DOI:10.19651/j.cnki.emt.2519447

基于 ACS-YOLO 的复杂背景下线束端子缺陷检测算法*

胡浩宇 王淑青 王许佳 夏杨威
(湖北工业大学电气与电子工程学院 武汉 430068)

摘 要:针对线束端子压接过程中,检测场景复杂导致遮挡和模糊的检测难点,提出基于改进 YOLOv11 的线束端子缺陷检测模型 ACS-YOLO。该模型设计 C2PSA_EFA 模块,结合 C2PSA 与边缘增强特征注意力机制,通过 Sobel 算子提取并融合边缘信息,增强线束端子不规则缺陷捕捉能力;引入注意力尺度序列融合模块 ASF-YOLO 改进 Neck 部分,引入多尺度特征融合机制,提升模型对线束端子缺陷特征的检测能力;引入 SlideLoss 分类损失函数,根据样本的难易程度调整损失权重,提升模型对困难样本的检测能力。实验结果表明,ACS-YOLO 模型相较于原始模型的准确率、召回率、mAP_{50} 分别提升 6.1%、1.0%、3.1%,可有效应用于线束端子缺陷检测任务。

关键词:线束端子缺陷检测;YOLOv11;C2PSA_EFA;边缘信息;ASF-YOLO;SlideLoss
中图分类号:TP391.4;TN911.73 **文献标识码**:A **国家标准学科分类代码**:510.4050

Wire harness terminal defect detection algorithm in complex backgrounds based on improved ACS-YOLO

Hu Haoyu Wang Shuqing Wang Xujia Xia Yangwei
(School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, China)

Abstract: Aiming at the difficulty of occlusion and blur detection caused by complex detection scene in the process of wire harness terminal crimping, a wire harness terminal defect detection model ACS-YOLO based on improved YOLOv11 is proposed. The C2PSA_EFA module is designed in this model. Combining C2PSA with edge-enhanced feature attention, sobel operator is used to extract and fuse edge information to enhance the irregular defect capture ability of harness terminals. The attention scale sequence fusion module ASF-YOLO is introduced to improve the Neck part, and the multi-scale feature fusion mechanism is introduced to improve the detection ability of the model to the wire harness terminal defect features. The SlideLoss classification loss function is introduced to adjust the loss weight according to the difficulty of the sample, and the detection ability of the model to difficult samples is improved. The experimental results show that the accuracy, recall rate and mAP_{50} of the ACS-YOLO model are increased by 6.1%, 1.0% and 3.1% respectively compared with the original model, which can be effectively applied to the harness terminal defect detection task.

Keywords: harness terminal defect detection; YOLOv11; C2PSA_EFA; edge information; ASF-YOLO; SlideLoss

0 引 言

在现代工业系统中,线束作为电力传输与信号控制的核心载体,其重要性堪比“神经血管网络”。随着电气化、智能化和轻量化技术的快速发展,线束系统已从简单的导线捆扎演变为集成化、高性能的复杂组件[1]。在汽车领域,随着新能源汽车和智能驾驶技术的快速发展,作为全球最大的新能源汽车生产国,我国线束行业呈现企业数量多、产业集中度低、技术水平差异大的典型特征。线束不仅承载着传统燃油车的电气连接,更是电动汽车高压能量传输的关键载体,其安全性和可靠性直接影响整车性能[2-3]。在航空航天领域,由于线缆规格多样、连接器种类繁杂,加之设计变更频繁,航空航天线束正朝着模块化设计、智能监测和新型复合材料应用方向发展[4]。线束作为电气连接的核心元件,压接端子的质量直接决定了线束系统的长期稳定性。压接工艺通过金属端子的塑性形变与导线导体形成机械互锁,这种微观尺度的物理结合需要实现金属界面的间隙贴合,从而确保电能传输的低损耗与信号传递的高保真[5]。

收稿日期:2025-07-25
*基金项目:国家自然科学基金(62306107)项目资助

在压接过程中,端子与导体的位置偏移、压力波动或模具磨损会导致金属接触面未完全咬合,导致端子出现异常形变造成电路虚接问题,最终可能引发短路、信号中断或系统功能失效。因此对线束端子进行缺陷检测是提升线束产品可靠性的必然选择[6]。

在当下工业实践中,端子的缺陷检测仍然高度依赖人工目视检查,操作员需借助放大镜观察端子形变,但受限于视觉分辨率与人体注意力衰减,人在疲劳作业时的漏检和误检率会随时间上升,这为线束的应用场景埋下了严重的安全隐患,因此需要一种更高效的自动检测方式[7]。

随着计算机视觉技术在目标检测领域的发展,为线束端子缺陷检测提供了新思路,常见的目标检测算法有单阶段算法、双阶段算法和基于 Transformer 架构的检测算法。DETR[8]算法结合 CNN 与 Transformer 结构进行目标检测,但存在收敛速度缓慢,训练时间长的问题。双阶段算法代表如 Fast R-CNN[9]和 Faster R-CNN[10]首先需要生成候选区域,然后对候选区域进行分类和定位,因此需要消耗更多的计算资源,不适合部署在要求高实时性的线束缺陷检测场中。单阶段算法如 SSD[11]与 YOLO[12-14]直接进行类别判断和边界框预测,在保证准确率的情况下具有更快的检测速度,更适合用于线束端子缺陷检测场景中。

针对线束端子缺陷检测问题,袁海兵等[15]提出一种基于改进 YOLOv7 的线束端子缺陷检测模型,通过在主干添加归一化注意力模块(normalization-based attention module,NAM);颈部网络引入通道级特征金字塔网络(channel-wise feature pyramid network, CFPNet);采用 SioU 损失函数替换为 CioU 损失函数,提升了模型的检测效果。Zhou 等[16]提出一种基于 YOLOv8 改进的轻量级网络 ILN-YOLOv8,通过增加 P2 检测层改进 FPN 和 PAN 结构;主干网络引入高效局部注意力机制(efficient local attention,ELA);使用 MPDioU 损失函数代替 CioU 损失函数;引入 slim-neck 架构改进颈部网络,在保证参数量下降的同时提升了检测精度。上述研究在端子缺陷检测领域取得了显著进展,但其验证场景多基于静态理想环境。在实际工业生产线中,线束端子常因传送带或机械臂抓取运动产生动态模糊,或因多线束重叠时产生的局部遮挡损失部分端子细节,进而干扰模型识别细微缺陷的能力。林哲等[17]结合 YOLO 与语义分割算法实现了遮挡情况下的物体识别与抓取,为复杂背景下线束端子缺陷检测提供了思路。

针对上述问题,本文提出了一种基于 YOLOv11 模型的目标检测算法并进行优化设计,YOLOv11 对比其他模型具有更高的准确率和更快的推理速度,因此选择在此模型基础上进行改进,本文主要工作为:

1)构建了一个面向工业复杂场景的线束端子缺陷数据集,该数据集采用 VMS 工业相机采集,设计运动模糊和线束堆叠复杂场景,以提高模型的泛化能力。

2)提出 C2PSA_EFA 模块,融合边缘增强特征注意力机制(edge-enhanced feature attention,EFA),融合了 L2 池化和 Sobel 边缘检测技术;引入 ASF-YOLO[18]模块改进颈部网络,融合多尺度特征;引入 Slide Loss[19]损失函数,优化模型对困难样本的学习效果。

