[11]等。R-CNN 在 2014 年由 Ross Girshick 等提出。R-CNN 会预先定义一个候选区域,然后利用高容量的卷积神经网络从候选区域中提取特征,从而达到非常好的检测效果。2015 年由 Ren 等提出了 Faster R-CNN, Faster R-CNN 在 Fast R-CNN 的基础上引入了区域生成网络^[12](region proposal

network, RPN), 主要的作用是可以帮助网络自动生成高质量的候选区域,而且 RPN 可以与网络共享全图的卷积特征,从而大大加速了候选区域的生成,保证了整个目标检测网络既能快速处理,又可以保持高准确率。

针对缝头检测的生产需求,需要检测速度达到实时或者近实时。虽然两阶段算法系列的 Faster R-CNN 在检测速度方面相比于之前的版本有了长足的进步,但是依然会慢于单阶段算法,而单阶段算法中 YOLO 是其中的速度最快的算法之一,更适合于本课题的应用需求。YOLO 模型的核心思想是将目标检测问题简化为一个端到端的回归问题,通过一次性前向传播就可以完成对图像中多个目标的检测。这个系列从 16 年诞生起到现在仍然在不停的更新迭代,作为 YOLO 系列表现比较出色的算法模型, YOLOv8^[13]相比于之前版本的 YOLO,采用了更高效的特征提取和目标检测方法。YOLOv8 通过对主干网络的改进: Focus 结构优化、C3 (CSP bottleneck with 3 convolutions) 模块^[14]替换成 C2f(faster implementation of CSP bottleneck with 2 convolutions) 模块^[15]等,在优化网络结构获得轻量化的同时,又可以获得丰富的梯度流信息,能够更好的提取特征,非常适合用于提取布匹缝头的特征信息。而且将检测头换成解耦检测头进一步加快了收敛,提高了检测精度,能够更轻松的应对复杂环境下的目标检测,故选用 YOLOv8 作为本文的基准网络。

为了提高缝头在复杂环境下的检测准确率,本文采用改进的 YOLOv8 算法对缝头进行检测。首先,需要建立缝头的数据集,收集工业生产环境下的布匹缝头图片,通过对图片的数据增强后得到布匹缝头的数据集。之后对 YOLOv8 的主干网络进行了改进,根据可变形卷积网络^[16](deformable convolutional networks v4, DCNv4)在视觉任务中的出色表现,结合其优点改进 C2f 模块,提出 C2f-DCNv4 模块,强化网络对不同尺寸特征的提取。在 YOLOv8 颈部网络采用加权双向金字塔(bidirectional feature pyramid network, BiFPN)^[17]结构作为特征融合的方式,通过这种双向架构,将低层的细节特征和高层的语义信息能够有效地相互影响。其次在颈部网络引入一次聚合的跨阶段局部网络^[18](VoV-GSCSP)模块,减少模型的参数量,在一定程度上解决了过拟合的问题。最后设计了一个小目标检测层,用于提取小尺寸的特征信息,提高对小目标检测的精确度。根据改进后的 YOLOv8-DVB 模型在训练集以及测试集上的表现,评估模型的性能,与 YOLOv8n 模型相比较,模型的检测精度和收敛速度以及鲁棒性都得到了提升。

1 相关原理

1.1 YOLOv8 算法介绍

YOLOv8 算法作为 YOLO 系列表现非常出色的算法之一,由 Ultralytics 在 2023 年发表的。作为 YOLO 系列非常出色的算法模型之一,它具备 YOLO 系列其他算法的

优点,并且支持所有以前的版本,可以在众多版本中随意转换,其网络结构如图 1 所示。

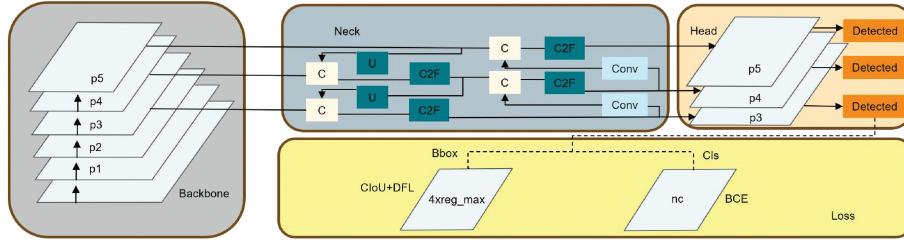


图 1 YOLOv8 结构图

Fig. 1 YOLOv8 structure diagram

YOLOv8 相比之前的版本有几个较大的改动。对于头部网络部分,使用了解耦头^[19]的设计,将分类和回归任务解耦开来,保证分类和回归任务的梯度不会互相干扰,从而提高了检测的精度。然后引入了 Anchor-Free^[20]机制,不再依赖预定义的锚框。通过直接预测目标框的中心点尺寸,避免了锚框尺寸选择的问题,简化了模型设计,并减少了后续处理步骤,在保持高精度的同时,进一步提高了速度和效率。其次是骨干网络方面,YOLOv8 沿用跨阶段局部网络(cross stage partial network,CSPNet)^[21]的思想,使用 C2f 模块替换掉了 C3 模块,这样既保留了丰富的梯度流信息,又能很好的缩减骨干网络中最大 stage 的 blocks 数,实现模型的轻量化。最后在损失函数方面,主要体现在采用了改进的交并比(intersection over union, IoU)损失^[22]来衡量预测边界框和真实框之间的差异,主要使用 CIoU、DIoU、GIoU。在分类损失方面引入了 Focal Loss^[23],解决类别不平衡问题。并且引入了动态标签分配策略,用于更智能地分配训练过程中的正负样本,提高了目标定位的准确性。

除了以上的这些改进之外,其他的结构基本延续之前的版本,骨干网络上依然是以快速空间金字塔池化(spatial pyramid pooling-fast,SPPF)^[24]模块作为最后一层,在颈部网络依然选用特征金字塔网络(feature pyramid network,FPN)和路径聚合网络(path aggregation network,PAN)^[25]结合作为特征融合方法,整体的网络保持很高的灵活性。

1.2 算法存在的问题

综合考虑算法应用的实时性和精确度之后,本文采用计算量和参数量较小且速度适中的 YOLOv8n 作为算法模型。但是在现实场景的应用过程中发现该模型存在一些问题。首先,YOLOv8n 算法使用传统卷积提取动态尺寸特征时,适应能力差,导致在提取不同尺寸大小的缝头的特征时,效果欠佳。其次,模型在复杂背景下对缝头的实时检测不理想,检测速度依然有待提升,且对于小目标的检测结果也存在进步空间,容易出现错检漏检等问题。

