

# 基于双维自注意力的跨阶段融合燃料棒 外观缺陷检测方法

张小刚, 俞东宝, 汤慧, 朱永利, 曹微

(中核北方核燃料元件有限公司, 内蒙古 包头 014035)

**摘要:** 压水堆核燃料棒是核电站普遍采用的一种燃料元件, 燃料棒的生产质量关乎核电站的运行安全, 对燃料棒的外观进行检测极其重要。针对生产过程中燃料棒外观缺陷背景复杂、特征提取难的特点, 本文提出了一种基于双维自注意力的跨阶段融合模型(DSCFM), 引入了一个新颖的多尺度特征融合结构(MFFS)作为模型的颈部结构, 用于高效地处理和融合不同层次的特征信息。此外, 为了进一步挖掘和利用底层的结构信息, 并对深层信息进行有效整合, 提出了一个具有双维度特性的自注意力特征融合模块(DSAF)。最后, 结合多尺度反卷积结构, 实现了对特征的有效上采样和优化, 显著提高了模型对不同尺度信息的捕获能力。在燃料棒外观缺陷数据集上进行验证, 实验表明 DSCFM 相较于其他检测模型, 能够快速精准地识别缺陷, 其 mAP 值达到了 82.0%。

**关键词:** 燃料棒; 缺陷检测; 多尺度; 双维度; 自注意力

中图分类号: TL99; TP181 文献标志码: A 文章编号: 1000-6931(2025)00-0001-09

doi: 10.7538/yzk.2024.youxian.0926

## Cross Stage Fusion Fuel Rod Appearance Defect Detection Method Based on Dual Dimensional Self Attention

ZHANG Xiaogang, YU Dongbao, TANG Hui, ZHU Yongli, CAO Wei

(China North Nuclear Fuel Co., Ltd., Baotou 014035, China)

**Abstract:** Pressurized water reactor nuclear fuel rods are a type of fuel element commonly used in nuclear power plants. The production quality of fuel rods is related to the safe operation of nuclear power plants. It is very important to inspect the appearance quality of fuel rods. In view of the complex background and difficult feature extraction of fuel rod appearance defects in the produce scene, this paper proposed a novel cross stage fusion model based on dual dimensional self attention (DSCFM). The model used Extended-ELAN as the backbone network to extract features and designed a novel multi-scale feature fusion structure (MFFS) as the neck structure of the model to efficiently process and fuse feature information at different levels. The purpose of the MFFS design is to optimize the information flow between features at different levels. A large amount of detailed information was retained through a detailed fusion mechanism, while strengthening the model's ability to understand the complexity of the scene. In addition, to further mine and utilize the underlying structural information and effectively integrate deep information, a self attention feature fusion module with dual dimensional characteristics (DSAF) was proposed. DSAF expanded the deep and underlying features into dual

dimensional feature maps, and used its own transposed matrix to generate channel and spatial attention maps, thereby accurately enhancing the expression of key information, while suppressing irrelevant noise and optimizing the feature fusion process. Through this dual dimensional self attention mechanism, DSAF dynamically adjusts feature responses, effectively captures long-term dependencies, and enhances the model's adaptability and interpretation capabilities for complex scenes. Finally, combined with a multi-scale deconvolution structure, the DSCFM achieves effective upsampling and optimization of features, significantly improving the model's ability to capture information at different scales and its robustness in various visual tasks. The results are verified on a fuel rod appearance defect dataset, and the experiments show that compared with other detection models, the DSCFM can quickly and accurately identify defects, with an mAP of 82.0% and a recall rate of 77.9%.

**Key words:** fuel rod; defect detection; multi-scale; dual dimensional; self attention

燃料棒<sup>[1]</sup>是核燃料组件<sup>[2]</sup>的核心部分,燃料棒在生产环节中极有可能与其他物体发生碰撞或者刮擦而产生缺陷。其中碰撞容易使燃料棒棒身表面出现磕伤,刮擦容易使燃料棒棒身表面出现划伤或者黏附异物。加之在役期间核燃料组件长期处于高温、高压及强中子辐射场等复杂环境条件下,燃料棒中的芯块会出现肿胀和变形,一旦燃料棒表面具有缺陷,燃料棒包壳管极有可能出现破裂,造成重大安全事故,所以燃料棒的质量对燃料组件以及堆内运行有着至关重要的影响。燃料棒外观缺陷类型多,表现形式复杂,受检测技术水平的制约,长期以来燃料棒外观质量的控制措施只能依靠检测人员在强光下目视检查,缺陷结果无法定量,且受检测人员经验影响,无法保证连续检测的灵敏度和稳定性<sup>[3]</sup>。其检测结果主要取决于检验人员本身的认知、经验、责任心以及视觉疲劳程度等因素,存在人因失误的风险。