3)设计模型对比实验、消融实验和改进模块的对比实验,验证本文改进模型的有效性。

1 数据集采集与制作

对于线束端子缺陷样本的采集,使用万濠自动影像测量仪 VMS-2010F 进行图像拍摄,使用设备如图 1 所示。该设备拥有彩色 CCD 摄像机、放置金属台、环形表面光以及底光。镜头倍率为 0.7～4.5 倍,基本满足拍摄需求。

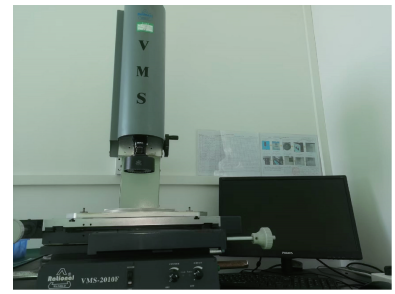


图 1 图像采集设备

Fig. 1 Image acquisition device

根据工厂实际的线束生产标准,本文将线束端子类型分为 5 类,分别为正常压接(pass)、未压接端子(uncrimping)、端子浅打(shallow)、端子深打(deep)、端子后仰(tilt)。所有端子类型如图 2 所示。每种端子类型包含 300 张 1 024×1 024 分辨率的原始图像,覆盖不同角度和光照条件,并使用 python 脚本对随机图片添加了高斯噪声。

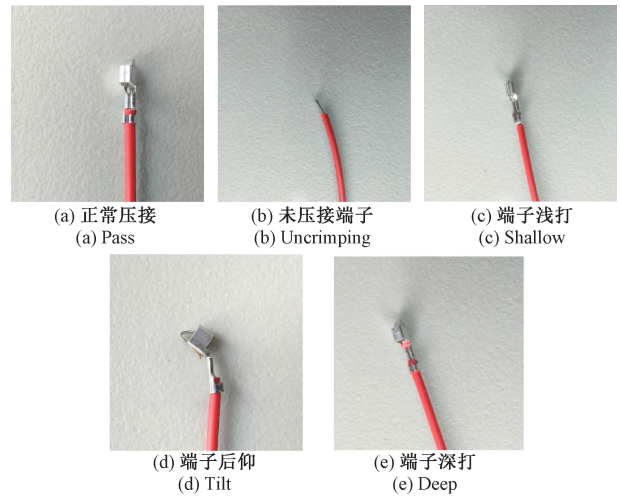
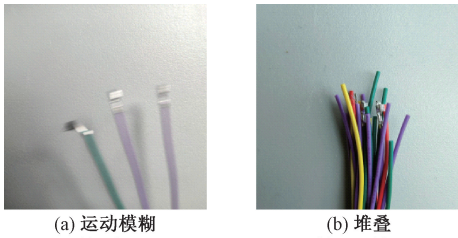


图 2 5 种不同类型端子图像

Fig. 2 Images of five different types of terminals

为了模拟实际工况中传送带或机械臂抓取线束运动所产生的运动模糊，通过摆动线束的方式，使用工业相机捕捉运动中的模糊图像，剔除因甩动幅度过大导致的关键特征完全丢失的样本，得到 500 张 1 024×1 024 分辨率的运动模糊图像。对于线束堆叠数据集的构建，从所有端子中随机选取带 3～5 个带端子的线束头部与线束尾部堆混合摆放，构建端子头部特征被遮挡的情况，最终得到 500 张 1 024×1 024 分辨率的线束堆叠图像。

运动模糊与线束堆叠样本如图 3 所示，使用 LabelImg 对所有图片进行标注，并对标注后的 1 000 张图片进行随机添加旋转变换、高斯噪声、光照变换等图像增强方法，最终将此部分数据集扩充到 2 000 张。



(a) 运动模糊　(b) 堆叠
(a) Blur　(b) Stacking

图 3　运动模糊与堆叠线束图像

Fig. 3　Images with blur and stacking wire harnesses

综上本数据集共包含 3 500 张图像以及对应标签，按照 8∶1∶1的比例将数据集划分为训练集、验证集和测试集，训练集包含 2 800 张图像和对应标签，验证集包含 350 张图像和标签，测试集包含 350 张图像和标签。

2　相关模型及其改进

2.1　YOLOv11 模型

YOLOv11 延续了 YOLO 系列“骨干网络(Backbone)-特征金字塔(Neck)-检测头(Head)”的架构[20]。如图 4 所示，骨干网络和特征金字塔部分采用 C3k2 模块替代 C2f 模块，C3k2 内部 Bottleneck 模块替换为了 C3k 模块，使模型具有更高的灵活性。同时在快速空间金字塔池化层(spatial pyramid pooling fast，SPPF)之后引入了 C2PSA 层，更有效地捕捉到细节信息。在检测头部分 YOLOv11 模型采用了深度可分离卷积(depthwise separable convolution，DWConv)，通过分离空间与通道处理，降低模型的参数量和计算量。

尽管 YOLOv11 在速度和精度上取得了较好的效果，但在实际的工业缺陷检测任务中仍存在以下不足之处。Backbone 中的 C3k2 模块感受野有限，难以有效捕捉被遮挡、模糊或堆叠区域中的缺陷特征，尤其是在端子排列密集

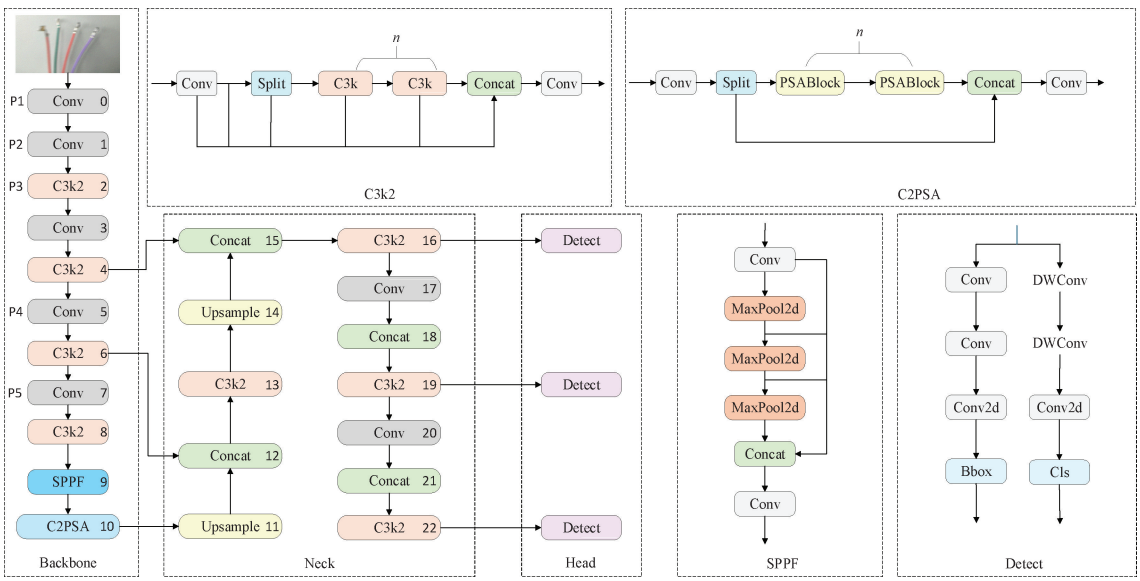


图 4　YOLOv11 整体结构

Fig. 4　Overall architecture of YOLOv11

或有干扰物时性能下降；模型缺乏对通道与空间维度的联合建模机制，容易出现特征权重分配不均的问题，难以聚焦关键区域；YOLOv11 默认采用的二元交叉熵损失在困难样本学习中存在局限，容易受到类别不平衡的干扰，导致模型缺乏对困难端子检测样本的识别能力。因此基于上述需要改进的部分，本文将采用 YOLOv11 作为基准模型进行相关模块的改进，具体改进部分将在下文中进行详细的介绍。