2 改进模型

针对以上 YOLOv8n 模型存在的不足,提出一种改进

网络模型 YOLOv8-DVB。图 2 为 YOLOv8-DVB 模型结构图。首先将骨干网络中最后一层的 C2f 模块替换为 C2f-DCNv4 结构。其次使用跨阶段部局部网络模块 VoV-GSCSP 将颈部网络的 C2f 模块替换。然后在 YOLOv8n 的颈部网络使用 BiFPN 实现特征融合。最后在颈部网络设计一个小目标检测层。C2f-DCNv4 模块增强网络对于动态特征的适应能力。VoV-GSCSP 模块采用更合理的架构和多尺度信息融合来减少参数量和计算量,能够在不影响精确度的情况下提高速度。BiFPN 可以减少因为比例变化和复杂场景造成的检测误差,从而增强模型的通用性。小目标检测层针对网络对小目标检测效果差而设计的,提升网络对小尺寸特征的提取能力。

2.1 C2f-DCNv4

传统卷积在进行卷积操作时,采样点是固定且规则的,这样设计的好处就是可以很好的处理静态的或者有一定规则的目标,但是布匹缝头检测一般都是动态任务,且不同种类布匹间的缝头也存在着尺寸差异,传统卷积针对此类检测任务可能效果不佳。而可变形卷积主要针对检测目标的尺寸和位置进行动态采样,从而能够更好的提取不同尺寸目标的特征。

DCNv4 是一种改进的卷积神经网络架构,其目的在于提高神经网络卷积操作的动态性能,主要通过引入可变形卷积和可变形感兴趣区域(region of interest pooling,ROI)池化^[26]来增强模型的表现。传统卷积核在进行卷积操作时,其采样点是固定且规则的。在可变形卷积中,通过引入一个偏移量(offset)学习机制,每个卷积核采样点都可以从原始位置出发自由以移动,从而达到变形的目的。这种方式的卷积可以很好的应对目标的几何形变,更好的捕捉目标的形状与姿态。对于可变形卷积,假设输入特征图 $x \in R^{H \times W \times C}$,首先将输入特征图按照通道数分成多个组,每个组内会进行独立的卷积操作。对于分组后的特征图使用普通卷积层对其进行预处理,计算出偏移量 Δp_{gk} 和调制量 Δm_{gk} ,然后根据偏移量将规则分布的采样点进行移动。在得到所有动态调整后的采样点值后,对输入特征图进行卷积操作,得到这一组的输出特征图。该过程如图 3 所示。

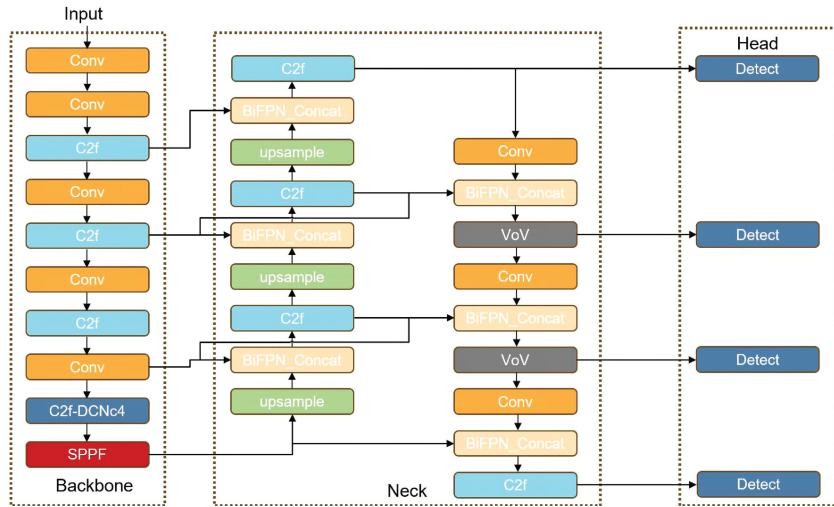


图 2 YOLOv8-DVB 结构图

Fig. 2 YOLOv8-DVB structure diagram

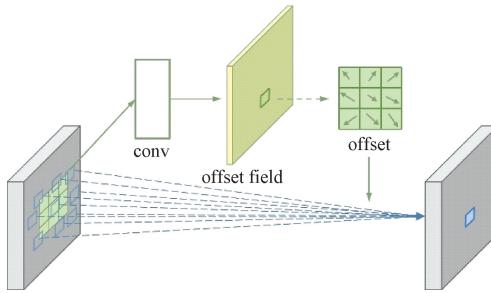


图 3 可变形卷积过程

Fig. 3 Deformable convolutional process

在得到每个组的输出后,在通道维度上将这些组的输出拼接起来形成最终的输出特征图,整个过程用公式描述如式(1)和(2)所示。

$$y_g = \sum_{k=1}^K m_{gk} x_g (p_0 + p_k + \Delta p_{gk}), \quad (1)$$

$$y = concat([y_1, y_2, \dots, y_G], axis = -1), \quad (2)$$

其中, g 表示组数, x_g 和 y_g 表示第 g 组的输入和输出特征图, K 表示采样点个数, Δp_{gk} 和 m_{gk} 表示第 g 组第 k 个采样点的偏移量和调制量, p_0 则表示卷积操作的中心点, 最终所有组的输出 y_g 在通道维度上拼接得到最终的输出特征图 y 。

DCNv4 同时还引入了可变形 ROI 池化, 池化层作为卷积神经网络中用于数据降维的有效手段, 对于特征提取至关重要。传统池化层主要的池化操作是在固定大小的窗口内进行最大池化或者平均池化, 以实现对特征图的下采样操作。而可变形 ROI 池化首先会针对输入特征图先生成感兴趣区域 ROI, 将每个 ROI 区域分为 $H \times W$ 的固定网格, 针对每一个网格的位置, 通过全连接层生成归一化的偏移量 Δp_k , 然后根据偏移量的大小对网格位置进行移动。通过上一步的移动得到了新的采样点的值, 对每个网

