

聚焦边缘与多尺度特征的轻量化违禁品检测^{*}

赵小涛^{1,2} 李新伟^{1,2}

(1.河南理工大学电气工程与自动化学院 焦作 454000;2.河南省煤矿装备智能检测与控制重点实验室 焦作 454000)

摘要: 针对X射线安检图像中背景复杂、尺度多变、小尺寸目标难以检测等挑战,提出一种聚焦边缘与多尺度特征的轻量化违禁品检测算法LEM-YOLO。首先,设计轻量化边缘特征增强模块(LEFE),增强模型的边缘特征提取能力。其次,设计高效多级特征融合金字塔网络,利用动态上采样(Dysample)和层次尺度特征金字塔网络(HS-FPN),增强多尺度特征融合并减少计算冗余,同时设计动态特征编码模块(DFE),保留小尺寸目标的全局信息。最后,使用Shape-IoU作为边界框回归损失函数,聚焦边框形状和自身尺度,提升目标定位精度。在公开数据集SIXray上进行实验,结果表明,LEM-YOLO在违禁品检测中的mAP达到了94.63%,比原算法提升了2.56%,同时模型体积下降了50.67%,更好地满足了违禁品检测场景的需求。

关键词: 违禁品检测;边缘特征;高效多级特征融合金字塔;轻量化;YOLOv8

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:** 520.6040

Lightweight contraband detection focusing on edge and multi-scale features

Zhao Xiaotao^{1,2} Li Xinwei^{1,2}

(1. School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China;

2. Henan Key Laboratory of Intelligent Detection and Control of Coal Mine Equipment, Jiaozuo 454000, China)

Abstract: To tackle challenges such as complex backgrounds, scale variability, and difficulty in detecting small objects in X-ray security inspection images, a lightweight contraband detection algorithm named LEM-YOLO is proposed, focusing on edge and multi-scale features. Firstly, a lightweight edge feature enhancement (LEFE) module is designed to strengthen edge feature extraction. Secondly, an efficient multi-level feature fusion pyramid network is developed, incorporating dynamic upsampling (Dysample) and the hierarchical scale feature pyramid network (HS-FPN) to enhance multi-scale feature fusion while reducing computational redundancy. Additionally, a dynamic feature encoding (DFE) module is used to preserve global information for small objects. Finally, Shape-IoU is employed as the bounding box regression loss function, concentrating on boundary shape and scale to improve localization accuracy. Experimental results on the public SIXray dataset demonstrate that LEM-YOLO achieves a mAP of 94.63%, a 2.56% increase over the original algorithm, while reducing model size by 50.67%, making it more suitable for contraband detection scenarios.

Keywords: contraband detection; edge features; efficient multi-level feature fusion pyramid; lightweight; YOLOv8

0 引言

安检是识别机场、地铁站、海关等重要公共场所及敏感部门潜在风险的关键措施之一。目前,违禁品检测主要依赖安检设备获取行李的X射线图像,并由经过专业培训的安检人员进行人工检查。然而,在客流高峰时期,包裹内的物品错综复杂,安检人员面临误检、漏检的风险,从而危及人身安全。因此,开发准确、快速的智能化违禁品检测方法

已经成为安检领域的重要挑战^[1]。

近年来,以卷积神经网络(convolutional neural networks,CNN)为代表的深度学习技术在图像分类、目标检测、语义分割等多个领域得到了广泛应用,并取得了显著成果^[2]。将卷积神经网络应用于X射线安检图像分析,是推动智能安检技术发展的关键因素。从实际工程应用的角度来看,X射线图像中的违禁品检测属于目标检测任务。安检人员不仅需要识别包裹中违禁品的类别,还需要精准

定位违禁品在行李中的具体位置,以便及时排除潜在的危险源。在违禁品目标检测领域,相关学者从不同角度研究了违禁品图像检测的难点,主要集中在以下方面:

违禁品图像背景复杂,物品轮廓常常相互重叠。对此,董乙杉等^[3]提出了一种改进 YOLOv5 的 X 光违禁品检测算法,通过引入融合注意力机制,增强了网络对违禁品重要特征的提取能力,抑制了冗余信息的干扰。张良等^[4]在 YOLOv5s 的基础上引入了 Transformer 模块,以增强对局部信息的获取能力,并设计了感受野自适应模块,提升了网络在复杂背景下的检测精度。Liu 等^[5]设计了自适应空间模块和通道注意力模块,通过关注空间信息和通道信息,增强了模型对重要特征的提取能力,提高了算法的检测精度。李松等^[6]提出了一种改进 YOLOv7 的 X 射线图像违禁品检测算法,针对检测过程中因物体遮挡导致的干扰,通过增强网络对不同通道特征的关注度,提高了算法的检测性能。上述算法在处理违禁物品遮挡方面具有一定的能力,但在检测边缘信息较弱的违禁品图像时,效果仍不理想。

违禁品的尺度变化较大,且小尺寸目标难以检测。对此,曹洋等^[7]设计了一种基于 ResNet50 的自适应学习多尺度违禁品识别算法,通过使用空洞卷积与可变形的空洞金字塔结构来增强模型的多尺度特征提取能力。穆思奇等^[8]基于 YOLOv4,结合空洞卷积和注意力机制,并采用特征图融合的策略,提高了模型的多尺度特征提取能力。Sima 等^[9]提出了一种多尺度特征注意力网络,将金字塔特征提取结构与自注意力模块相结合,以捕捉全局信息与多尺度特征之间的相关性。粟兴旺等^[10]通过引入可变形卷积来适应物体的尺度变化,并构建混合卷积注意力模块,以获取全局上下文信息,提升模型的形变感知能力。王海群等^[11]基于 YOLOv8,通过增加目标检测头来实现对小尺寸违禁品的精准检测,并采用不同卷积核生成多个特征图,从而使模型能更好地捕捉不同尺度的目标信息。颜志明等^[12]在 YOLOv8 的基础上,通过采用密集连接的方法进行金字塔池化,从而实现多尺度特征的提取,提升了模型对不同大小违禁品的检测精度。上述算法在一定程度上缓解了违禁品尺度多变的问题,但也增加了计算成本,导致模型体积变大。