2.2　改进 C2PSA 模块

考虑到线束端子检测过程中受到的遮挡、运动模糊和噪声等复杂影响，并且 C3K2 模块感受野受限，导致 YOLOv11 有时难以准确识别缺陷特征，因此需要引入注意力机制来提升模型对目标的感知能力。由于骨干网络中 C2PSA 本身嵌入金字塔切片注意力(pyramid squeeze

attention,PSA),其核心是将部分卷积特征送入由多头自注意力模块(multi-head self attention, MHSA)和前馈网络(feed forward network, FFN)构成的子模块,另一部分则作为残差直接传递,最终通过卷积进行融合来增强全局建模能力。然而 MHSA 通过计算查询(Query)和键(Key)之间的点积来建立全局依赖关系,但这一过程是位置无关的,需要引入额外的位置编码,并且每个注意力头需独立学习权重,导致整体参数量显著增加,增加了训练和部署负担。因此,本文将 C2PSA 模块中的 MHSA 注意力机制模块替换为边缘增强特征注意力机制(edge-enhanced feature attention,EFA)模块,构成 C2PSA_EFA 模块,整体结构如图 5 所示。

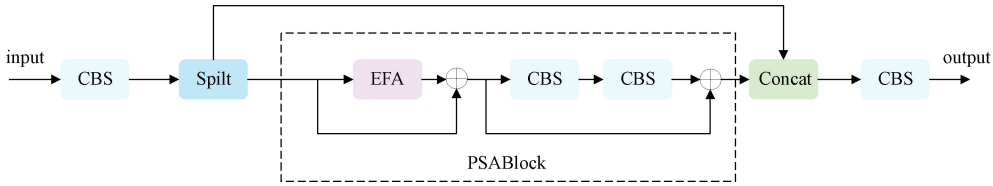


图 5 C2PSA_EFA 模块结构

Fig. 5 Structure of the C2PSA_EFA module

具体而言,EFA 模块基于卷积块注意力模块[21](convolutional block attention module,CBAM)上改进,通过串联通道注意力和空间注意力使模型能够更好的捕捉到图像中的重要信息,并且不需要额外的参数量和计算量。为了使 EFA 模块能够更好的捕捉在运动模糊和线束堆叠情况下对关键缺陷信息的识别能力。通过通道自适应增强、边缘引导、特征强化的方式帮助网络更精确的聚焦于目标区域,EFA 网络结构如图 6 所示。

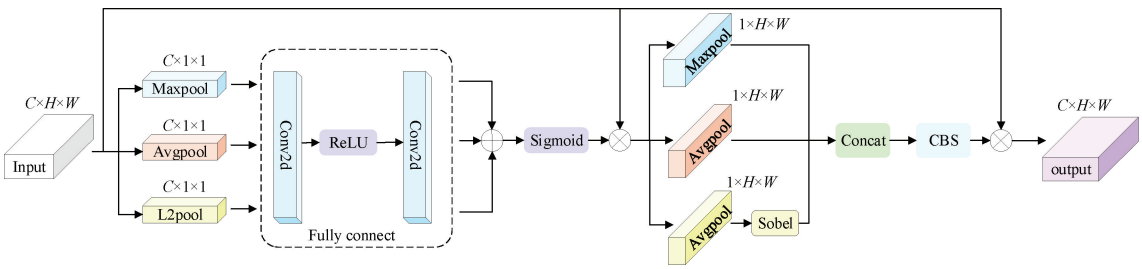


图 6 EFA 注意力机制结构

Fig. 6 Architecture of the EFA attention mechanism

EFA 模块首先将输入特征图分别经过平均池化层、最大池化层和 L2 池化层进行压缩,生成 3 个 $C\times1\times1$ 的池化特征图。其中 L2 池化通过计算欧几里得范数来进行特征聚合,通过开方求和抑制噪声的线性叠加。L2 池化能够有效抑制噪声的干扰,保留了更多的特征信息,公式如式(1)所示。

$$X_{L2}(i,j)=\sqrt{\frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}X(i,j)^2} \tag{1}$$

其中,H、W 是特征图的高和宽,$X(i,j)$ 为通道特征图在 (i,j) 的值。接着经过多层感知机进行卷积,然后将 3 个通道的注意力相加得到通道特征权重 W_c,具体公式为:

$$\begin{cases}X_{\text{avg}}=\dfrac{1}{H\times W}\displaystyle\sum_{i=1}^{H}\sum_{j=q}^{W}X(i,j)\\ X_{\max}=\max_{i=1,j=1}^{H,W}X(i,j)\\ W_c=\sigma(MLP(X_{\text{avg}})+MLP(X_{\max})+MLP(X_{L2}))\end{cases} \tag{2}$$

其中,σ 为 Sigmoid 函数,X 为输入特征图,MLP 为多层感知机计算。最终将原始特征图与特征权重相乘得到通道注意力模块输出 X',公式如式(3)所示。

$$X'=W_c\cdot X \tag{3}$$

接着在前两个分支对通道特征 X' 分别进行平均池化和最大池化,在第 3 个分支经过平均池化之后引入 Sobel 算子计算垂直方向上的梯度,使模型能够捕捉到端子图像的像素变化,从而关注端子的边缘特征。公式如式(4)~(5)所示。

$$X'_s=\boldsymbol{K}_s*AvgPool(X') \tag{4}$$

$$\boldsymbol{K}_s=\begin{bmatrix}-1&-2&-1\\0&0&0\\1&2&1\end{bmatrix} \tag{5}$$

图 7 展示了用 Sobel 算子在垂直方向上的边缘特征提取可视化效果,Sobel 算子在垂直方向上提取的边缘特征图展现出较为清晰的结构轮廓,该边缘特征随后与通道平均池化图融合,作为空间注意力权重生成的输入,在突出目标区域的同时抑制了背景干扰,有效增强了模型对复杂场景中细粒度目标的感知能力。

综上在特征融合阶段,输入特征图分别通过 3 种池化层得到 3 个尺寸为 $1\times H\times W$ 的特征图,然后将特征图进

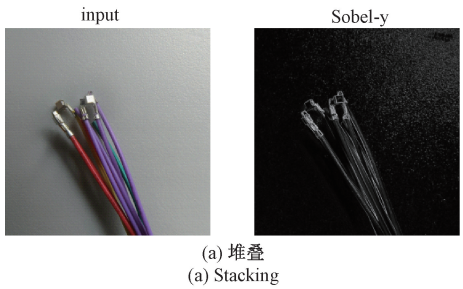


图 7　Sobel 算子边缘提取效果

Fig. 7　Edge extraction results using Sobel operator

行拼接，经过 7×7 的卷积层、批归一化层（batch normalization，BN）和 SiLU 激活函数之后得到空间权重 W_s，具体公式如式(6)所示。

$$\begin{cases} X'_{\text{avg}} = \dfrac{1}{C}\sum_{j=1}^{C} X'(i,j) \\ X'_{\max} = \max_{C=1}^{C} X'(i,j) \\ W_s = \sigma(f^{7\times7}(Concat(X'_{\text{avg}}, X'_{\max}, X'_s)) \end{cases} \tag{6}$$