格采样点的采样特征值进行最大池化操作或者平均池化操作, 获得该感兴趣区域的输出, 最后将所有的感兴趣区域的池化结果按照顺序拼接后得到最终输出, 过程如图 4 所示。与传统池化相比而言, 可变形 ROI 池化能够更好地适应物体的复杂形变、局部扭曲、弯曲等情况, 保留更多的细节和结构信息, 尤其在处理复杂背景或非刚性物体时表现更佳。

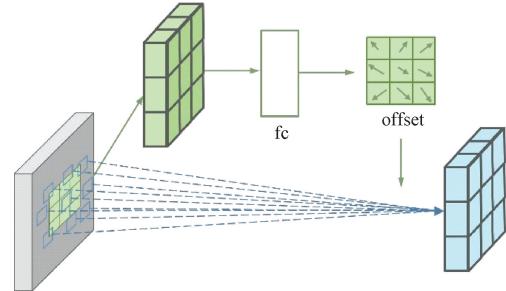


图 4 可变形 ROI 池化过程

Fig. 4 Deformable ROI pooling process

C2f 模块的主要思想是通过多尺度特征提取, 逐步从粗层级特征到细层级特征进行建模。粗层级特征指的是在较大尺度下捕捉整体的图像结构和轮廓信息, 细层级特征指的是在小尺度下聚焦与细节信息, 例如边缘、纹理等。这种从粗到细的处理方式有助于网络捕捉全局信息的同时也能保留细节特征。而 DCNv4 通过引入偏移量动态调整卷积核的采样位置, 能够适应输入图像中的几何变形和不规则形状, 更好的适应不同尺度的目标。如果将二者结合起来, 能够更有效的捕捉目标物体的局部特征, 同时保持对空间变形的鲁棒性。

为了有效结合 C2f 模块和 DCNv4, 需要先确定结合的位置。对于骨干网络而言, 浅层网络主要作用是提取基础的边缘信息和纹理信息, 此时几何变形不明显, 不需要

DCNv4 的灵活性,适合传统卷积。而深层网络主要负责处理更高阶的特征,比如物体的形状和姿态等,此时 DCNv4 的动态性能就能很好的适应复杂的几何变形和多样性。在此基础上,只需对深层的 C2f 模块进行改进即可。本文选择将第 8 层的 C2f 模块改进为 C2f-DCNv4,将 C2f 模块中 bottleneck 一个 3×3 的卷积层改为可变形卷积,并使用可变形 RoI 池化对特征图进行下采样,改进后的 C2f-DCNv4 结构如图 5 所示。

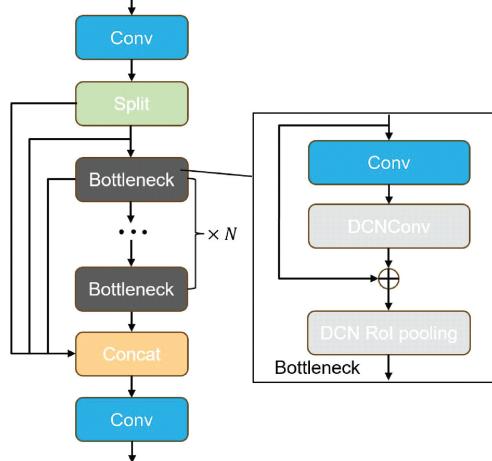


图 5 C2f-DCNv4 结构

Fig. 5 Structure of C2f-DCNv4

Bottleneck 结构^[27]中包含有两个卷积层,第 1 个卷积层的作用主要是通过通道数减半的方法来降低参数量,为第 2 层卷积提供了预处理后的基础特征。而第 2 层卷积层是 bottleneck 结构中最主要的特征提取层,具备较大的感受野,负责捕捉局部和全局特征。通过将可变形卷积引入第 2 层卷积层,使得卷积的采样位置可以根据输入图像的特征动态调整,这使得网络在处理不规则目标、复杂背景或变形目标时更具鲁棒性。同时在 bottleneck 中,由于第 1 层的降维操作已经减少了通道数,意味着计算量相对而言有一定的降低。而在第 2 层加入可变形卷积后,不会显著增加计算负担,却可以有效提升网络在复杂场景下的特征提取能力。最后为了进一步增强模型对目标尺寸变化的适应能力,加入了可变形 RoI 池化,能够进一步调整池化窗口,从而在目标区域内更加精准地提取特征。这样组合能显著提升网络对于目标检测这类任务的精度和鲁棒性,特别是在处理形变大,复杂背景或多尺度的目标时,效果尤为显著。

2.2 VoV-GSCSP 模块

为了提高卷积神经网络的计算速度,通常需要在骨干网络部分对输入图像进行一系列的预处理,将输入图像的空间信息转换为信道信息。但是在特征每次升维和降维的过程中,空间维度的压缩和通道数的变化都有可能导致部分语义信息丢失,从而大大降低准确率。

针对这个问题,本文引入一种轻量级卷积 GSConv,该

卷积既能够降低计算复杂度,又可以最大限度的保留通道之间的连接性,从而改善特征融合。其结构如图 6 所示。其中 Conv 表示普通卷积,用于生成少量的特征图。DWConv^[28]表示深度卷积,是深度可分离卷积的一个组成部分,用于生成附加的 Ghost 特征图。最后标准卷积生成的特征图和 Ghost 特征图通过拼接形成最终输出。这些特征图综合了重要的语义信息,同时减少了计算成本。

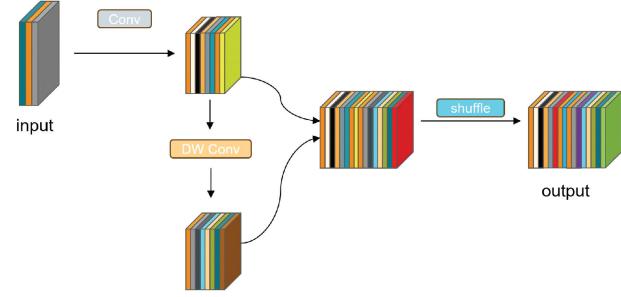


图 6 GSConv 结构

Fig. 6 Structure of GSConv

图 7 表示的是基于 GSConv 的瓶颈层 GS Bottleneck^[29]。GS Bottleneck 模块通过 Ghost Module 的设计思想来减少冗余计算。传统的 Bottleneck 模块通常包含多层卷积和特征图缩减操作,而 GS Bottleneck 使用 GSConv 来生成较少但有效的特征,使模型在实现轻量化的同时仍能实现出色的特征提取。