现有的违禁品检测算法通常模型体积大、计算成本高,在边缘设备上部署困难。对此,Song 等^[13]在 YOLOv5 中引入了 CGhost 模块替代 C3 模块,以减少特征提取过程中的参数量,并设计 Stem 模块以提取重要特征,从而提高了算法的检测性能。左景等^[14]提出了一种多分支轻量化卷积结构来改进违禁品检测模型,采用轻量化双分支结构来降低模型的参数量,并通过信息补偿分支来弥补轻量化带来的精度损失。上述算法在轻量化效果上表现不够显著,同时检测精度较低,难以满足实际应用中对违禁品检测的要求。

以上研究表明,现有的 X 光违禁品检测算法主要面临

以下几个问题:首先,违禁品图像中背景与物体往往交织在一起,使得普通目标检测算法难以提取关键特征,从而导致对违禁品的识别不准确;其次,违禁品的尺度变化较大,图像中包含较多小尺寸目标,同一类型的违禁品可能以不同位置和大小的形式出现,这增加了检测的难度并导致检测精度下降;最后,现有的违禁品检测算法通常模型体积大,计算成本高,难以在边缘设备上高效部署。针对上述问题,本文以 YOLOv8s 为基线模型,提出了一种聚焦边缘与多尺度特征的轻量化违禁品检测算法 LEM-YOLO,其主要贡献为:

1)设计 EFE_C2f 替换骨干网络中的 C2f,通过将 Sobel 算子与 PConv^[15]有效结合,充分提取边缘特征和空间特征,增强了模型对边缘信息的关注度,有效抑制复杂背景信息的干扰。

2)重新设计 YOLOv8 的特征融合部分,提出了高效多级特征融合金字塔网络(EM-FPN),在实现模型轻量化的同时,利用动态上采样 Dysample^[16]减少关键信息的丢失,并采用动态特征编码(dynamic feature encoding, DFE)模块对不同级别的语义特征进行动态融合,保留小尺寸目标的全局信息,增强模型对尺度变化信息的敏感度。

3)将 CIoU 替换为 Shape-IoU^[17]作为新的回归损失函数,使得损失计算更聚焦于边框形状和自身尺度,从而提升了目标的定位精度。

1 YOLOv8 模型

YOLOv8 是一种先进的深度学习模型,相较于 YOLOv5, YOLOv8 集成了更多先进的架构,能够在检测精度与速度之间实现更好的平衡。YOLOv8 的模型结构包括 3 个主要部分:骨干网络(Backbone)、特征融合网络(Neck)和检测头(Head)。骨干网络负责图像特征的提取,由 CBS(卷积+BN+SiLU)、C2f(跨阶段部分块)和 SPPF(空间金字塔池化层)组成。这些模块的组合不仅提高了特征提取的效率,还优化了模型的计算复杂度。其中,C2f 模块通过分割特征图并将其部分通过密集卷积块,提升了特征重用能力,并降低了计算量,而 SPPF 模块则通过多尺度的池化操作,增强了网络对不同尺度目标的检测能力。特征融合网络不仅继承了特征金字塔网络(FPN)的顶到底路径,还额外增加了底到顶的路径,以提高语义特征的传播和细节特征的提取能力,增强了不同层级特征之间的信息流动和融合。检测头将 Anchor-base 换成了 Anchor-free 解耦头,省去了传统 YOLO 模型中依赖的预定义锚框,通过直接预测目标的中心点和边界框,从而简化了模型的结构。

2 LEM-YOLO 算法

针对违禁品图像背景复杂、物体尺度变化大、小尺寸目标难以检测、部署困难等问题,本研究以 YOLOv8s 为基准模型,在 Backbone 与 Neck 部分进行轻量化设计,并对回归

损失进行优化,提出了LEM-YOLO算法,如图1所示。首先,在Backbone中引入了轻量级边缘特征增强跨阶段部分模块(EFE_C2f),替换原有的C2f模块,从而提升了模型对边缘信息和空间信息的综合提取能力。其次,在Neck部分设计了高效多级特征融合金字塔网络EM-FPN,通过结

合动态上采样Dysample和层次尺度特征金字塔网络HS-FPN^[18],实现了上下文信息的高效融合,同时通过DFE模块增强模型对小尺寸目标的语义特征提取能力。最后,对于回归损失使用Shape-IoU来提高边界框的定位效果与检测精度。