其中，$f^{7\times7}$ 表示 7×7 的卷积计算。将空间权重与通道特征图相乘，最终得到尺寸为 $C\times H\times W$ 的空间注意力加权后的特征图 X''，公式如式(7)所示。

$$X'' = W_s \cdot X' \tag{7}$$

综上，C2PSA_EFA 模块在通道注意力模块中结合平均池化、最大池化与 L2 池化特征聚合方式，同时在空间注意力模块中引入了基于 Sobel 算子的边缘引导机制，最终得到包含丰富边缘信息和空间信息的特征图，能有效增强模型对端子的细微缺陷、遮挡干扰与模糊轮廓的识别能力。

2.3　颈部网络改进

针对不同类型端子以及不同复杂背景的多尺度问题，本文引入一种注意力尺度序列融合的模块，即 ASF-YOLO（attention scale sequence fusion YOLO）模块来改进颈部网络，其整体结构如图 8 所示。

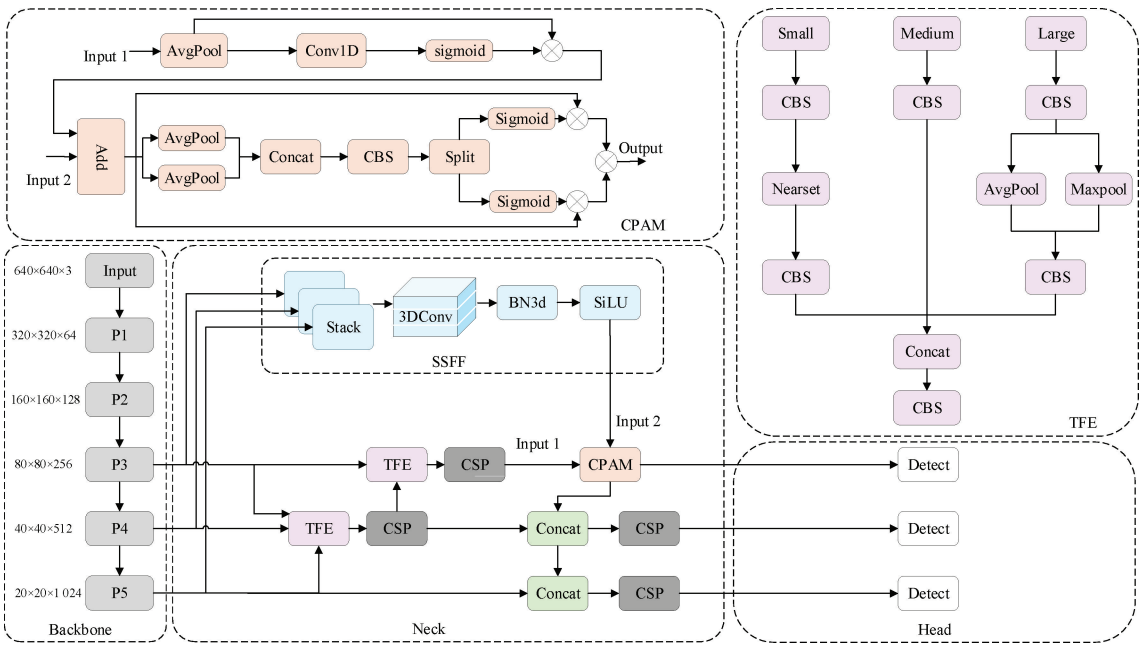


图 8　ASF-YOLO 整体架构

Fig. 8　Overall architecture of ASF-YOLO

该模型结合了注意力机制与多尺度特征融合策略，通过设计尺度序列特征融合（scale sequence feature fusion，SSFF）模块将主干网络的多尺度特征在路径聚合网络（path aggregation network，PANet）中进行三维卷积融合，同时构建三重特征编码器（triple feature encoder，TFE）模块，用来捕获缺陷样本的精细空间信息。在此基础上在 P3 分支引入通道-位置注意力机制（channel and position attention mechanism，CPAM），用于融合 SSFF 与 TFE 模块输出的多尺度特征，从而提高不同尺度缺陷的检测和分割精度。

SSFF 模块主要负责融合多尺度特征图和浅层特征图的详细信息，以主干网络输出的多尺度特征图 P3、P4 和

P5 作为输入,借助高斯滤波器对每层特征图进行平滑处理,具体公式如式(8)~(9)所示。

$$F_\sigma(i,j)=\sum_u\sum_v f(i-u,j-v)\times G_\sigma(u,v) \tag{8}$$

$$G_\sigma(u,v)=\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{u^2+v^2}{\sigma^2}} \tag{9}$$

其中,f 表示二维输入图像,经过二维高斯滤波器 G_σ 进行卷积平滑,得到平滑特征图 F_σ。SSFF 模块采用最近邻插值法将 P4 和 P5 上采样与 P3 相同的尺寸,通过 1×1 卷积统一通道数为 256。随后使用 unsqueeze 将每层特征图由三维张量[H,W,C]扩展为四维张量[1,H,W,C]。最后通过三维卷积、三维批归一化和 SiLU 激活函数实现不同尺度之间的特征融合。

对于三重特征编码模块(TFE)。相比于传统特征金字塔融合只对小尺寸的特征图进行上采样,TFE 通过将原特征图分成大、中、小 3 种尺寸,大尺寸经过卷积后引入最大池化(Max pooling)与平均池化(Average pooling)混合结构进行下采样。小尺寸经过卷积后使用最近邻插值法进行上采样。将 3 种不同尺寸特征图通过卷积调整为相同空间尺寸和通道数之后,在通道维度上拼接,具体公式如式(10)所示。

$$F_{TFE}=Concat(F_l,F_m,F_s) \tag{10}$$

其中,F_{TFE} 表示模块输出的特征图,F_{TFE} 通道数为 $3C$,分辨率与中尺寸特征图 F_m 相同。

在特征融合的基础上,ASF-YOLO 提出通道-位置注意力机制(CPAM)来增强关键特征的表达能力,CPAM 包含通道注意力网络和位置注意力网络。其中通道注意力网络以 TFE 模块输出的特征作为输入,位置注意力网络则接收通道注意力输出与 SSFF 全局特征的叠加结果。