随着科学技术的发展,在高端工业制造领域利用机器视觉技术<sup>[4]</sup>进行缺陷检测<sup>[5]</sup>已经开始流行起来。机器视觉检测技术以其检测结果准确直观,可靠性和重复性好,能保证连续检测的灵敏度和稳定性的特点,可以有效地取代现有燃料棒人工检测的方式,避免人因失误,方便生产过程的掌控和优化。目前,基于深度学习<sup>[6]</sup>的工业缺陷目标检测技术以其高精度、高效率的特点成为研究的热点。一方面,当前的两阶段检测网络通常分为候选区域提取和候选区域分类回归<sup>[7]</sup>。自从 Girshick 等<sup>[8]</sup>提出区域卷积网络(regional CNN, R-CNN)并成功应用于目标检测领域以来,许多改进算法快速发展而出,旨在提升检测的速度和准确

性。Ren 等<sup>[9]</sup>提出 Faster R-CNN,其核心创新在于引入了区域建议网络(RPN),实现了高效的候选区域生成,从而显著提高了对象检测的速度和性能。Zhang 等<sup>[10]</sup>提出了一种基于帧间差分和时空上下文的改进 Faster R-CNN 算法,通过改进 RPN 模块来提高 Faster R-CNN 的训练和测试速度,实现目标的实时检测。陈法法等<sup>[11]</sup>以感兴趣区域校准代替感兴趣区域池化,减少 Faster R-CNN 模型自身量化产生的缺陷定位误差。另一方面,一阶段检测网络<sup>[12]</sup>也取得了显著进展,如 YOLO 系列算法<sup>[13]</sup>的实现与应用,具备更高的检测速度与精准率。Adli 等<sup>[14]</sup>提出一种基于 YOLOv5 的新模型,通过 Swin Transformer v2 块将基于窗口的自注意力机制集成到 Backbone 中,提高全局和局部信息的利用率。YOLOv6<sup>[15]</sup>通过在骨干网络中引入 RepVGG 结构,实现了在训练阶段提高模型表达能力和训练效率,并在推理阶段通过结构重参数化加速推理过程,从而显著提升了整体检测性能。Li 等<sup>[16]</sup>通过引入空间金字塔池化快速跨阶段部分连接模块和全局注意力机制改进 YOLOv7,以提高特征提取精度。戴林华等<sup>[17]</sup>对 YOLOv8n<sup>[18]</sup>进行改进,减少特征图计算冗余,降低模型的参数量。然而,燃料棒外观缺陷表现形式复杂、背景复杂的特点对基于深度学习的图像检测技术提出了更高要求。首先燃料棒在生产过程中,其表面缺陷表现为划伤、磕伤及异物等多种形式,而划伤、磕伤及异物这 3 种缺陷出现频次最高。这些缺陷的尺寸、形状、分布位置各不相同,划伤通常表现为燃料棒表面出现的线性或曲线状痕迹,可能是单条或多条平行线,磕伤通常表现为燃料棒

表面的局部凹陷, 形状可能为圆形、椭圆形或不规则形, 异物缺陷通常表现为燃料棒表面附着或嵌入的颗粒、纤维或其他外来物质。其次, 燃料棒表面的反光特性、材质不均匀性增加了检测的难度。例如, 反光会导致图像过曝或阴影, 材质不均匀性会被误判为缺陷, 这些背景复杂性使得缺陷的准确提取和识别变得极具挑战。

本文提出一种双维自注意力<sup>[19]</sup>的跨阶段融合模型(DSCFM), 设计多尺度特征融合<sup>[20]</sup>结构(MFFS)和双维度自注意力特征融合模块(DSAF), 通过多尺度反卷积<sup>[21]</sup>结构优化特征上采样, 有效提升模型在处理多尺度信息<sup>[22]</sup>的能力, 具备较好

的燃料棒外观缺陷识别速度和精度。

### 1 模型架构

DSCFM 整体架构如图 1 所示, 采用 Extended-ELAN 作为主干网络以提取特征, 并引入了一个新颖的 MFFS 作为模型的颈部结构, 用于处理不同层次的特征信息。此外, 为了进一步挖掘和利用底层的结构信息, 并对深层信息进行有效整合, 引入一个 DSAF。最终, 结合多尺度反卷积结构, DSCFM 实现了对特征的有效上采样和优化, 显著提高了模型对不同尺度信息的捕获能力及其在各种视觉任务中的鲁棒性。