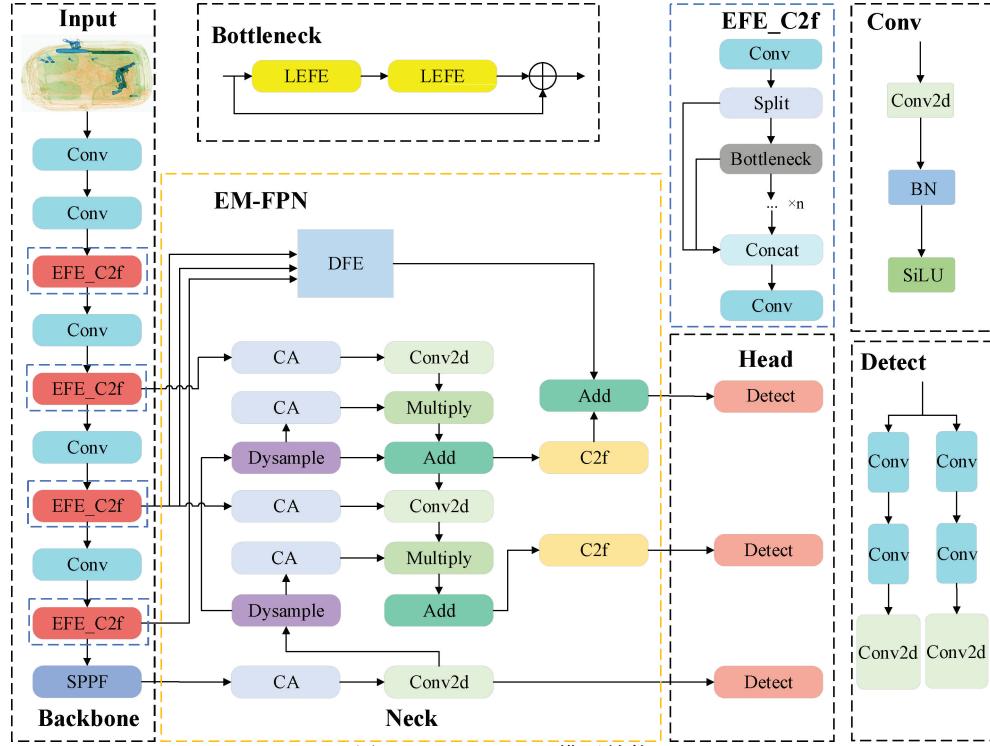


图1 LEM-YOLO模型结构

Fig. 1 Structure of the LEM-YOLO model

2.1 骨干网络改进

在YOLOv8的骨干网络中,C2f模块通过将输入特征图分为两部分,其中一部分直接进行卷积操作,另一部分通过跨阶段连接与前一部分融合。这种设计使得特征在多个层级间得到复用,从而提升了模型的效率和表示能力。不同于普通图像,X射线图像通过密度差异来显示物体的轮廓,边缘特征是区分物体类别的关键信息。在面对背景复杂的X射线安检图像时,物体之间经常出现相互遮挡和重叠的现象,导致边缘特征不明显。然而,C2f模块与常规特征提取模块类似,对于复杂背景图像的边缘信息提取效果不佳,难以有效提取违禁品的边缘特征,因此不适合用于违禁品检测任务。

为了解决这一问题,本研究设计了轻量化边缘特征增强模块(lightweight edge feature enhancement module, LEFE)其结构如图2所示。

LEFE模块主要由提取边缘信息的SobelConv分支、提取空间信息的PConv分支和跳跃连接组成。首先,Sobel滤波器以3D组卷积的形式对输入特征图在水平方向和垂直方向上进行边缘信息的高效提取,保留了违禁品的关键边缘特征。其次,PConv分支仅对部分通道进行卷

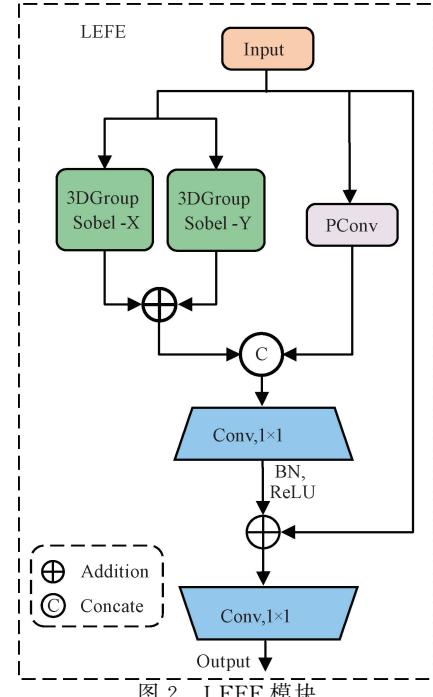


图2 LEFE模块

积操作,在保留全局空间细节的同时,最大限度地减少计算成本。最后,将这两种特征在通道维度上拼接,通过卷积层进一步融合后再与保留原始特征的跳跃连接相结合,在突出边缘的同时,确保不丢失物体的整体特征。这种针对违禁品检测任务设计的创新模块使得模型在处理复杂的 X 射线图像时,能够精确地捕捉到违禁品的边界特征,从而提高识别的准确性。特别是在面对形状复杂、边缘模糊或重叠的物体时,LEFE 模块通过其边缘增强功能,显著提高了模型对图像中遮挡物体边缘信息的敏感性。

将 LEFE 模块嵌入到 C2f 模块中得到了 EFE_C2f 模块,如图 3 所示。与 C2f 相比,EFE_C2f 模块提升了特征提取的效率,降低了计算开销,适合在资源受限的环境中进行实时检测。

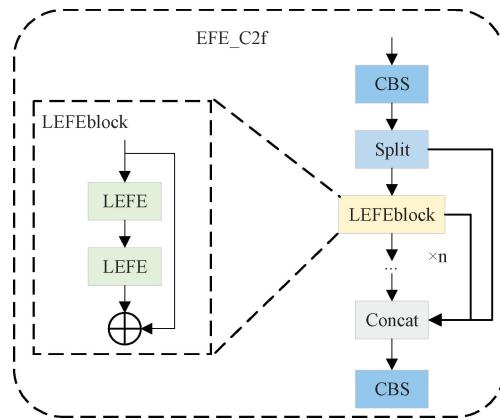


图 3 EFE_C2f 模块

Fig. 3 EFE_C2f module

1) Sobel 算子

图像的特征提取通常从边缘开始,并逐步构建更高层次的特征描述。为有效提取边缘特征,在主干网络中引入 Sobel 算子进行图像的边缘检测。Sobel 算子是一种边缘检测算法,利用一阶微分等效算子检测图像的梯度并确定边缘位置。由于在运算中引入了类似局部平均的处理,Sobel 算子对噪声具有平滑效果,能够有效减少噪声的影响。Sobel 算子包括水平方向和垂直方向的卷积模板,其公式分别如式(1)和式(2)所示,用于计算图像在 X 轴和 Y 轴方向上的梯度。

$$A_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 2 & 1 & 1 \end{bmatrix} \quad (1)$$

$$A_2 = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \quad (2)$$

卷积模板中的像素对应于图像的每个像素点,通过将周围像素值与模板系数相乘并求和,从而得到新的灰度值。接着,对图像中的每个像素点进行离散卷积运算 $g(x, y)$,如式(3)所示,并用两个卷积模板的最大卷积结

果来替代中心像素的灰度值。

$$g(x, y) = \begin{bmatrix} \frac{\partial f}{\partial \xi} & \frac{\partial f}{\partial \psi} \end{bmatrix}^T \quad (3)$$

最后,根据设定的阈值 T ,将处理后的灰度值与阈值进行比较,如果最大值超过阈值,该像素被视为图像的边缘,否则视为背景。网络经过 Sobel 算子处理后的效果图如图 4 所示。