通道注意力网络首先对输入特征进行全局平均池化,在此基础上通过大小为 k 的一维卷积核实现跨通道交互的捕获,具体公式如式(11)与(12)所示。

$$C=\psi(k)=2^{\gamma\times k-b} \tag{11}$$

$$k=\phi(C)=\left|\frac{\log_2(C)+b}{\gamma}\right|_{odd} \tag{12}$$

其中,$\gamma=2$、$b=1$,卷积核尺寸 k 与通道维数 C 相关,odd 表示最近邻的奇数。

位置注意力机制使用双轴编码机制,将输入特征分别沿宽度方向 z_w 和高度方向 z_h 进行平均池化,具体计算公式为:

$$z_w(i)=\frac{1}{H}\sum_{i=1}^{H}E_c(i,j) \tag{13}$$

$$z_h(j)=\frac{1}{W}\sum_{j=1}^{W}E_c(i,j) \tag{14}$$

其中,W 为输入特征图的宽度,H 为输入特征图的高度,$E_c(i,j)$ 是特征图位置值。将双轴特征拼接后经过 1×1 的卷积得到位置坐标 $P(a_w,a_h)$ 之后,通过 split 操作分割特征得到分割输出的宽度权重 s_w 和高度权重 s_h,具体公式如式(15)~(16)所示。

$$s_w=Split(a_w) \tag{15}$$

$$s_h=Split(a_h) \tag{16}$$

最终 s_w 和 s_h 分别通过 Sigmoid 激活函数,赋予权重之后得到 CPAM 模块的输出,如式(17)所示。

$$F_c=\boldsymbol{E}\times s_w\times s_h \tag{17}$$

其中,$\boldsymbol{E}$ 代表通道和位置的权重矩阵。

2.4 SlideLoss 损失函数

在实际端子缺陷检测当中,端子缺陷检测样本分为简单样本和困难样本。简单的缺陷样本通常清晰容易分辨和定位,而困难的缺陷样本会包含运动模糊、线束遮挡和噪声等因素。如果模型依赖对容易样本的训练,模型对困难的样本的识别能力不足,使得此类样本更容易被误判为背景类,最终会造成漏检和误检,降低模型整体检测精度。并且困难样本包含更丰富的特征变化和噪声干扰信息,这些信息能够提升模型的鲁棒性,在训练过程中主要学习容易样本会因学习过多显著特征而产生过拟合,这表明即使模型在训练集上表现很好,但难以适应真实工业场景中复杂多变的场景,泛化能力不足。因此,训练模型能够识别困难样本能更好的提升模型的泛化能力和检测精度。目前,YOLOv11 等主流网络采用的类别分类损失为 BCEWithLogitsLoss 损失函数,该函数将 Sigmoid 激活函数与二元交叉熵损失函数(binary cross entropy loss,BCELoss)结合,其公式如式(18)所示。

$$l_{bce}=-\lambda_n[y_n\cdot\lg(\sigma(x_n))+(1-y_n)\cdot\lg(1-\sigma(x_n))] \tag{18}$$

其中,x_n 为模型对第 n 个样本标签的预测值,y_n 为第 n 个样本标签的真实值,λ_n 为手动给定的样本权重,σ 为 Sigmoid 激活函数。尽管 BCEWithLogitsLoss 通过结合激活函数与二元交叉熵损失函数简化计算流程,但缺乏对困难样本的预测,从而影响模型整体性能与泛化能力,因此本文在 YOLOv11 模型中引入 SlideLoss 来解决这个问题,

SlideLoss 的核心是引入了自适应的动态阈值 μ,其值为交并比(intersection of union,IoU)的平均值。IoU 值反映了预测框与真实框的重叠程度,高 IoU 样本通常特征显著,样本量较多,而低 IoU 样本则通常特征模糊或背景干扰严重,样本量较少,因此 IoU 在[0,1]内呈现偏态分布,动态阈值 μ 通常位于密度分布区间。本文使用 YOLOv11 模型在验证集上测得 IoU 分布如图 9 所示。

如图 9 所示,对于 IoU 过低的样本,如当 $0.1<\mathrm{IoU}<0.6$ 时,此类样本分布基本在 $\mathrm{IoU}<\mu-0.1$ 的范围,视为困难样本,其包含过多的噪声和非典型特征,因此赋予较低的权重。对于分布在 $\mu-0.1<\mathrm{IoU}<\mu$ 的样本视为模糊样本,通常位于模型分类决策边缘附近,能显著提升模型对复杂特征判别能力,赋予其最高的权重。对于分布在 $\mathrm{IoU}>\mu$ 的样本视为容易样本,样本 IoU 越高则赋予的权重值越低。

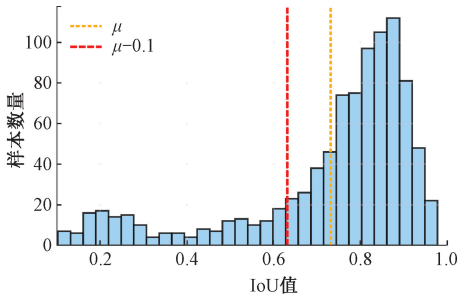


图 9　验证集上 IoU 分布

Fig. 9　IoU distribution on the validation set

SlideLoss 加权函数具体公式如式(19)所示。

$$f(x)=\begin{cases}1, & x\leqslant\mu-0.1\\ e^{1-\mu}, & \mu-0.1<x<\mu\\ e^{1-x}, & x\geqslant\mu\end{cases}\tag{19}$$

SlideLoss 加权函数图像如图 10 所示。最终，SlideLoss 在 BCEWithLogitsLoss 损失的基础上引入该加权函数，构造得到完整损失公式如式(20)所示。

$$I_n=l_{bce}\cdot f(x_n)\tag{20}$$

SlideLoss 通过自适应加权机制有效优化了模型对困难样本的特征提取能力。在线束端子缺陷检测任务中，该损失函数能够使模型聚焦于线束遮挡、运动模糊等缺陷样本，提升了模型对各类复杂缺陷的检测精度。