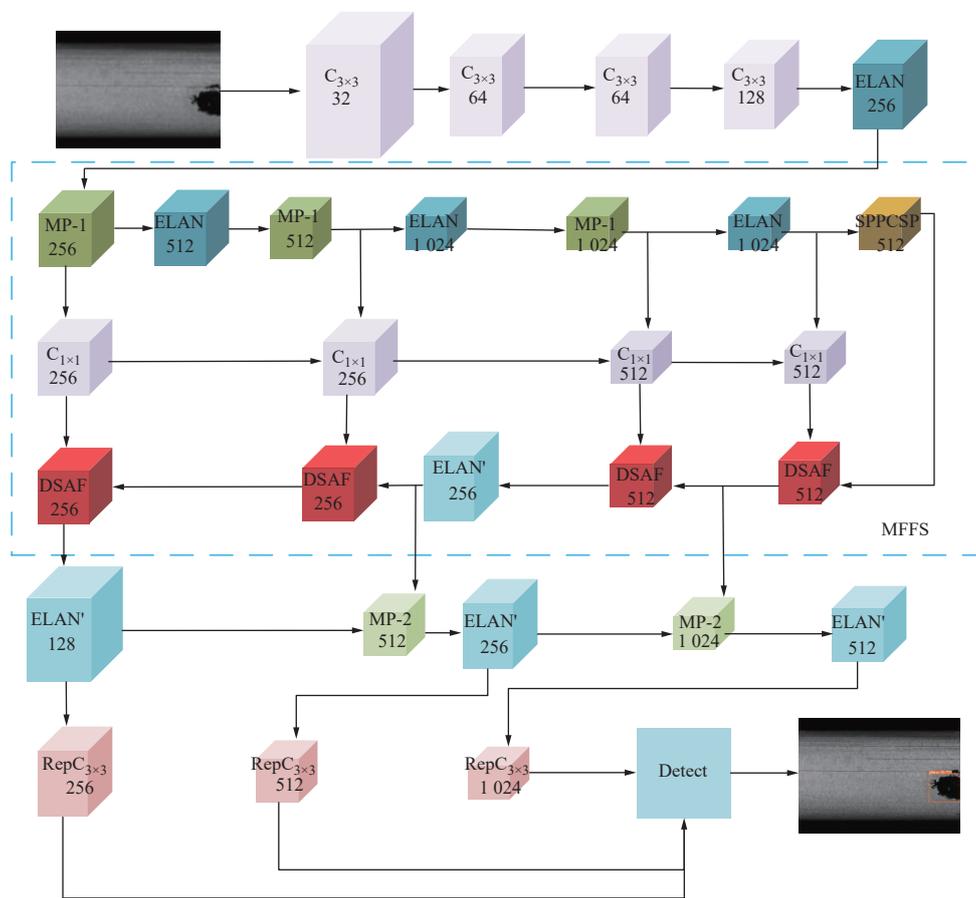


图 1 DSCFM 整体架构

Fig. 1 Structure of DSCFM

模型的主干网络由卷积层、ELAN 层、最大池化层(MP)和 SPPC 模块构成, 图 1 中  $C_{3 \times 3}$  32 表示卷积核大小为  $3 \times 3$ , 卷积个数为 32。输入图像经过 4 个卷积层后进入 ELAN 层, 得到的新特征被送入 MP-1 层和 ELAN 层, MP-1 层和 ELAN 层具

体结构如图 2 所示。MP-1 层的作用是进行下采样, 由两条路径拼接而成。ELAN 是一个高效的网络结构, 它通过控制最短和最长的梯度路径, 使网络能够学习到更多的特征, 并且具有更强的鲁棒性。第 1 条分支是经过 1 个  $1 \times 1$  的卷积做通道



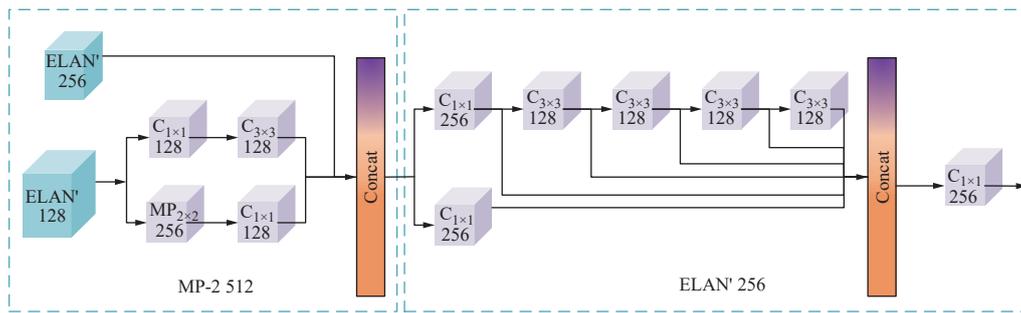


图4 MP-2和ELAN'的结构

Fig. 4 Structure of MP-2 and ELAN'