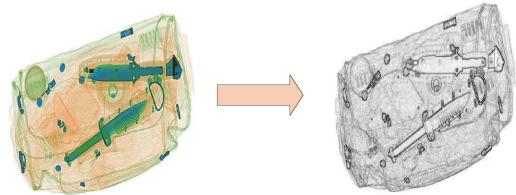


图 4 边缘检测图

Fig. 4 Edge detection diagram

2) PConv

X 射线图像中物体与违禁品相互遮挡,部分违禁品的边缘信息较少,仅仅使用边缘检测算子难以对违禁品特征有效提取。普通卷积能够提取图像中的空间信息和纹理信息,但这种方式往往会引入大量冗余信息,导致计算资源增加,模型体积变大。为了优化这一问题,使用幻影卷积(ghost convolution, GConv)和深度可分离卷积(depthwise separable convolution, DSConv)成为一种有效的策略。然而,这些卷积方法虽然在降低模型复杂性方面表现出色,但也削弱了卷积层对复杂特征的捕捉能力,导致在处理高维特征或复杂纹理信息时,无法充分提取细节特征,从而导致检测精度的下降。部分卷积(PConv)在保留常规卷积提取空间特征能力的同时,减少了冗余计算,在计算量与计算速度之间达到了良好的平衡。

PConv 的结构如图 5 所示,特征图的不同通道之间往往存在高度相似性,PConv 通过仅对部分输入通道应用常规卷积,而对其余通道不进行操作,从而减少了计算量和内存访问。普通卷积与 PConv 的 FLOPs 如式(4)、(5)所示。

$$FLOPs = h \times w \times k^2 \times c^2 \quad (4)$$

$$FLOPs = h \times w \times k^2 \times c_p \quad (5)$$

其中, h, w, k, c 分别代表了特征图的高、宽、卷积核、通道数, c_p 为选择的部分通道数。如果 $c_p = c/4$, PConv 的 FLOPs 仅为普通卷积的 $1/16$ 。

2.2 颈部网络改进

在违禁品数据集中,违禁品检测任务面临尺度多变和小尺寸目标难以检测的挑战。这种问题源于不同类型的违禁品尺寸差异较大,即使是相同类型的违禁品在不同安检机下的成像也会出现尺寸差异。

YOLOv8 的颈部网络通过 FPN+PAN 结构进行特征融合,然而,这种技术没有进行特征选择,只是将多个特征层的像素值简单相加。当面对 X 射线图像中违禁品的多

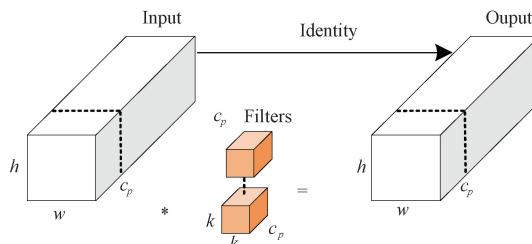


图 5 PCConv 结构
Fig. 5 PCConv structure

变尺度时,容易受到尺度冗余信息的干扰,无法有效提取违禁品尺度变化的共性特征。

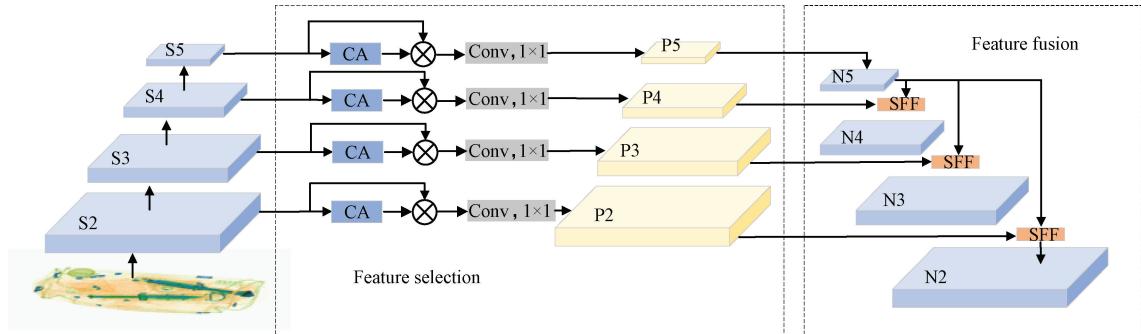


图 6 HS-FPN 结构
Fig. 6 HS-FPN structure

为此,本研究在 HS-FPN 的基础上提出了高效多级特征融合金字塔网络 (efficient multi-level feature fusion pyramid network, EM-FPN),结构如图 1 所示。EM-FPN 使用 DySample 进行动态上采样,由于其动态调整的特性,可以有效地减少图像放大不均匀的情况,生成的图像质量更高。针对小尺寸目标不易检测的特点,设计了动态特征编码模块 (dynamic feature encoding module, DFE),通过动态提取高、中、低层融合后的特征,加强了模型对违禁品小尺寸目标的特征提取能力。EM-FPN 根据违禁品固有的尺寸差异进行设计,增强了模型对违禁品尺度多变信息的敏感度,进一步提升了违禁品对不同尺度目标的检测能力。

1) DySample

上采样是网络架构的关键组成部分,它可以将低分辨率特征提升到与高分辨率特征相同的尺度,使模型更准确地定位和识别物体的细节。随着动态网络的普及,一些动态上采样模块(如 CARAFE)在多个任务上取得了良好的性能提升,但它们引入了大量的计算成本,限制了其在违禁品检测任务中实时应用的可能性。为了解决这一问题,本文在颈部网络中引入了 DySample 动态上采样模块,这是一种轻量、高效的动态上采样算子,能够以较低的计算资源实现高效的上采样,通过仅保留差异显著的像素值,可以有效减少数据量,降低计算复杂度,提高上采样速度和效率。DySample 使用动态因子抽样法进行特征图处理

层次尺度特征金字塔网络 HS-FPN 专门针对白细胞的尺度差异进行设计,具备一定的多尺度特征提取能力。HS-FPN 的结构如图 6 所示,主要包括特征选择模块和特征融合模块。HS-FPN 通过引入特征选择模块,将高级语义特征作为权重对低级特征进行过滤。然后特征融合模块将过滤后的特征与高级语义特征逐点相加,实现多尺度特征融合,从而提高了模型的特征表达能力。然而,HS-FPN 在上采样时使用反卷积,这容易导致像素变化剧烈,生成不均匀的特征图,干扰模型对 X 射线图像中违禁品特征的准确捕捉。此外,HS-FPN 缺乏针对小尺寸目标的专门设计,难以精确提取小尺寸目标的细节信息。