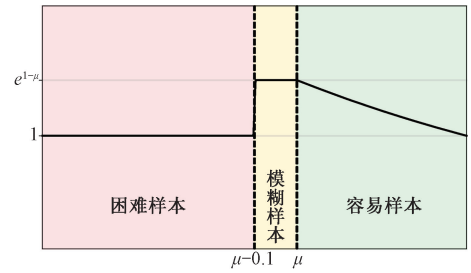


图 10　SlideLoss 加权函数

Fig. 10　Weighting function of SlideLoss

2.5　ACS-YOLO 模型

基于上述改进，本文得到了最终的改进模型 ACS-YOLO，模型结构如图 11 所示。

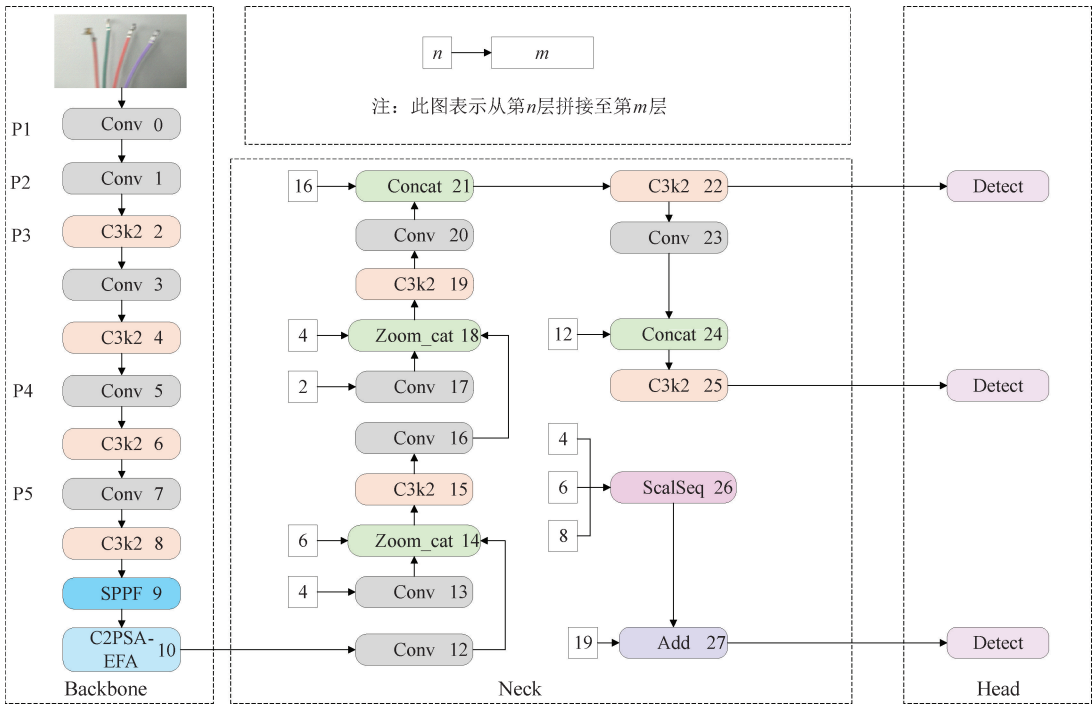


图 11　ACS-YOLO 模型结构

Fig. 11　Architecture of the ACS-YOLO model

首先在 YOLOv11 的 Backbone 部分，为提升特征表征能力，将 Backbone 末端 C2PSA 模块替换为 C2PSA_EFA 注意力模块。该模块引入了 L2 池化与 Sobel 边缘提取器，L2 池化使模型具有更强的噪声鲁棒性，Sobel 算子使模型能够更加关注目标对象的边缘特征，有效提升模型在遮挡严重、图像模糊或存在高噪声等复杂场景中对端子缺陷特征的提取能力。Neck 部分在原有结构的基础上引入了 ASF-YOLO 模块，该模块融合注意力机制与尺度序列特征融合策略，针对端子在实际检测中常出现的尺寸差异大、边缘模糊、背景干扰等问题，构建更有效的多尺度特征融合方式。同时改进了损失函数部分，引入自适应加权损失函数 SlideLoss，该损失函数通过动态阈值区分困难样本与容易样本，并对困难样本赋予更高权重，引导模型重点学习遮挡、运动模糊等难以识别的缺陷区域，从而提升模

型精度和鲁棒性。

总的来说,C2PSA_EFA 模块增强了 Backbone 深层特征的判别能力,提高了提取输入数据特征的能力。ASF-YOLO 模块则有效补足了 Neck 在多尺度信息融合方面的不足,使模型能够更全面地理解尺寸不一、形态各异、背景复杂的端子缺陷。最终搭配 SlideLoss 损失函数,更加关注困难样本,提升模型对复杂缺陷类型的学习能力与泛化性能,三者共同构成 ACS-YOLO 的核心改进策略,最终提升模型的检测效果。

3 实验结果与分析

3.1 实验环境

本文实验基于 Windows11 操作系统,硬件配置为 NVIDIA GeForce RTX 4070 Ti Super GPU,显存为 16 G,python 版本为 3.12.7,CUDA 版本为 12.6,模型训练参数设置如表 1 所示。

表 1 模型训练参数

Table 1 Model training parameters

参数	值
训练轮次	200
批次大小	16
图片尺寸	640×640
学习率	0.01
优化器	SGD

3.2 评价指标

为评估模型性能,本文应用多种指标如精度(precision)、召回率(recall)、单类别精度(average precision, AP)和平均精度(mean average precision, mAP)来评估改进 YOLOv11 在端子缺陷数据集上的效果。对于每个类别计算不同置信阈值下的精度和召回率,计算公式为:

$$\begin{cases} P = \dfrac{TP}{TP + FP} \\ R = \dfrac{TP}{TP + FN} \end{cases} \tag{21}$$

其中,TP 为正确检测的数量,FP 为误检的数量,FN 为漏检的数量。

AP 为 PR 曲线下的面积,表示某类标签的平均精度,mAP 为 AP 的算术平均即所有类别的平均精度,计算公式为:

$$AP = \int_0^1 P(R)\mathrm{d}R \tag{22}$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \tag{23}$$

3.3 消融实验分析

为了验证每个改进策略对模型性能的提升作用,本文以 YOLOv11n 为基准设计消融实验,实验过程保持模型参数和训练条件一致,在测试集上验证模型性能,消融实验结果如表 2 所示,其中"√"表示添加该改进模块。

表 2 消融实验结果

Table 2 Results of ablation experiments

序号	ASF	C2PSA_EFA	SlideLoss	Params/10^6	FLOPs/G	P/%	R/%	mAP_{50}/%
1	—	—	—	2.58	6.3	86.1	81.7	88.4
2	√	—	—	2.95	7.6	87.2	82.3	89.5
3	—	√	—	2.54	6.3	88.6	84.5	90.6
4	—	—	√	2.58	6.3	86.0	82.1	89.2
5	√	√	—	2.91	7.6	90.2	82.6	90.8
6	√	√	√	2.91	7.6	92.2	82.7	91.5

如表 2 中结果所示,第 1 组表示未改进的 YOLOv11n 模型,参数量为 2.58 M,计算量为 6.3 G,mAP 为 88.4%。第 2 组为单独引入 ASF-YOLO 模块后,参数量提升 0.37 M,计算量提升 1.3 G,这主要是由于 ASF-YOLO 引入了特征融合结构,增加了特征图之间的计算过程,尽管复杂度有所上升,模型的 mAP 提升了 1.1%,证明了其在多尺度特征融合方面的有效性;第 3 组为在 Backbone 末端改进的 C2PSA_EFA 模块,参数量与计算量稍微减少,mAP 提升了 2.2%,证明了该模块能在保持轻量的同时有效提升了模型对模糊、遮挡等复杂背景中缺陷区域的感知能力;第 4 组为添加 SlideLoss 损失函数改进策略,在不增加额外参数的前提下,mAP 提升了 0.8%,证明该损失函数对于困难样本给予了更高关注,使得模型对复杂背景下的缺陷检测效果更好。

当 ASF-YOLO、C2PSA_EFA、SlideLoss 3 个模块同时被引入时,C2PSA_EFA 模块首先引入多类型池化与边缘引导机制,强化模型对关键区域的空间感知与关注能力,ASF 模块进一步增强了模型对复杂缺陷的特征表达能力,Slide Loss 则作为训练阶段的损失函数优化策略,利用动态阈值机制聚焦于困难样本,使模型更有效地关注复杂缺陷。因此 mAP 相比于基准模型提升了 3.1%,参数量和计算量虽略有增加,但整体控制在合理范围之内,证明了本文算法在线束端子检测上的有效性。

3.4 模型对比试验分析

为了进一步验证本文提出的模型的优越检测性能,本文选用 RT-DETR、YOLOv5n、YOLOv8n、YOLOv10n、

YOLOv11n、YOLOv12n 这 6 种主流原始模型以及 MSB-YOLO 和 CGA-YOLO 两种最新基于 YOLOv11 的改进模型进行了对比试验，将所有对比模型设置为统一规格，其中带有后缀“n”的模型表示轻量级模型，网络尺寸为原始模型的 1/3，在测试集上统一验证模型效果，对比实验结果如表 3 所示。

表 3　不同模型对比实验结果

Table 3　Comparative experimental results of different models

模型	Params/10^6	FLOPs/G	P/%	R/%	mAP_{50}/%	FPS/fps
RT-DETR-l	31.99	103.5	92.9	82.5	89.6	32.7
YOLOv5n	2.50	7.1	88.2	77.9	88.0	57.8
YOLOv8n	3.01	8.1	87.5	81.2	88.9	60.1
YOLOv10n	2.26	6.5	83.2	76.9	85.1	54.9
YOLOv11n	2.58	6.3	86.1	81.7	88.4	56.1
YOLOv12n	2.56	6.3	84.8	73.8	83.6	39.7
MSB-YOLO[22]	2.37	7.3	86.1	75.0	85.5	36.9
CGA-YOLO[23]	2.55	6.3	84.5	82.5	88.3	51.3
ACS-YOLO	2.91	7.6	92.2	82.7	91.5	54.7