于网络中得到的多组特征进行融合,充分利用了浅层特征的细节信息和深层特征的语义信息。这种设计不仅保留了丰富的空间细节,还增强了对复杂场景的理解能力,从而提升了模型对多尺度目标的检测性能。其次,MFFS引入了ELAN和SPPCSP模块对深层特征 $C_5$ 进行双重处理,生成 $C'_5$ 、 $C''_5$ 。SPPCSP模块通过多尺度池化和通道数处理,实现了不同尺度特征的融合,进一步增强了特征的表达能力。这种双路径设计有效提升了模型对多尺度目标的适应性和检测精度。另外,MFFS通过对 $C_3$ 、 $C_4$ 、 $C_5$ 进行通道数调整,增加了网络的非线性表达能力,并结合双维自注意力机制进行特征融合,进一步优化了特征间的信息流动。这种动态优化机制能够更好地捕捉目标的关键特征,同时抑制背景干扰,提高了模型的鲁棒

性。最后,在预测端MFFS通过MP-2融合浅层和深层特征,同时处理空间信息和语义信息,增强了模型对目标的定位和分类能力。ELAN'通过卷积结果的拼接,进一步增强了长距离依赖关系,提升了模型对复杂场景的建模能力。

1.2 双维度自注意力特征融合模块

DSAF通过将深层和底层特征分别展开为二维特征图,利用自身的转置矩阵生成通道和空间的注意力图,进而精准地加强关键信息的表达,同时抑制无关噪声,优化了特征融合过程。通过这种双维自注意力机制,DSAF能够动态地调整特征响应,有效捕获长期依赖关系,增强了模型对复杂场景的适应性和解释能力。DSAF的具体结构如图5所示。将深层特征 $X_1 \in \mathbf{R}^{C \times H \times W}$ 和浅层特征 $X_2 \in \mathbf{R}^{C \times H \times W}$ 进行融合,其中: $X_1$ 包含重要的语义信

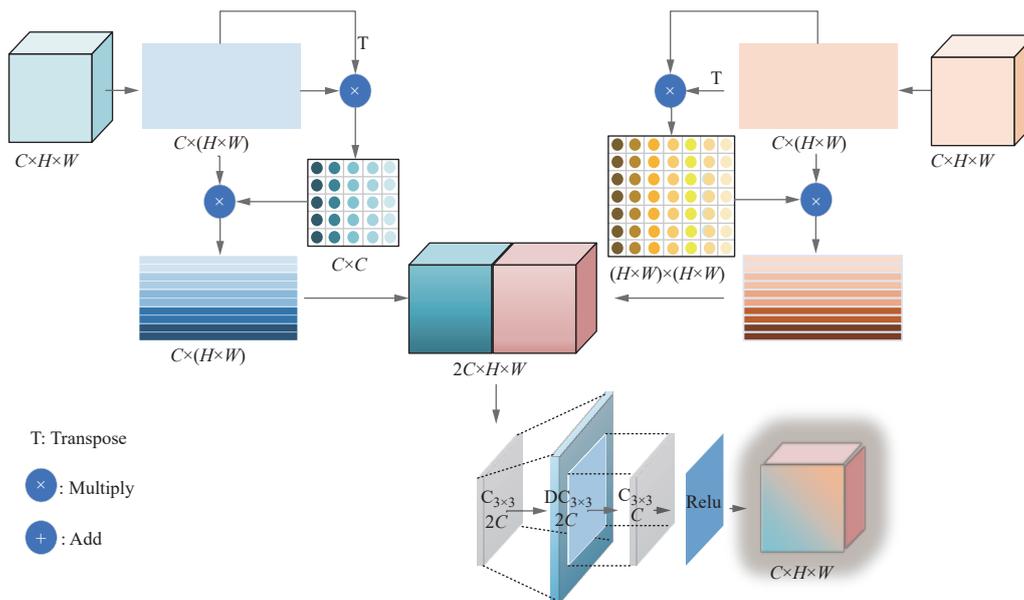


图5 DSAF结构

Fig. 5 Structure of DSAF

息;  $X_2$  包含详细的空间信息。

首先, 通过全局平均池化将输入  $X_1$  和  $X_2$  拉伸为二维矩阵, 得到  $X'_1, X'_2 \in \mathbf{R}^{C \times (H \times W) \times 1}$ ,  $X'_1$  和  $X'_2$  可以由下式表示:

$$X'_1, X'_2 = \text{Reshape}(X_1, X_2) \quad (3)$$

式中, Reshape 为更改矩阵形状操作。

其次, 将生成的  $X'_1$  和  $X'_2$  分为 2 个分支进行处理, 映射到不同的向量空间后计算每个位置的注意力分布, 显然相较于加性模型, 点积模型具有更好的计算效率。具体来说, 将二维矩阵进行转置, 得到  $X_1^T, X_2^T \in \mathbf{R}^{(H \times W) \times C \times 1}$ , 然后  $X'_1$  与  $X_1^T$  相乘,  $X'_2$  与  $X_2^T$  相乘, 分别获得  $X_1$  在通道维度上的节点特征、 $X_2$  在空间维度上的节点特征, 由下式表示:

$$\begin{cases} A_1 = X'_1 X_1^T \\ A_2 = X'_2 X_2^T \end{cases} \quad (4)$$

式中:  $A_1$  为通道维度上的节点特征,  $A_1 \in \mathbf{R}^{C \times C \times 1}$ ;  $A_2$  为空间维度上的节点特征,  $A_2 \in \mathbf{R}^{(H \times W) \times (H \times W) \times 1}$ 。