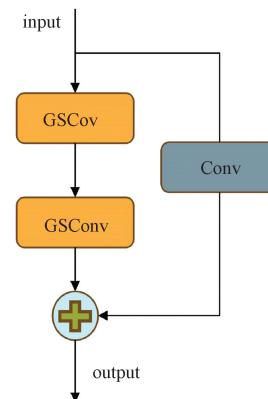


图 7 GS bottleneck 结构

Fig. 7 Structure of GS bottleneck

图 8 表示的是以 GSConv 和 GS Bottleneck 为基础构建的 VoV-GSCSP 模块。VoV-GSCSP 模块利用残差连接来淡化输入的特征图,该结构包含两个分支,其分支 1 经过普通卷积和多个残差结构的 GS Bottleneck,分支 2 经过普通卷积后与特征图连接,目的是利用连接以及添加特征融合策略来增加小目标的检测效果并降低计算成本。VoV-GSCSP 模块的使用不仅减少了参数量,同时在一定程度上解决了过拟合问题,增强了网络的非线性表示能力。在 YOLOv8 的颈部合理使用 VOV-GSCSP 不仅减少了模型的参数量和计算量,还一定程度上提高了模型检测的准确率。

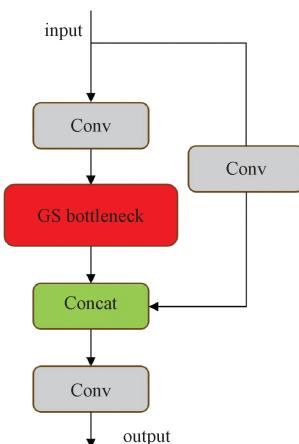


图 8 VoV-GSCSP 结构
Fig. 8 Structure of VoV-GSCSP

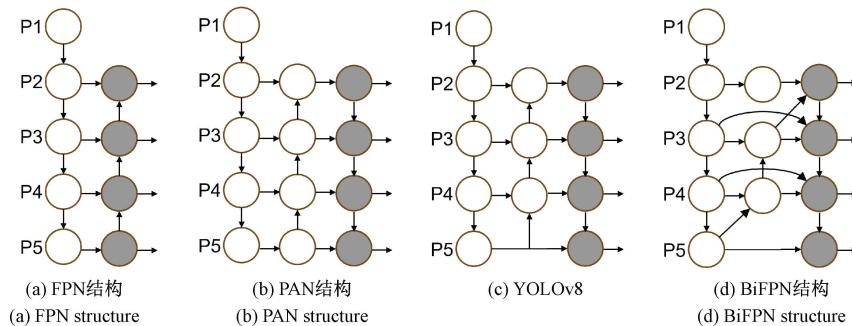


图 9 特征网络
Fig. 9 Feature network

YOLOv8 采用的特征融合方式为 FPN+PAN，通过将两者结合，FPN 提供了自顶向下的特征融合，PAN 提供了自底向上的特征汇聚，这使得最终的特征图能够同时包含高层语义信息和低层细节信息。其结构图如图 9(c)所示。

然而，所有的特征融合方式都对小目标的识别检测效果较差。因为在特征提取的过程中，小目标很容易受正常大小目标影响。网络会删除不明显目标，使得小目标信息不断减少，最终导致对小目标检测的漏检。为了解决这一问题，本文使用了 BiFPN 作为特征融合的方式。BiFPN 结构图如图 9(d)所示。

BiFPN 引入了可学习权重，允许网络自适应地为不同层次的特征分配权重，很好的解决了 FPN 或 PAN 在特征融合时对所有特征等权处理的问题。加权机制通过学习不同尺度特征的权重，在特征融合时根据重要性调整它们的贡献，这样就可以确保在特征融合的过程中，重要的特征信息不被稀释。

2.4 小目标检测层

为了进一步提升对小目标检测的效果，本文增加了一个小目标层和一个 P2 检测头。YOLOv8 网络结构中包含有 3 个不用缩放比例的特征图，分别用于检测小型、中型和大型目标。随着网络深度的增加，特征图变得更小、更

2.3 BiFPN

为了融合不同尺度的特征信息，YOLOv8 中的颈部网络使用 FPN+PAN^[30]的结构作为特征融合的方式。FPN 作为一种灵活高效的多尺度特征提取架构，广泛用于目标检测、语义分割等计算机视觉任务。

FPN 通过自顶向下的相同路径，成一系列具有通道数但不同空间分辨率的特征图，每个特征图代表不同的几何图形。这些几何图形组成了一个特征金字塔，能够处理不同尺寸的特征图的目标。其结构如图 9(a)所示。而 PAN 则是在继承 FPN 的自顶向下路径的基础上引入了自底向上的路径，该路径通过逐层下采样的方式，将浅层的细节特征传递至深层。这样不仅可以使浅层的细节信息被深入掌握，也能够提升模型针对大目标的检测能力。PAN 的结构如图 9(b)所示。

抽象，并包含更多语义信息。小尺寸特征图通常用于检测大型目标，因为越往高层，特征图拥有的感受野就越大。而大尺寸特征图在定位目标方面更为精确，一般用于检测小目标。因此，在颈部结构加入一个小目标检测层，获取更大尺寸的特征图，以提高网络检测小型目标的能力。改进后的颈部网络结构如图 10 所示。未改动之前，YOLOv8 对应的特征图大小分别为 20×20 、 40×40 、 80×80 ，检测尺寸大小分别为 32×32 、 16×16 、 8×8 的目标。增加小目标检测层之后，在原有结构的基础上增加了 4 倍下采样以及相应的 160×160 输出映射。

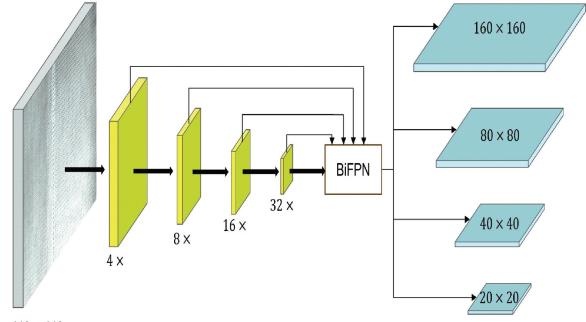


图 10 加入小目标检测层

Fig. 10 Add a small target detection layer

3 实验与结果

3.1 实验环境

本实验用到的计算机操作系统为 Windows11 64 位, 处理器型号是 i7-11800H, 显卡型号为 NVIDIA GeForce RTX3060, 主频 2.3 GHz, 16 G 机带 RAM, CUDA 版本为 11.8, 编程语言为 Python3.8, 深度学习框架式 Pytorch2.0.1。