的过程如图 7 所示。

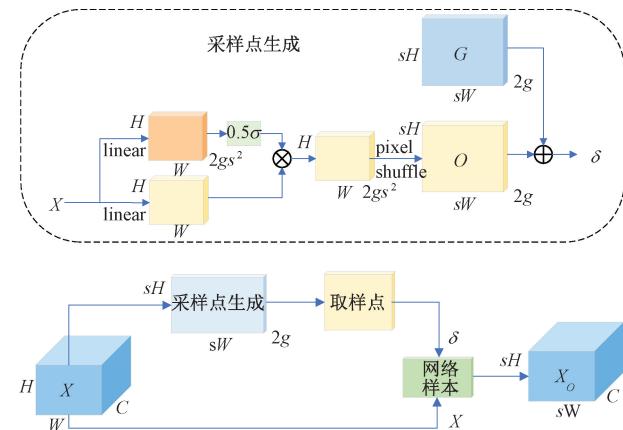


图 7 基于动态因子的 DySample 模块
Fig. 7 DySample module based on dynamic factors

给定上采样比例因子 s 和大小为 $C \times H \times W$ 的特征映射 X ,经过两个线性层后相乘,再经过像素重洗 (pixel shuffle)生成偏移量 O ,如式(6)所示。

$$O = P(0.5\sigma(\text{linear}(X)) \cdot \text{linear}(X)) \quad (6)$$

其中, P 为像素洗牌, σ 为 sigmoid 函数。

偏移量 O 和原始采样网格 g 共同构成采样集 δ ,最后将采样集 δ 与原始特征映射 X 来生成上采样后的特征图 X_o ,如式(7)、(8)所示。

$$\delta = g + O \quad (7)$$

$$X_o = \text{grid_sample}(\delta, X) \quad (8)$$

2) DFE

YOLOv8 结构中的 P3 检测头通过多层特征融合, 包含丰富的底层特征信息, 对小目标的精确定位至关重要。然而, 在违禁品数据集中, 由于存在较多的小尺寸目标, 仅通过原有的 P3 检测层难以充分利用低层特征中丰富的空间信息和通道信息。许多学者通过额外引入 P2 检测层来进一步提升小目标的检测能力, 但这种方法会导致模型参数量和计算量显著增加^[19]。因此, 本研究设计了 DFE 模块, 通过有效融合特征图中的语义信息, 进一步动态捕捉融合后高分辨率的低层特征信息, 增强模型对小尺寸目标的敏感度。相比传统的添加 P2 检测层的方法, DFE 显著减少了计算成本, 其结构如图 8 所示。

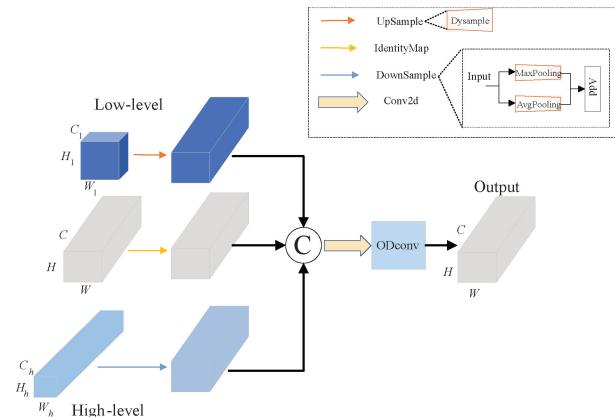


图 8 DFE 结构

Fig. 8 DFE structure

DFE 模块主要由高层、中层和低层 3 个特征图分支组成。首先, 在高层特征图中全局平均池化和全局最大池化被用来计算每个通道的平均值和最大值。最大池化的主要目的是从每个通道中提取最相关的数据, 而平均池化是为了均匀地从特征图中获取所有数据, 以最小化过度损失。高层特征图通过将这两种池化方法结合使用进行下采样, 从每个通道中提取最代表性的信息, 同时最小化信息损失。其次, 底层特征图使用 DySample 模块进行上采样, 有助于保持低分辨率图像的丰富局部特征, 并防止小目标特征信息的损失。然后, 将底层、中层和高层特征图调整为相同尺寸后进行拼接, 并通过卷积层处理。最后, 经过 ODConv^[20] 进行动态高效的特征提取, 进一步提升模型的特征表达能力。ODConv 在卷积过程中根据输入数据的特征动态地调整卷积核的形状和大小, 以适应不同的输入数据, 它分别引入空间维度、输入通道维度、输出通道维度、核空间维度这四种注意力机制, 使模型能在不同维度上全面提取融合后小尺寸违禁品的关键特征与全局信息。ODConv 的多维注意力机制如图 9 所示。