由表 3 结果可知，ACS-YOLO 在平均精度均值 mAP_{50} 指标上比基准模型 YOLOv11n 提升了 3.1%，且在其他基准模型中取得了最好的效果，参数量相较于 YOLOv11n 上涨 0.33 M 但低于 YOLOv8n 和 RT-DETR 模型。YOLOv10n、YOLOv12n、MSB-YOLO 和 CGA-YOLO 参数量虽然较小但精度有所下降，并且召回率比较低，表明这些模型在线束堆叠场景下未能识别出更多的端子，不适用于实际检测当中。从准确率方面来看 ACS-YOLO 仅用较少的参数量获得了 92.2%的准确率，与 RT-DETR 持平但远高于其他模型，表明其在运动模糊场景下的端子识别效果更好。从实时性能来看，ACS-YOLO 达到了 54.7 fps，检测速度略低于 YOLOv8n 但基本高于其他基准模型和改进模型，基本满足工业检测实时性要求。

图 12 为表 3 中部分模型在训练过程中迭代相同次数的训练结果对比，可以看出，在迭代相同次数时 ACS-YOLO 在 mAP_{50} 指标上领先于其他基准模型，收敛速度也快于其他模型，迭代 80 轮的效果已经超过其他模型。综上所述，ACS-YOLO 以较小的计算代价实现了优于其他原始模型的检测精度，对具有运动模糊和堆叠端子的目标检测更具有优势。

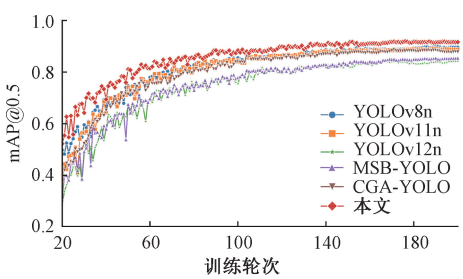


图 12　迭代轮数相同时的对比试验

Fig. 12　Comparative experiments at the same training epochs

3.5　不同改进 C2PSA 对比实验

为验证本文改进的 C2PSA_EFA 模块在运动模糊及堆叠遮挡场景下对端子图像特征提取的优越性，本文设计了多种改进 C2PSA 的对比实验。选取的注意力模块包括高效通道注意力[24]（efficient channel attention，ECA）、全局注意力机制[25]（global attention mechanism，GAM）、坐标注意力[26]（coordinate attention，CA）以及 CBAM 4 种典型模型，分别嵌入 C2PSA 中同一位置，在测试集上进行横向对比，对比结果如表 4 所示。

表 4　不同注意力机制对比结果

Table 4　Comparison results of different attention mechanisms　%

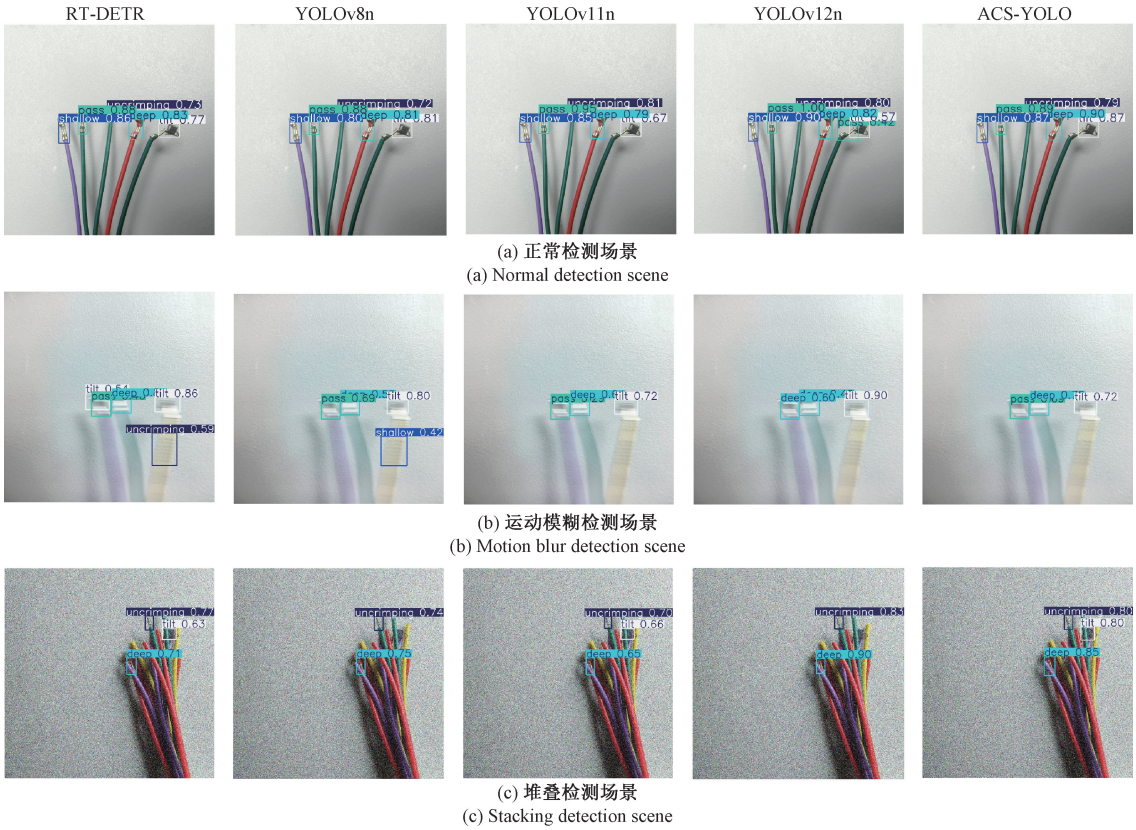
模型	P	R	mAP_{50}
C2PSA_GAM	84.7	79.0	88.4
C2PSA_ECA	84.9	80.9	88.1
C2PSA_CBAM	84.9	81.9	89.3
C2PSA_CA	88.1	80.5	89.4
C2PSA_EFA	88.6	84.5	90.6

相比于直接在骨干网络添加注意力机制，通过改进 C2PSA 得到的 C2PSA_EFA 注意力机制在 mAP_{50} 表现最优，基本提升超 1.0%，证明了 C2PSA_EFA 能提高特征提取能力的有效性。

3.6　模型对比结果可视化

为了直观的比较 ACS-YOLO 算法与其他算法在复杂场景下的线束端子缺陷检测效果，本文选用不同场景下的图片对比实验，所有场景包括具有正常情况下的线束、运动模糊下的线束和多根堆叠的线束，实验结果如图 13 所示。图 13(a)为正常检测场景下的不同模型检测效果对比，结果表明除了 YOLOv12n 模型出现了错检情况，其余