初始学习率设定为 0.01, 批大小设定为 8, 轮次为 200 轮, 输入图像尺寸为 640×640 , 权重衰减系数为 0.000 05, IoU 阈值设定为 0.5。

3.2 数据集和预处理

由于目前这方面的研究较少, 所以不存在这方面的开源数据集, 本文所使用的数据集是自建数据集。在不同的工业环境下采集各种布匹缝头的图片 419 张, 标签为: 缝头(Stitch), 部分布匹缝头的图例如图 11 所示。

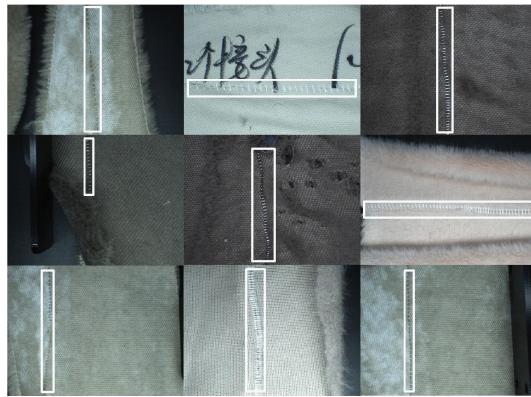


图 11 缝头样本

Fig. 11 Seam samples

为了进一步提升数据量, 扩充数据集, 对原有的数据集进行数据增强操作。将翻转、裁剪、平移、调整亮度和添加噪声的方式随机组合对数据集进行扩充, 扩充后的数据集图片数量达到 838 张。通过扩充数据集, 能进一步提升模型的泛化能力, 减少过拟合出现的可能。将扩充后的数据集按照 7:2:1 的比例划分为训练集、测试集和验证集。

本文数据集所用图片大多为工业环境下的布匹缝头图片, 图片大小都调整为了 640×640 。数据集通过增强后总共是 836 张, 然后对数据集中的图片进行标注。训练时采用的是 Pytorch 框架, 使用 YOLOv8n 原模型和 YOLOv8-DVB 进行训练, 在数据集中训练 200 个轮次, 超过 20 个轮次训练参数没有得到改善的话, 就直接结束训练。

3.3 评价指标

为了更好的对改进后的模型进行检验, 选取召回率 R(Recall)、准确率 P(Precision)、平均精度 mAP(Mean average precision)、参数量(Params)、Giga 浮点运算每秒(GFLOPs)以及每秒预测的帧数 FPS 作为评价指标。

3.4 C2f-DCNv4 模块有效性分析

C2f-DCNv4 模块能够利用偏移量实现卷积采样点的动态调整, 可以更好的适应不同尺寸特征的提取, 极大地增强了模型的鲁棒性。为了测试 C2f-DCNv4 模块在骨干网络中的位置对于模型检测效果的影响, 本文设计如表 1 所示的有效性实验。对于骨干网络, 深层网络一般需要处理更高阶的特征, 更适合 C2f-DCNv4 的发挥。C2f-DCNv4-1 表示将骨干网络中第 1 个 C2f 层替换为 C2f-DCNv4 模块, 其他结构不做变动。以此类推, 后面的 C2f-DCNv4-2、C2f-DCNv4-3、C2f-DCNv4-4 依次是将骨干网络中第 2、3、4 个 C2f 层替换为 C2f-DCNv4 模块, 其他位置保持原有结构。

由表 1 可知, 在骨干网络中不同位置的 C2f 层引入 DCNv4 后, 都能够实现降低模型的计算成本的同时提高模型检测的检测效果。而且, 随着引入 DCNv4 的 C2f 层不断加深, 检测精度也在逐步上升。这是因为 DCNv4 能够对目标形变做出动态调整, 相比于标准卷积可以更有效的提取到目标特征, 使得模型可以更加准确的拟合不同形状大小的布匹缝头, 提升模型对形状尺寸差别较大的缝头检测能力。同时可以看到计算量和参数量都在下降, 这是因为 C2f-DCNv4 通过灵活的采样点分布, 能够更有效的捕捉有意义的特征, 减少不必要的特征计算和冗余卷积操作, 从而减少模型的参数量, 达到降低计算量的结果。FPS 的增加意味着每秒能够预测的帧数也在增加, 说明随着 C2f-DCNv4 层次的加深, 单张的预测时间也在缩短, 进一步验证了模型的有效性。

表 1 C2f-DCNv4 模块有效性实验

Table 1 C2f-DCNv4 module validity experiments

模型	mAP/%	Params/M	GFLOPs	FPS
YOLOv8n	77.0	3.51	7.5	36
+C2f-DCNv4-1	78.1	3.33	7.0	39
+C2f-DCNv4-2	78.0	3.34	7.2	40
+C2f-DCNv4-3	79.1	3.16	7.1	45
+C2f-DCNv4-4	80.3	3.09	6.8	47

3.5 VoV-GSCSP 模块有效性分析

VoV-GSCSP 模块是由 GS Bottleneck 模块构成, GS Bottleneck 模块则是以 GSConv 为基础构成。GSConv 能够通过减少不必要的冗余计算, 保持模型的轻量化设计。GSConv 卷积核大小可以是 3×3 或 1×1 , 可以根据具体的应用需求而定。

表 2 为 GSConv 卷积大小对比实验, 用于验证 VoV-GSCSP 模块中卷积设计的有效性。GSConv 能够通过深度卷积这种廉价操作来从这些基本特征生产 Ghost 特征, 扩充通道。这些特征与基本特征有很高的相似性, 但是计算成本大大降低, 使得模型在保持性能的同时大大降低了

计算开销。同时,卷积核的尺寸也会对计算成本造成影响,由表 2 可知当 GSConv 卷积核大小越小时,能进一步降低模型的计算负担的同时保持较好的性能。

3.6 消融实验

为了验证本课题所用到的改进方法对于模型性能的影响,根据改进方案进行了消融实验,由于本文所用的实验模型是由 YOLOv8n 改进而来,消融实验选取 YOLOv8n 原模型进行对比,得到消融实验结果如表 3 所示。