2.3 损失函数改进

作为目标检测定位中的重要组成, 边界框回归损失直

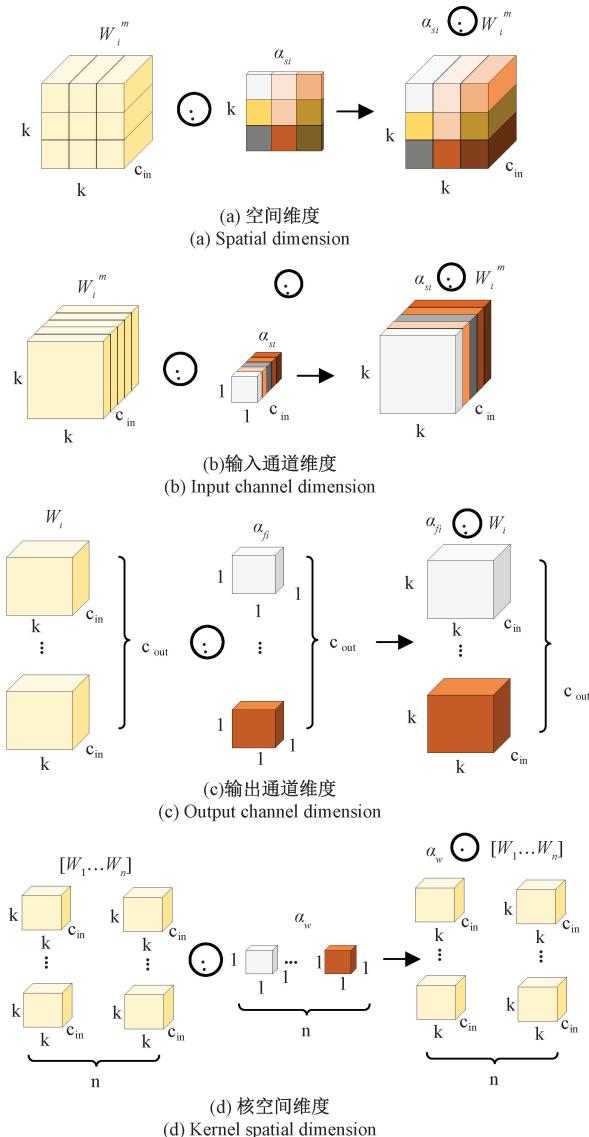


图 9 ODConv 多维注意力机制

Fig. 9 ODConv multi-dimensional attention mechanism

接影响模型对目标位置的预测精度。YOLOv8 中使用的 CIoU 边界框回归损失函数考虑了框中心之间的距离和长宽比的差异, 提高了收敛速度和整体性能, 然而当处理尺度多变的违禁品时, CIoU 忽略了具有相同长宽比但尺寸不同的框之间的大小差异, 导致回归精度的下降。因此, 本文将 Shape-IoU 作为新的损失函数, 该方法不仅考虑了真实框和预测框的距离、形状和角度的影响, 还结合了边界框自身的形状和尺寸来计算损失, 提升模型对违禁品尺寸变化的适应性, 从而使边界框回归更准确。Shape-IoU 如图 10 所示。

假设 $b_{gt}(x_c^{gt}, y_c^{gt})$ 是 GT 框的中心点, $b(x_c, y_c)$ 是预测框的中心点。Shape-IoU 的公式为:

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (9)$$

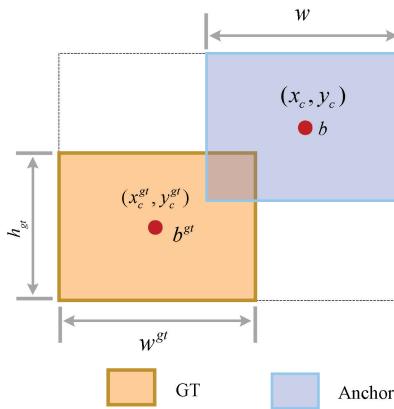


图 10 Shape-IoU

Fig. 10 Shape-IoU

$$ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (10)$$

$$hh = \frac{2 \times (h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (11)$$

$$d^{shape} = hh \times (x_c - x_c^{gt})^2 / c^2 + ww \times (y_c - y_c^{gt})^2 / c^2 \quad (12)$$

$$\Omega^{shape} = \sum_{t=w,h} (1 - e^{-wt})^4 \quad (13)$$

$$w_w = hh \times \frac{|w - w^{gt}|}{\max(w, w^{gt})} \quad (14)$$

$$w_h = hh \times \frac{|w - w^{gt}|}{\max(w, w^{gt})} \quad (15)$$

$$L_{Shape-IoU} = 1 - IoU + d^{shape} + 0.5\Omega^{shape} \quad (16)$$

式中: $scale$ 为缩放因子, ww 和 hh 分别表示水平方向和垂直方向上的权重系数。

3 实验设计和结果分析

3.1 数据集和环境

本研究采用 SIXray 数据集^[21]来验证改进模型的有效性。SIXray 数据集由中国科学技术大学(USTC)发布,是专门用于违禁品检测的 X 射线图像数据集。该数据集具有复杂的图像背景,违禁品往往被衣物、包袋等其他物品遮挡,增加了检测的难度,更贴近实际安检场景。SIXray 数据集包含超过 100 万张图像,其中仅有 8 929 张包含违禁品,充分反映了真实安检中违禁品样本稀少的特点。该数据集中的违禁品类别涵盖枪支、刀具、扳手、钳子和剪刀等 5 种常见的危险物品,部分样本如图 11 所示。SIXray 数据集按照 8:1:1 的比例随机划分为训练集、验证集和测试集。

本研究的实验平台运行在 Ubuntu 20.04 系统上,主要编程语言为 Python 3.8 版本。实验所使用的硬件设备包括 Intel Core i7-8700K@3.7 GHz 六核 CPU、16 GB 运行内存,以及 NVIDIA GeForce RTX 2080 8 GB 显存 GPU。实验中的 Batch size 设置为 8,优化器选择了 SGD,

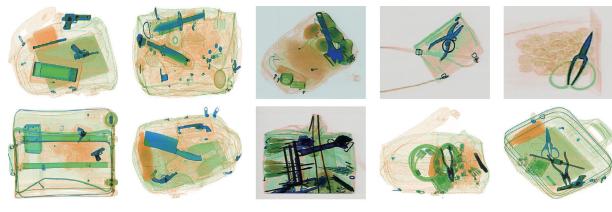


图 11 SIXray 数据样本示例

Fig. 11 Example samples from the SIXray dataset

初始学习率为 0.01,权重衰减系数为 0.0005,Epoch 设置为 300。为确保实验的公正性,所有算法在实验过程中均采用相同的实验环境、数据集和数据增强技术。

3.2 评价指标

采用单个目标的平均检测精度(AP)、IoU 阈值为 0.5 时所有目标的平均检测精度均值(mAP50)、参数量(Parameters)、计算量(FLOPs)、模型体积(Model Size)、每秒检测帧数(FPS)作为评价指标,以综合评估模型性能。AP 与 mAP 的公式为:

$$AP = \frac{TP}{TP + FP} \quad (17)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (18)$$

其中, TP 代表实际为正样本且预测为正类的数量; FP 表示实际为负样本但预测值为正类的数量; AP_i 为第 i 类违禁品的 AP 值; N 为违禁品类别个数。