然后, 将节点特征与原二维矩阵相乘, 再投影回图像特征并在第 1 维度进行拼接, 得到新的富含丰富上下文信息和语义的特征图。接着, 使用反卷积将特征图大小变为原来的 2 倍, 提高分辨率并用卷积使模型学习更加丰富的特征细节, 最后用激活函数将输入特征映射到非线性特征空间, 缓解梯度消失问题。

### 1.3 损失函数

预测头的损失函数包括回归损失  $L_{\text{box}}$ 、置信度损失  $L_{\text{conf}}$  和分类损失  $L_{\text{cls}}$ 。因此, 整体损失函数为:

$$L = \sum_i^Z (L_{\text{box}} + L_{\text{conf}} + L_{\text{cls}}) = \sum_i^Z \left( \sum_j^{B_i} L_{\text{Clou}_j} + \sum_j^{S_i \times S_i} L_{\text{conf}_j} + \sum_j^{B_i} L_{\text{cls}_j} \right) \quad (5)$$

式中:  $Z$  为检测层个数;  $i, j$  为累加数字;  $L_{\text{Clou}}$  为位置损失函数;  $B$  为标签分配到先验框的目标个数;  $S \times S$  为该尺度被分割成的网格数。  $L_{\text{box}}$  和  $L_{\text{cls}}$  对每个目标计算,  $L_{\text{conf}}$  对每个网格计算。

$L_{\text{conf}}$  表示预测目标框内存在目标的概率, 定义如下:

$$L_{\text{conf}} = - \frac{\sum (o_i \ln \hat{c}_i + (1 - o_i) \ln(1 - \hat{c}_i))}{N} \quad (6)$$

$$\hat{c}_i = \text{Sigmoid}(c_i) \quad (7)$$

式中:  $o_i$  为预测目标框和真实标注框的交并比 IoU;  $c$  为预测值;  $N$  为正样本和负样本的总数; Sigmoid 为激活函数。

$L_{\text{cls}}$  表示预测目标的类别, 定义如下:

$$L_{\text{cls}} = - \frac{\sum_{i \in \text{pos}} \sum_{j \in \text{class}} (O_{ij} \ln \hat{C}_{ij} + (1 - O_{ij}) \ln(1 - \hat{C}_{ij}))}{N_{\text{pos}}} \quad (8)$$

$$\hat{C}_{ij} = \text{Sigmoid}(C_{ij}) \quad (9)$$

式中:  $C_{ij}$  为将目标  $i$  预测为第  $j$  类的概率;  $O_{ij} \in \{0, 1\}$  为包含第  $i$  类的目标框中是否存在第  $j$  类目标, 存在为 1, 否则为 0;  $N_{\text{pos}}$  为正样本的数量; pos 为正样本; class 为类别。

CIoU 损失的计算公式如下:

$$\begin{cases} L_{\text{Clou}} = 1 - \text{IoU} + \frac{\rho^2}{c^2} + \alpha v \\ \text{IoU} = \frac{|G \cap B|}{|G \cup B|} \\ v = \frac{4}{\pi^2} \left( \arctan \frac{w_g}{h_g} - \arctan \frac{w_b}{h_b} \right)^2 \end{cases} \quad (10)$$

式中:  $\rho$  为计算两个框的中心点距离;  $c$  为两个框边界的最远距离;  $\alpha$  为权重系数;  $v$  为用于衡量框的长宽比一致性的参数;  $B, G$  分别为预测框和真实框;  $w_g, h_g$  分别为真实框的宽和高;  $w_b, h_b$  分别为预测框的宽和高。

CIoU 损失的计算过程如图 6 所示。

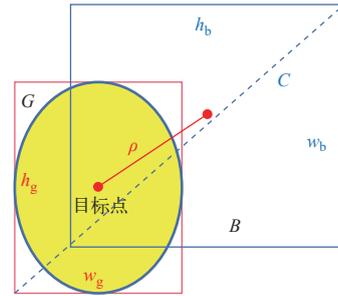


图 6 CIoU 损失的计算过程  
Fig. 6 Process of calculating CIoU loss

## 2 实验结果与分析

### 2.1 实验数据

本文所用的燃料棒外观缺陷数据集图像共 892 张, 外观缺陷包括划伤、磕伤及异物 3 种缺陷。采用 Mosaic、Mixup 等方式进行数据增强后, 数据集扩充到 4 025 张。其中, 训练集、验证集、测试集占比约为 8 : 1 : 1, 即训练集 3 219 张, 验证集 403 张, 测试集 403 张。