表 3 YOLOv8 的不同改进消融实验结果

Table 3 Results of different improved ablation test of YOLOv8

C2f-DCNv4	VoV-GSCSP	BiFPN	P2	P/%	R/%	mAP/%	Params/M	GFLOPs	权重文件/MB
✓				77.0	65.4	73.8	3.51	7.5	8.1
	✓			80.3	73.1	83.9	3.09	6.8	7.5
		✓		77.6	73.2	79.2	3.02	5.5	7.0
			✓	79.1	74.7	76.7	3.54	7.6	8.3
				✓	80.2	76.2	75.2	2.98	9.3
✓	✓			81.3	83.1	76.4	2.81	6.2	6.7
✓	✓	✓		80.7	83.2	77.6	3.35	6.6	6.9
✓	✓		✓	83.5	79.3	80.4	3.12	7.6	8.1
✓		✓	✓	83.8	84.1	80.9	3.84	6.4	6.8
✓	✓	✓	✓	84.9	73.2	84.7	2.77	5.3	6.2

根据表 3 可以看出,本文在 YOLOv8n 模型的基础上进行改进,将模型的 C2f 结构与 DCNv4 模块进行融合,进一步提升卷积核对复杂目标几何形变的适应能力,能够更好地捕捉到局部特征的变化,从而增强模型对复杂目标检测的准确率。相较于原模型,改进后的模型 mAP 提升了 3.3%,同时模型的计算量减少了 9.33%,模型参数量减少了 11.97%,模型大小降低 7.4%;通过在颈部网络加入 VoV-GSCSP 模块来实现模型的轻量化。VoV-GSCSP 模块通过 GSConv 卷积利用廉价的深度卷积或者线性变换操作从基本特征中生成 Ghost 特征,利用 Ghost 特征填补剩余通道。这些附加特征与基本特征具有较高的相似性,但是计算成本却大大降低。在少量提升 mAP 的同时模型计算量下降了 26.67%,参数量减少了 11.97%,模型大小减少了 13.58%;将原来的特征融合方式由 FPN+PAN 改为 BiFPN,BiFPN 的核心思想是双向特征融合,即允许特征在自顶向下和自底向上两个方向上传递。同时 BiFPN 在特征融合时引入加权机制,允许网络自适应地为不用层次的特征分配权重,这样可以确保在特征融合的过程中,重要的特征信息不会被稀释。能够在牺牲极少的参数量和计算量的代价下将 mAP 提升 2.1%。最后通过增加小目标检测层,可以有效的增强模型对小尺寸特征的提取能力,提升模型对小目标检测的精确度。小目标检测层的加入,会导致计算量出现 24% 的上升,同时能够将 mAP 提升 3.2%,并且模型参数量降低 15.1%,模型大小也下降了

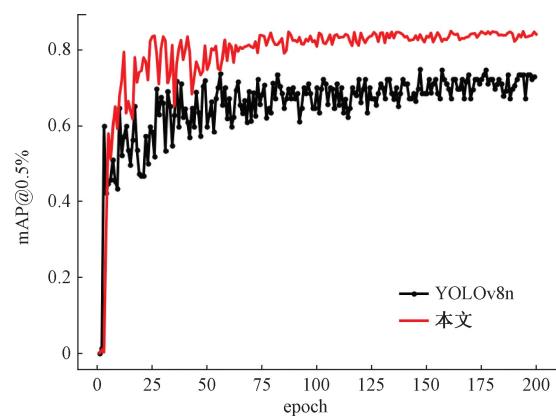
表 2 GSConv 卷积核大小对比实验

Table 2 GSConv convolution kernel size comparison experiments

模型	mAP/%	Params/M	GFLOPs
YOLOv8n	77.0	3.51	7.5
GSConv(3×3)	77.4	3.27	6.1
GSConv(1×1)	77.6	3.02	5.5

9.88%,总体而言牺牲了部分计算成本,换取较高的性能提升和模型的轻量化。组合使用多个模块,将多个模块结合起来取长补短,能够提升 mAP 的同时帮助模型实现轻量化。改进后的模型在平均精度和轻量化之间都有所提升,同时兼顾了性能和轻量化之间的平衡。

图 12(a)和(b)分别为模型改进前后的 mAP@0.5% 和 Loss 对比曲线图,从图 12(a)可以看出,改进后的模型不仅平均精度更高,同时曲线也更稳定。图 12(b)的 Loss 曲线则反应了改进后的模型有着更快的收敛速度,并且与原模型相比也更加平滑。



(a) 平均精度@0.5%对比曲线

(a) mAP@0.5% comparison curve

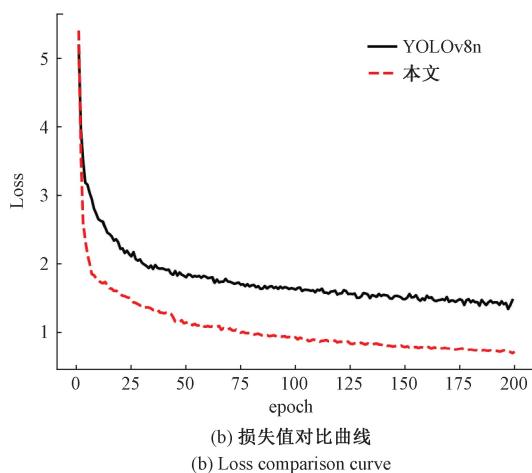


图 12 改进前后 mAP@0.5% 与损失对比曲线

Fig. 12 Comparison curve of mAP@0.5% and loss before and after improvement

3.7 不同模型的对比实验分析

为了进一步验证本文提出的 YOLOv8-DVB 模型的性能, 对比目前比较主流的几种卷积神经网络模型: YOLOv7、YOLOv5、YOLOX-X 以及 Faster R-CNN 等, 其中 YOLOv8-FastNet 和 YOLOv8-MobileNet 是两种在 YOLOv8 基础上改进效果比较好的模型, 分别将主干网络替换成了 FastNet 和 MobileNet。在参数一致的情况下, 在自建布匹缝头数据集上进行训练和测试, 最后得到各个网络的性能结果如表 4 所示。

与同是 YOLO 系列的 YOLOv5n、YOLOv7n 和 YOLOv8n 等轻量级算法模型比较, YOLOv8-DVB 算法能够在不影响模型的轻量化的情况下提升模型的检测性能, 准确率与之相比分别提升了 20%、17.3%、12.2%、7.9%, 召回率分别提升了 17.7%、16.7%、2.5%、7.8%, mAP 提升了 19.9%、16.5%、9.5%、10.9%。各方面相比于原模型都有大幅度的提升。与 YOLO 系列的改进版本 YOLOX