3.3 对比实验

1) EFE_C2f 模块对比实验

为验证 EFE_C2f 模块设计的有效性,本文采用了普通卷积(Conv)、幻影卷积(GhostConv)、深度可分离卷积(DSConv)部分卷积(PConv)4 种卷积方式,并结合 Sobel 算子进行了对比实验,结果如表 1 所示。

表 1 卷积模块性能对比

Table 1 Performance comparison of convolution modules

卷积类型	Params/M	GFLOPs	mAP50/%
Conv	10.21	25.5	93.47
GhostConv	9.83	24.6	92.62
DSConv	9.52	23.8	93.65
PConv	9.48	23.7	93.84

实验结果显示,当 Sobel 算子与 4 种卷积方式结合时,普通卷积(Conv)的参数量和计算量最大,但其检测效果并不理想,mAP 值仅比最低的幻影卷积(GhostConv)高出 0.55%。在使用幻影卷积时,尽管参数量和计算量略低于普通卷积,但 mAP 值最低,主要原因是在处理复杂背景时,GhostConv 仅对特征图进行简单的线性变换,无法充分

提取违禁品的关键特征。采用深度可分离卷积(DSConv)时,与普通卷积相比,参数量和计算量分别减少了 0.69 M 和 1.7 GFLOPs,且 mAP 值比幻影卷积高出 1.03%。使用部分卷积(PConv)时,参数量和计算量相较于普通卷积减少最多,分别降低了 0.73 M 和 1.8 GFLOPs,同时 mAP 值比普通卷积、幻影卷积和深度可分离卷积分别高出 0.37%、1.22% 和 0.19%。综上所述,与其他 3 种卷积方式相比,使用 PConv 卷积使模型的综合检测效果达到了最优。

2) 颈部网络对比实验

为了评估 EM-FPN 的性能,选择了当前五种主流的特征融合网络结构(PAN-FPN、BIFPN、GDFPN、ASF、HSFPN)作为 YOLOv8s 的颈部网络,与 EM-FPN 进行对比,观察不同颈部网络结构对模型性能的影响。通过分析表 2 可知,相较于 YOLOv8 的原始颈部网络(PAN-FPN),EM-FPN 在 mAP 这一关键指标上有显著提升,提升幅度达 1.38%,明显优于其他几种颈部网络结构。同时,在计算复杂度方面,得益于动态上采样模块(Dysample)和高效网络融合结构的设计,EM-FPN 显著减少了计算资源的消耗,其参数量和计算量均为最低,且模型体积最小,较好地满足了模型轻量化的要求。在检测速度方面,尽管 EM-FPN 的 FPS 略低于原始颈部网络,但仍能满足实时检测的需求。

表 2 不同颈部网络性能对比

Table 2 Performance comparison of different neck networks

网络	Params/M	GFLOPs	mAP50/%	FPS	Size/MB
PAN-FPN	11.13	28.4	92.07	118	22.5
BIFPN	7.37	25	93.14	107	15.0
GDFPN	13.61	29.9	92.36	96	27.7
ASF	11.29	30.1	92.81	106	22.8
HSFPN	7.14	23.9	92.76	114	14.5
EM-FPN	6.97	22.6	93.45	110	14.1

为进一步验证 EM-FPN 在检测小尺寸违禁品目标方面的有效性,本研究将 EM-FPN 与传统的小目标检测增强方法(额外添加一个 P2 检测头)进行对比,实验结果如表 3 所示。基准模型 YOLOv8 在加入小目标 P2 检测头后,mAP 提升了 1.19%,这表明原有的网络结构在检测小尺寸违禁品目标方面存在不足,有较大的改进空间。尽管加入 P2 检测头的方法提高了模型的检测精度,但同时也显著增加了计算资源的消耗,导致参数量、计算量和模型体积的增大,以及检测速度的明显下降。相比之下,EM-FPN 的参数量、计算量和模型体积分别仅为 P2 检测头的 50.25%、69.32% 和 51.6%,且 EM-FPN 的 mAP 值表现更为优越,比 P2 检测头提升了 0.19%。这些结果验证了 EM-FPN 在检测小尺寸违禁品目标方面的有效性。

表 3 小尺寸目标检测方法对比

Table 3 Comparison of small object detection methods

方法	Params/M	GFLOPs	mAP50/%	FPS	Size/MB
—	11.13	28.4	92.07	118	22.5
P2	13.87	32.6	93.26	99	27.3
EM-FPN	6.97	22.6	93.45	110	14.1

3) 损失函数对比实验

为了验证 Shape-IoU 的有效性,本研究将其与当前流行的回归损失函数(C-IoU、D-IoU、S-IoU、E-IoU)进行了对比实验,结果如表 4 所示。相比 Shape-IoU,其他损失函数的精度均较低,验证了 Shape-IoU 的优越性。

表 4 不同损失函数对比

Table 4 Comparison of different loss functions

IoU	mAP50/%
C-IoU	92.07
D-IoU	91.76
S-IoU	92.37
E-IoU	92.52
Shape-IoU	92.74

4) 与主流模型对比实验

为了评估不同算法在各类违禁品检测中的性能,本文选择了几种当前较为先进的算法与本文提出的 LEM-YOLO 算法在 SIXray 数据集上进行对比实验。从检测精度和轻量化两个方面进行对比的实验结果如表 5、6 所示。

从表 5 可以看出,本研究提出的 LEM-YOLO 模型的检测精度(mAP50)为 94.63%。与 YOLO 系列中最新的 YOLOv10 m 算法相比,mAP 提高了 1.15%,且显著高于其他主流算法。相较于基准模型 YOLOv8s,LEM-YOLO 在五种违禁品类别的检测精度(AP)上均有所提升,这表明该算法在提升违禁品检测精度方面表现出色,能够满足高精度检测的需求。

观察表 6 可知,在轻量化方面,LEM-YOLO 相较于其他算法在参数量、计算量和模型体积上均为最低,充分满足了轻量化的要求。虽然由于部分复杂设计导致推理时间延长,LEM-YOLO 的 FPS 略低于基准模型,但其检测速度仍然保持在每秒 105 帧,满足了实时检测的需求。

图 12 进一步直观分析了 LEM-YOLO 的检测性能,圆圈大小表示 Params 开销。与其他算法相比,LEM-YOLO 在精度和轻量化方面均展现出显著优势,适用于轻量级实时检测违禁品目标的场景。