## 2.2 实施细节与评价指标

本文实验所使用的代码环境为 Python 3.8, 使用 PyTorch 1.8、CUDA 11.2 深度学习框架来构建网络, 并在 GPU 型号为 NVIDIA GeForce RTX 3090(显存 24G)、CPU 型号为 Intel(R) Core(TM) i9-9900K CPU@3.6 GHz 的服务器上进行训练。批次训练数据量设置为 16, 训练次数设置为 300。在训练阶段, 输入图像的大小被调整为 640 像素×640 像素, 具体实验参数信息如表 1 所列。

表 1 实验参数信息

Table 1 Experimental parameter information

参数	参数值
训练次数	300
批次训练数据量	16
学习率	0.01
动量	0.937
权重衰减	0.000 5

为了全面评估模型的表现, 本文实验使用 precision(精准率)、recall(召回率)、AP(平均精度)、mAP(mean average precision, 平均精度均值, 即各类别 AP 的平均值)、FPS(frames per second, 每秒帧率, 即每秒处理的张数)这些指标来评价模型的性能和鲁棒性。其中精准率、召回率、平均精度和平均精度均值的计算公式如下:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (11)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (12)$$

$$\text{AP} = \int_0^1 p(r)dr \times 100\% \quad (13)$$

$$\text{mAP} = \frac{\sum_{i=1}^N \text{AP}}{N} \times 100\% \quad (14)$$

式中: TP 为模型正确预测为正例的样本数量; FP 为模型错误预测为正例的样本数量; FN 为模型错误预测为反例的样本数量;  $N$  为样本的类别数;  $p$  为精准率;  $r$  为召回率。

## 2.3 结果对比

表 2 所列为不同模型的 AP、mAP、FPS 等指标的对比, 可以看出, 在其他 YOLO 系列检测模型的结果中 YOLOv5s 的 mAP 最高, 达到了 80.9%, 效果最好; 而与本文所提的 DSCFM 相比, DSCFM 的 mAP 达到了 82.0%, 比 YOLOv5s 的 mAP 高出 1.1%, 这说明 DSCFM 检测精度水平更高, 鲁棒性更好。

在与其他 YOLO 系列模型的性能对比中, YOLOv7 的 FPS 为 121  $\text{s}^{-1}$ , 速度最低; YOLOv5s 的 FPS 为 270  $\text{s}^{-1}$ , 速度最高; 而本文模型的 FPS 达到了 195  $\text{s}^{-1}$ , 虽然略低于 YOLOv5s, 但在保证较高检测精度的同时, 仍然展现出显著的实时性优势。相较于 YOLOv7, 本文模型的 FPS 提升了 61.2%, 显著优于其性能; 而与 YOLOv5s 相比, 本文模型在速度上仅稍逊一筹, 但在复杂场景下的检测精度和鲁棒性方面更具优势。这种平衡了速度与精度的设计, 使得本文模型在实际应用中具有更强的竞争力和实用性, 尤其适用于对实时性和准确性要求较高的燃料棒外观质量视觉检测场景。

## 2.4 主观实验

不同模型检测结果如图 7 所示, 在异物缺陷检测中, 所有模型均能有效检出, 但在磕伤和划

表 2 与其他先进检测模型的结果对比

Table 2 Comparison with other advanced detection model

模型	precision/%	recall/%	AP/%			mAP/%	FPS/ $\text{s}^{-1}$
			划伤	磕伤	异物		
YOLOv5s	80.9	82.8	73.1	79.2	90.3	80.9	270
YOLOv5n	82.3	70.4	66.8	78.5	81.0	75.4	227
YOLOv5m	88.1	73.4	65.3	85.7	87.4	79.5	230
YOLOv5l	82.7	68.3	63.2	81.9	84.3	76.5	228
YOLOv5x	87.0	69.2	66.1	83.6	86.4	78.7	212
YOLOv7	75.7	68.8	69.1	71.2	83.9	74.7	121
YOLOv8n	86.9	75.2	76.1	73.1	90.4	79.9	188
DSCFM	81.7	77.9	71.4	84.0	90.6	82.0	195

伤缺陷检测中,本文所提的 DSCFM 展现出显著优势。对于磕伤缺陷, YOLOv5s 模型将细微磨痕误检为磕伤,出现了明显的误检情况,而 DSCFM 则能够准确区分细微磨痕与磕伤,避免了误检问题;对于划伤缺陷,其他模型仅能检测出一个划伤,存在漏检现象,而 DSCFM 能够准确检测出图