表 4 不同模型的缝头检测对比结果

Table 4 Comparative results of stitch detection with different models

模型	P/%	R/%	mAP/%	Params/M	GFLOPs	权重文件/MB
YOLOv3n	64.9	55.7	64.8	2.77	5.3	6.2
YOLOv5n	67.6	56.5	68.2	2.13	4.9	5.4
YOLOv7n	72.7	70.7	75.2	3.66	10.4	7.5
YOLOv8n	77.0	65.4	73.8	3.51	7.5	8.1
YOLOX-X	83.2	79.1	82.5	78.14	251.9	350.7
YOLOv8-FastNet	69.0	73.4	72.1	1.80	5.1	7.6
YOLOv8-MobileNet	76.7	77.3	72.7	2.90	5.7	9.3
SSD	67.5	60.5	67.9	102.79	272.8	67.4
RT-DETR	85.7	84.1	85.1	19.20	49.0	36.4
Faster R-CNN	54.3	63.2	56.5	13.75	37.0	371.2
YOLOv8-DVB	84.9	73.2	84.7	2.77	5.3	6.2

模型相比较, mAP 虽然提升不高, 但是模型的轻量化程度和计算量却远远优于 YOLOX。与替换了主干网络的 YOLOv8-FastNet 和 YOLOv8-MobileNet 相比, 虽然 YOLOv8-FastNet 参数量更少, 计算成本更低, 但是 YOLOv8-DVB 能以将 mAP 提升 12.6% 来弥补这一点。相比之下, 显然 YOLOv8-DVB 的综合性能更加优越。与比较新的实时检测模型 RT-DETR 进行比较, RT-DETR 模型虽然 mAP 要高出 0.4%, 但是参数量和计算量都有比

较大的劣势。与同为单目标检测的经典算法 SSD 相比较, 各方面优势都很大。与双阶段目标检测算法中的 Faster R-CNN 相比较也是处于优势。

实验结果表明, 与目前比较主流的算法以及经典算法表, YOLOv8-DVB 算法能够以比较低的计算量和参数量达到相当不错的性能和检测精度, 具有高效性和优异性。

3.8 实验检测效果验证

如图 13 所示, 分别是 YOLOv8n 模型和 YOLOv8-

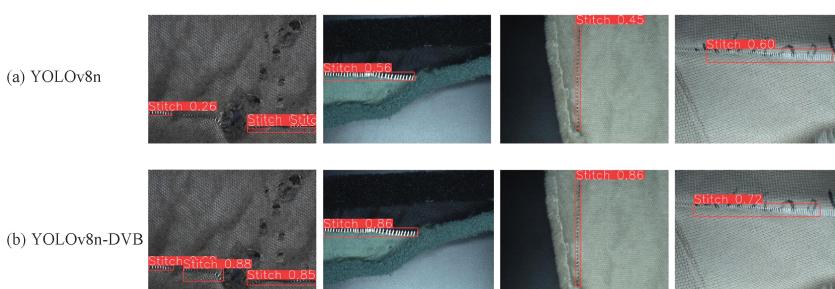


图 13 检测效果对比

Fig. 13 Comparison of detection effect

DVB 模型的检测结果,检测目标用矩形框表示,并标注了识别的标签和置信度。

图 13(a)为 YOLOv8 原模型的检测结果,图 13(b)为 YOLOv8-DVB 模型的检测效果。可以看出,相比于原模型,本文提出的改进模型在对缝头进行检测时可以保持更高的置信度,原本的漏检得到了改善。所有的实验结果表明本文所提出的基于 YOLOv8 改进的算法模型能够在复杂的工业环境下对布匹缝头做出更加有效的检测,并且能够保持极高的准确性,具有更优秀应用价值。

4 结 论

本研究建立了布匹缝头检测数据集,在 YOLOv8 模型骨干网络上引入 DCNv4 模块对 C2f 进行改进,增强网络对不同尺寸的缝头特征信息的提取。对于 YOLOv8 的颈部网络,首先使用 VoV-GSCSP 模块作为特征的融合方式,极大的减少了参数量和计算量,减少计算负担,达到模型的轻量化。然后更换特征网络,BiFPN 包含双向特征融合,从两个方向上传递特征信息,并且可以为不同层次的特征分配权重,保证重要的特征信息不会被过滤,进一步增强模型的检测精度。最后,为模型加入一个小目标检测层,用于专门提取小尺寸缝头特征,加强模型小尺寸特征的提取能力,从而增强模型对小尺寸缝头的检测能力。实验表明,与原模型相比,改进后的 YOLOv8-DVB 检测精度更高,对于复杂工业环境下的缝头检测效果更好,同时模型的计算成本和部署成本都实现了一定程度的降低,能够很好的满足实时检测的需求。为实现缝头的自动化检测提供了一种可行的算法方案,对纺织工业发展具有积极意义。