对 5 种类型端子检测效果均为良好且 ACS-YOLO 置信度总体均高于其他模型。图 13(b)为运动模糊场景下不同模型检测效果对比,结果表明 RT-DETR、YOLOv8n、YOLOv12n 出现了错检的情况,而 ACS-YOLO 没有错检,且置信度高于 YOLOv11n 模型。图 13(c)为堆叠场景下不同模型检测效果对比,结果表明 YOLOv8 和 YOLOv12 出现了漏检的情况,而 ACS-YOLO 在线束堆里识别出了所有类型的端子,且置信度显著高于其他模型。综合来看 ACS-YOLO 的检测精度更高,并且能更好的识别漏检测端子。



(a) 正常检测场景
(a) Normal detection scene

(b) 运动模糊检测场景
(b) Motion blur detection scene

(c) 堆叠检测场景
(c) Stacking detection scene

图 13 复杂场景下检测效果对比

Fig. 13 Detection performance comparison in complex scenarios

4 结 论

针对实际工况中线束产生运动模糊或者多根线束堆叠导致缺陷不明显的情况,本研究提出一种基于 YOLOv11 的改进线束端子缺陷检测模型。该模型使用 C2PSA_EFA 改进骨干网络;引入 ASF-YOLO 模块改进颈部网络;采用 SlideLoss 损失函数代替 YOLOv11 使用的 BCEWithLogitsLoss 分类损失函数。实验结果表明,改进后的 ACS-YOLO 模型相比于 YOLOv11 原模型的 mAP_{50} 提升了 3.1%,具有更高的准确性,参数量略微上涨但仍适用于模型部署,实时性能上看基本满足工业需求。然而本文的改进稍微增加了模型的参数量与计算量,因此后续可继续研究模型剪枝轻量化等方式减小模型的参数量和计算量,同时保持精度不变,考虑将模型部署至实际设备当中,还可考虑更全面的实际工况背景进行研究。

参考文献

[1] KAMBLE SUPRIYA S, KULKARNI ASHWINI A. Automatic optical inspection system for wiring harness using computer vision[C]. 2021 IEEE International Conference on Electronics, Computing and Communication Technologies(CONECCT), 2021: 1-5.

[2] NGUYEN H G, KUHN M, FRANKE J. Manufacturing automation for automotive wiring harnesses[J]. Procedia CIRP, 2021, 97: 379-384.

[3] TROMMNAU J, KÜHNLE J, SIEGERT J, et al. Overview of the state of the art in the production process of automotive wire harnesses, current research and future trends[J]. Procedia CIRP, 2019, 81: 387-392.

[4] 傅桂林, 刘晓蓉, 赵裔俊, 等. 航空线束制造现状及自动化转型方向研究[J]. 航空精密制造技术, 2023, 59(6): 61-62, 65.
FU G L, LIU X R, ZHAO Y J, et al. Research on current status and automation transformation direction of aviation wire harness manufacturing[J]. Aviation Precision Manufacturing Technology, 2023, 59(6): 61-62, 65.

[5] CAI Y M, ZHOU K, LI CH SH, et al. Application of ultrasonic flaw detection technology in crimping

quality inspection of line hardware[C]. 2019 IEEE Sustainable Power and Energy Conference(iSPEC), 2019: 647-651.

[6] 袁彬淦，钟铭恩，倪晶鑫. 线束端子压接后外观缺陷的视觉检测算法研究[J]. 系统仿真学报，2022，34(5)：1152-1159.
YUAN B G, ZHONG M EN, NI J X. Research on visual inspection algorithm of crimping appearance defects for wiring harness terminals[J]. Journal of System Simulation, 2022, 34(5): 1152-1159.

[7] 赵朗月，吴一全. 基于机器视觉的表面缺陷检测方法研究进展[J]. 仪器仪表学报，2022，43(1)：198-219.
ZHAO L Y, WU Y Q. Research progress of surface defect detection methods based on machine vision[J]. Chinese Journal of Scientific Instrument, 2022, 43(1): 198-219.

[8] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]. European Conference on Computer Vision, 2020: 213-229.

[9] GIRSHICK R. Fast R-CNN [C]. 2015 IEEE International Conference on Computer Vision(ICCV), 2015: 1440-1448.

[10] 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.

[11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multiBox detector [C]. European Conference on Computer Vision, 2016: 21-37.

[12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016: 779-788.

[13] 张志宏，张力玲，马婷婷，等. 轨道缺陷无损检测技术研究现状综述[J]. 电子测量技术，2025，48(18)：53-72.
ZHANG ZH H, ZHANG L L, MA T T, et al. A review of the research status of nondestructive testing technology for track defects [J]. Electronic Measurement Technology, 2025, 48(18):53-72.

[14] 山显英，张琳，李泽慧. 深度学习驱动下的目标检测研究进展综述[J]. 计算机工程与应用，2025，61(1)：24-41.
SHAN X Y, ZHANG L, LI Z H. Review of research progress in object detection driven by deep learning[J]. Computer Engineering and Applications, 2025, 61(1): 24-41.

[15] 袁海兵，赵凤胜，杨奕洋，等. 基于改进 YOLOv7 的线束缺陷检测研究[J]. 国外电子测量技术，2024，43(2)：165-173.
YUAN H B, ZHAO F SH, YANG Y Y, et al. Research on wire harness defect detection based on improved YOLOv7 [J]. Foreign Electronic Measurement Technology, 2024, 43(2): 165-173.

[16] ZHOU X J, KAN J CH, ROSELY N F L M, et al. ILN-YOLOv8: A lightweight image recognition model for crimped wire connectors[J]. IEEE Access, 2025, 13: 5193-5202.

[17] 林哲，潘慧琳，陈丹. 融合改进 YOLO 和语义分割的遮挡目标抓取方法[J]. 电子测量与仪器学报，2024，38(12)：190-201.
LIN ZH, PAN H L, CHEN D. Grasp method for occlusion method by fusing improved YOLO with semantic segmentation [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(12): 190-201.

[18] KANG M, TING CH M, TING F F, et al. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation [J]. Image and Vision Computing, 2024, 147: 105057.

[19] YU Z P, HUANG H B, CHEN W J, et al. YOLO-FaceV2: A scale and occlusion aware face detector[J]. Pattern Recognition, 2024, 155: 110714.

[20] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements[J]. ArXiv preprint arXiv: 2410.17725, 2024.

[21] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [J]. Lecture Notes in Computer Science, 2018, 11211: 3-19.

[22] CHEN Y M, YUAN X B, WANG J B, et al. YOLO-MS: Rethinking multi-scale representation learning for real-time object detection [J]. IEEE Trans Pattern Anal Mach Intelligence, 2025, 47(6): 4240-4252.

[23] YE W L, REN J J, LIANG J, et al. Automatic detection and localization of surface defects in high-speed railway ballastless track based on cascaded group attention and optoelectronic encoder[J]. Structural Health Monitoring, 2025, DOI: 10.1177/14759217251338866.

[24] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2020: 11531-11539.

[25] LIU Y CH, SHAO Z R, HOFFMANN N. Global attention mechanism: Retain information to enhance channel-spatial interactions[J]. ArXiv preprint arXiv: 2112.05561, 2021.

[26] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13708-13717.

作者简介

胡浩宇，硕士研究生，主要研究方向为深度学习、图像处理。
E-mail：467617629@qq.com

王淑青(通信作者)，教授，博士，主要研究方向为工业控制、机器视觉、图像处理。
E-mail：1836595319@qq.com

王许佳，硕士研究生，主要研究方向为深度学习、图像处理。
E-mail：1085529103@qq.com

夏杨威，硕士研究生，主要研究方向为深度学习、图像处理。
E-mail：15527137207@163.com