像中的两个划伤缺陷,表现出更高的检测精度和鲁棒性。这些结果表明, DSCFM 在复杂缺陷检测任务中具有更强的适应性和准确性,尤其在区分细微缺陷和避免漏检方面表现突出,充分体现了其在燃料棒外观缺陷检测中的实用价值和性能优势。

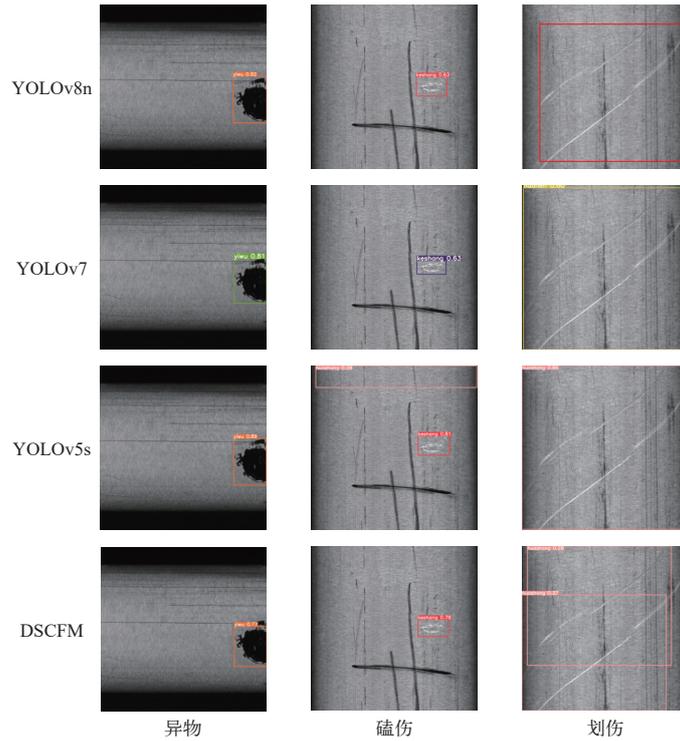


图7 不同模型检测结果可视化

Fig. 7 Visualization of detection result for different models

### 3 结论

本文提出了一种基于双维自注意力的跨阶段融合模型 DSCFM, 构建了一种新的多尺度特征融合结构 MFFS 作为模型颈部对特征进行融合处理。为了学习底层的结构信息, 并且对深层信息进行整合, 提出了一个双维度自注意力特征融合模块。实验表明, DSCFM 在燃料棒外观缺陷数据集上取得了先进的性能, 具有良好的检测能力, 能够有效抑制复杂背景的干扰, 高效地提取缺陷特征, 其 mAP 达到了 82.0%, 召回率达到了 77.9%。该研究方法可为今后燃料棒外观检测图像的高精度自动分析评价提供基础。

#### 参考文献:

[1] 汤琪, 王华才, 梁政强. 反应堆乏燃料棒的 X 射线数字

成像检测工艺及其应用[J]. 无损检测, 2020, 42(10): 14-17.

TANG Qi, WANG Huacai, LIANG Zhengqiang. Application of X-ray digital imaging technology to the examination of irradiated fuel rods from nuclear reactor[J]. *Nondestructive Testing Technology*, 2020, 42(10): 14-17(in Chinese).

[2] 高三杰, 肖湘, 秦梦淋, 等. 燃料组件破损棒自动超声检查系统关键技术研究[J]. 中国测试, 2024, 50(增刊 1): 296-300.

GAO Sanjie, XIAO Xiang, QIN Menglin, et al. Research on key technology of automatic ultrasonic inspection system for fuel assembly failure[J]. *China Measurement & Test*, 2024, 50(Suppl. 1): 296-300(in Chinese).

[3] 张小刚, 俞东宝, 汤慧, 等. 基于深度学习的 X 射线燃料棒端塞缺陷自动检测方法研究[J]. 原子能科学技术, 2024, 58(8): 1767-1776.