参考文献

- [1] 程汉权,熊继平,陈经纬.布匹瑕疵检测算法研究进展[J].计算机时代,2023(11): 16-21.
CHENG H Q, XIONG J P, CHEN J W. Advances in fabric defect detection algorithms[J]. Computer Era, 2023(11): 16-21.
- [2] 李保磊.布匹自动缝头机的设计与研究[D].上海:东华大学,2022.
LI B L. Design and research of automatic cloth sewing machine[D]. Shanghai: Donghua University, 2022.
- [3] 吴志华,钟铭恩,谭佳威,等.复杂纹理布匹五类典型瑕疵图像检测算法研究[J].电子测量技术,2023, 46(16):57-63.
WU ZH H, ZHONG M E, TAN J W, et al. Research on five types of typical defects image detection algorithms for complex textured fabrics[J]. Electronic Measurement Technology, 2023, 46(16): 57-63.
- [4] 徐业东,蔡亚恒,李严,等.基于 YOLOv5s 的轻量化架空输电线路鸟巢检测网络[J].电子测量技术,2024, 47(7):138-148.
XU Y D, CAI Y H, LI Y, et al. Lightweight overhead transmission line bird's nest detection network based on YOLOv5s [J]. Electronic Measurement Technology, 2024, 47(7): 138-148.
- [5] DONG CH, ZHANG K, XIE ZH Y, et al. An improved cascade RCNN detection method for key components and defects of transmission lines[J]. IET Generation, Transmission Distribution, 2023, 17(19): 4277-4292.
- [6] 曾乾,李博.基于改进 SSD 的青瓜检测算法[J].国外电子测量技术,2023,42(4):158-165.
ZENG Q, LI B. Cucumber detection algorithm based on improved SSD[J]. Foreign Electronic Measurement Technology, 2023, 42(4): 158-165.
- [7] 刘昶成,邵文权,李玲陶.基于改进 YOLOv5s 的电力作业人员安全帽检测算法研究[J].国外电子测量技术,2024,43(2):34-42.
LIU CH CH, SHAO W Q, LI L T. Research on safety helmet detection algorithm for power operators based on improved YOLOv5s[J]. Foreign Electronic Measurement Technology, 2024, 43(2): 34-42.
- [8] 李忠科,刘小芳.基于轻量级 YOLOv8n 网络的 PCB 缺陷检测算法[J].电子测量技术,2024, 47 (4): 120-126.
LI ZH K, LIU X F. PCB defect detection algorithm based on lightweight YOLOv8n network [J]. Electronic Measurement Technology, 2024, 47 (4): 120-126.
- [9] 涂新奎,郑少武,于善虎,等.基于对称形状生成的三维目标检测网络[J].仪器仪表学报,2023, 44 (6): 252-263.
TU X K, ZHENG SH W, YU SH H, et al. 3D object detection network based on symmetric shape generation [J]. Chinese Journal of Scientific Instrument, 2023, 44(6): 252-263.
- [10] 刘毅,于畅洋,李国燕,等.UAST-RCNN:遮挡行人的目标检测算法[J].电子测量与仪器学报,2022, 36(12):168-175.
LIU Y, YU CH Y, LI G Y, et al. UAST-RCNN: Object detection algorithm for blocking pedestrians[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(12): 168-175.
- [11] REN SH 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.
- [12] WANG Y SH, YANG G, LU H. A domain-adaptive tree-crown detection and counting method based on cascade region proposal networks [J]. Algorithms, 2023, 16(7): 309.
- [13] HUSSAIN M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection [J]. Machines, 2023, 11(7): 677.
- [14] SUN H, ZHANG W ZH, YANG SH, et al. Lightweight single-stage ship object detection algorithm for unmanned surface vessels based on improved YOLOv5[J]. Sensors, 2024, 24(17): 5603.
- [15] 潘海鸿,陈希良,钱广坤,等.基于改进 YOLOV8N 的轻量化茶叶嫩芽检测方法[J].电子测量技术,2024,47(7):149-156.
- PAN H H, CHEN X L, QIAN G K, et al. Research on lightweight tea sprout detection method based on improved YOLOv8n [J]. Electronic Measurement Technology, 2024, 47(7): 149-156.
- [16] LIANG R F, CEN Y G. Radar signal classification with multi-frequency multi-scale deformable convolutional networks and attention mechanisms[J]. Remote Sensing, 2024, 16(8): 1431.
- [17] CHUSHI Y, YOAN S. SAR ship detection based on improved YOLOv5 and BiFPN [J]. ICT Express, 2024, 10(1): 28-33.
- [18] WANG W, MENG Y Z, LI SH, et al. HV-YOLOv8 by HDPconv: Better lightweight detectors for small object detection [J]. Image and Vision Computing, 2024, 147: 105052.
- [19] THAMMASANYA T, PATIAM S, RODCHAROEN E, et al. A new approach to classifying polymer type of microplastics based on Faster-RCNN-FPN and spectroscopic imagery under ultraviolet light [J]. Scientific Reports, 2024, 14(1): 3529.
- [20] NASIM M, MUMTAZ R, AHMAD M, et al. Fabric defect detection in real world manufacturing using deep learning[J]. Information, 2024, 15(8): 476.
- [21] NING T, WU W T, ZHANG J. Small object detection based on YOLOv8 in UAV perspective[J]. Pattern Analysis and Applications, 2024, 27(3): 103.
- [22] TANG H Y, LIANG SH, YAO D, et al. A visual defect detection for optics lens based on the YOLOv5-C3CA-SPPF network model [J]. Optics Express, 2023, 31(2): 2628-2643.
- [23] MA SH B, LIU Y N, ZHANG Y P. Fabric defect detection based on improved lightweight YOLOv8n [J]. Applied Sciences, 2024, 14(17): 8000.
- [24] SUN H, TAN CH H, PANG S, et al. RA-YOLOv8: An improved YOLOv8 seal text detection method[J]. Electronics, 2024, 13(15): 3001.
- [25] YANG Y CH, LIU J, HE J T. A lightweight waxberry fruit detection model based on YOLOv5[J]. IET Image Processing, 2024, 18(7): 1796-1808.
- [26] MA Z F, WU ZH P, CAO Y G. End-to-end light license plate detection and recognition method based on deep learning[J]. Electronics, 2022, 12(1): 203.
- [27] 王蕾,党时鹏,潘丰.基于卷积神经网络的隐匿性旁路预测模型[J].计算机工程,2024,50(8):40-49.
- WANG L, DANG SH P, PAN F. Model for predicting concealed accessory pathway based on convolutional neural network [J]. Computer Engineering, 2024, 50(8): 40-49.
- [28] 张磊,李熙尉,燕倩如.基于改进 YOLOv5S 的综采工作面人员检测算法[J].中国安全科学学报,2023,33(7): 82-89.
- ZHANG L, LI X W, YAN Q R. Personnel detection algorithm in fully mechanized coal face based on improved YOLOv5s[J]. China Safety Science Journal, 2023, 33(7): 82-89.
- [29] GANG CH, CHAO P ZH, YANG J, et al. SGST-YOLOv8: An improved lightweight YOLOv8 for real-time target detection for campus surveillance [J]. Applied Sciences, 2024, 14(12): 5341.
- [30] 徐芳芯,樊蝶,马小陆.面向拥挤行人检测的改进 YOLOv7 算法[J].计算机工程,2024,50(3): 250-258.
- XU F X, FAN R, MA X L. Improved YOLOv7 algorithm for crowded pedestrian detection [J]. Computer Engineering, 2024, 50(3): 250-258.

作者简介

刘瑞明,博士,教授,主要研究方向为图像处理、模式识别等。

E-mail: raymondrlmliu@126.com

陈伦奥(通信作者),硕士研究生,主要研究方向为人工智能、模式识别。

E-mail: 2315662379@qq.com

徐春融,硕士研究生,主要研究方向为模式识别。

周韬,硕士研究生,主要研究方向为缺陷检测。