- ZHANG Xiaogang, YU Dongbao, TANG Hui, et al. Detection method of X-ray fuel rod end plug defect based on deep learning[J]. *Atomic Energy Science and Technology*, 2024, 58(8): 1767-1776(in Chinese).
- [4] 李伟, 刘鹏亮, 王从政, 等. 压水堆燃料辐照后燃料棒间距数据机器视觉测量方法研究[J]. *核动力工程*, 2022, 43(4): 53-59.
- LI Wei, LIU Pengliang, WANG Congzheng, et al. Research on machine vision measurement method of rod spacing data for PWR fuel after irradiation[J]. *Nuclear Power Engineering*, 2022, 43(4): 53-59(in Chinese).
- [5] 朱永利, 张小刚, 俞东宝, 等. 绕丝燃料棒自动外观缺陷检测系统设计[J]. *设备监理*, 2024(1): 38-41, 68.
- ZHU Yongli, ZHANG Xiaogang, YU Dongbao, et al. Design of an automatic appearance defect detection system for wire wound fuel rods[J]. *Plant Engineering Consultants*, 2024(1): 38-41, 68(in Chinese).
- [6] 吴之望. 基于深度学习的核燃料组件缺陷检测[D]. 北京: 华北电力大学, 2020.
- [7] 赵年甫, 王霖, 王向军, 等. 基于改进二阶段检测网络的长时跟踪重检测方法[J]. *应用光学*, 2023, 44(4): 768-776.
- ZHAO Nianfu, WANG Lin, WANG Xiangjun, et al. Redetection method for long-term tracking based on improved two-stage detection networks[J]. *Journal of Applied Optics*, 2023, 44(4): 768-776(in Chinese).
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. USA: IEEE, 2014.
- [9] REN S, HE K, 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.
- [10] ZHANG H, SHAO F, CHU W, et al. Faster R-CNN based on frame difference and spatiotemporal context for vehicle detection[J]. *Signal, Image and Video Processing*, 2024, 18(10): 7013-7027.
- [11] 陈法法, 刘咏, 潘瑞雪, 等. 多尺度特征融合改进 Faster RCNN 的铝材表面缺陷辨识[J]. *组合机床与自动化加工技术*, 2024(5): 166-170.
- CHEN Fafa, LIU Yong, PAN Ruixue, et al. Multi-scale feature fusion to improve faster RCNN for aluminum surface defect recognition[J]. *Modular Machine Tool & Automatic Manufacturing Technique*, 2024(5): 166-170(in Chinese).
- [12] 黄健宸, 王晗, 卢昊. 结合轻量化骨干与多尺度融合的单阶段检测器[J]. *中国图象图形学报*, 2022, 27(12): 3596-3607.
- HUANG Jianchen, WANG Han, LU Hao. One-stage detectors combining lightweight backbone and multi-scale fusion[J]. *Journal of Image and Graphics*, 2022, 27(12): 3596-3607(in Chinese).
- [13] 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). USA: IEEE, 2016.
- [14] ADLI T, BUJAKOVIĆ D, BONDŽULIĆ B, et al. A modified YOLOv5 architecture for aircraft detection in remote sensing images[J]. *Journal of the Indian Society of Remote Sensing*, 2024: 1-16.
- [15] LING H, ZHAO T, ZHANG Y, et al. Engineering vehicle detection based on improved YOLOv6[J]. *Applied Sciences*, 2024, 14(17): 8054-8068.
- [16] LI X, XUE S, LI Z, et al. A candy defect detection method based on StyleGAN2 and improved YOLOv7 for imbalanced data[J]. *Foods*, 2024, 13(20): 3343-3360.
- [17] 戴林华, 黎远松, 石睿. 基于改进 YOLOv8n 算法的水稻叶片病害检测[J]. *湖北民族大学学报(自然科学版)*, 2024, 42(3): 382-388.
- DAI Linhua, LI Yuansong, SHI Rui. Detection of rice leaf diseases based on improved YOLOv8n algorithm[J]. *Journal of Hubei Minzu University (Natural Science Edition)*, 2024, 42(3): 382-388(in Chinese).
- [18] GAO Z, YU X, RONG X, et al. Improved YOLOv8n for lightweight ship detection[J]. *Journal of Marine Science and Engineering*, 2024, 12(10): 1774-1794.
- [19] ZHANG Y, XU M, ZHU Q, et al. Improved YOLOv5s combining enhanced backbone network and optimized self-attention for PCB defect detection[J]. *The Journal of Supercomputing*, 2024, 80(13): 19062-19090.
- [20] 王宸, 杨帅, 周林, 等. 基于自适应多尺度特征融合网络的金属齿轮端面缺陷检测方法研究[J]. *电子测量与仪器学报*, 2023, 37(10): 153-163.
- WANG Chen, YANG Shuai, ZHOU Lin, et al. Research on metal gear end-face defect detection method based on adaptive multi-scale feature fusion network[J]. *Journal of Electronic Measurement and Instrumentation*, 2023, 37(10): 153-163(in Chinese).
- [21] 帅文强. 基于反卷积神经网络的遥感图像融合算法研究[D]. 西安: 西安电子科技大学, 2023.
- [22] 唐嘉男, 孟祥瑞. 基于多尺度信息提取和特征融合的皮肤镜图像分割算法[J]. *湖北民族大学学报(自然科学版)*, 2024, 42(2): 226-232.
- TANG Jianan, MENG Xiangrui. Dermoscopic image segmentation algorithm based on multi-scale information extraction and feature fusion[J]. *Journal of Hubei Minzu University (Natural Science Edition)*, 2024, 42(2): 226-232(in Chinese).