3.4 消融实验

为深入分析各改进点对模型性能的影响,本研究在基

表 5 不同模型检测精度对比

Table 5 Comparison of detection accuracy across different models

模型	AP					mAP50
	Gun	Knife	Wrench	Pliers	Scissors	
YOLOv8s	98.04	90.57	91.87	91.56	88.30	92.07
YOLOv9m ^[22]	98.21	90.34	91.92	95.84	89.47	93.16
YOLOv10m ^[23]	98.30	90.59	92.67	94.48	91.34	93.48
RTDETR-L ^[24]	98.98	91.25	93.37	95.27	91.33	94.04
YOLOv8-af ^[11]	97.80	91.70	91.60	91.60	90.80	92.70
YOLOv8-SP ^[12]	98.70	91.21	91.30	94.31	94.85	94.27
YOLO-ray ^[25]	98.60	89.70	89.90	95.00	92.60	93.20
YOLOv5-d ^[26]	99.10	86.80	90.80	96.30	91.00	92.80
LEM-YOLO(Ours)	98.34	91.59	94.23	95.75	93.24	94.63

表 6 不同模型轻量化对比

Table 6 Comparison of lightweight models

模型	Params/M	GFLOPs	FPS	Size/MB
YOLOv8s	11.13	28.4	118	22.5
YOLOv9m	20.15	77	82	41.1
YOLOv10m	15.31	58.9	95	33.4
RTDETR-L	28.45	100.6	54	59.0
YOLOv8-af ^[11]	—	—	—	—
YOLOv8-SP ^[24]	11.93	29.75	115	—
YOLO-ray ^[25]	6.66	15.3	85	—
YOLOv5-d ^[26]	12.20	—	87	—
LEM-YOLO	6.34	20.9	105	11.1

准模型 A 的基础上逐步引入各改进模块, 进行了消融实验, 结果如表 7 所示。首先, 将原本的 C2f 模块替换为 EFE_C2f 模块后, 模型 B 的参数量和计算量均有所减少, 同时 mAP 提升了 1.77%。EFE_C2f 模块的引入降低了计算资源消耗, 增强了模型对边缘特征的提取能力。其次, 在模型 B 的基础上进一步引入 EM-FPN 网络结构, 得到模型 C。与基准模型 A 相比, 模型 C 的参数量、计算量以及模型体积分别减少了 43.04%、26.41% 和 50.67%。此外, 相较于模型 B, 模型 C 的 mAP 提高了 0.62%, 表明

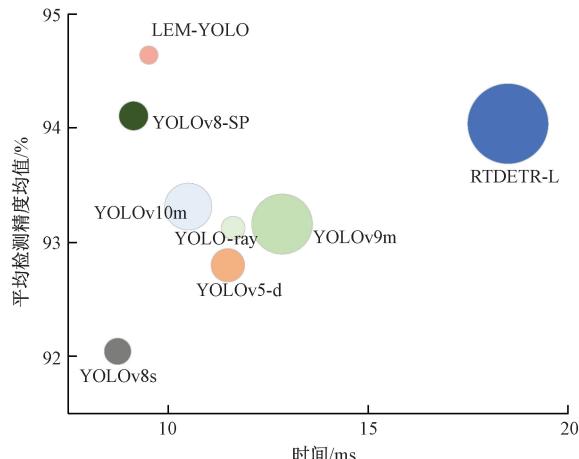


图 12 不同算法的性能、推理速度和 Params 比较

Fig. 12 Comparison of performance, inference speed, and Params of different algorithms

EM-FPN 增强了模型的多尺度特征提取能力。最后, 将 Shape-IoU 作为新的损失函数, 模型的 mAP 达到 94.63%, 比基准模型 A 提升了 2.56%, 且模型体积与计算资源消耗均降至最低。消融实验的结果验证了所提算法的有效性, LEM-YOLO 在违禁品检测中表现出优越的检测性能。

表 7 消融实验

Table 7 Comparative experiments

模型	EFE_C2f	EM-FPN	Shape-IoU	Params/M	GFLOPs	mAP50/%	Size/MB
A				11.13	28.4	92.07	22.5
B	V			9.47	23.5	93.84	19.2
C	V	V		6.34	20.9	94.46	11.1
D	V	V	V	6.34	20.9	94.63	11.1

3.5 可视化分析

为了直观分析模型改进后的效果, 选择基准模型与改

进模型在违禁品检测中进行可视化对比, 测试结果如图 13 所示。其中图 13(a)是单样本图像, 图 13(b)是多样本图

像,图 13(c)是被遮挡严重的样本图像,图 13(d)是尺度变化大的样本图像。

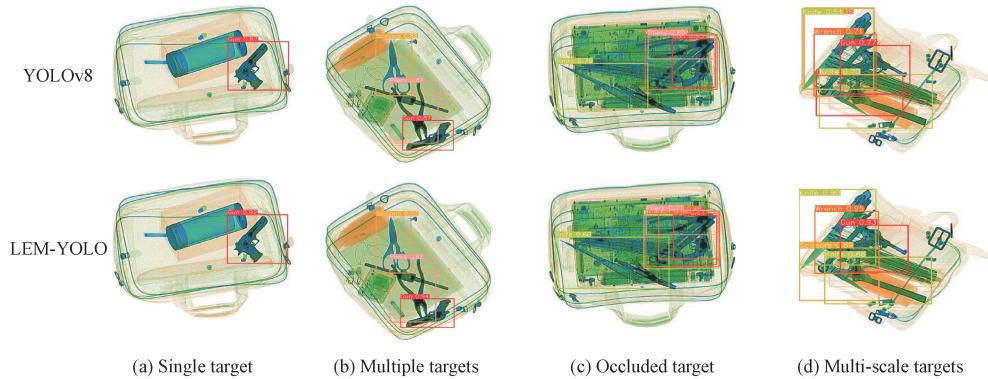


图 13 改进前后模型检测结果对比

Fig. 13 Comparison of detection results before and after model improvements

从图 13 可以看出,在单样本和多样本图像的检测中,相较于 YOLOv8,LEM-YOLO 的置信度分数更高,检测效果更优。对于遮挡严重的样本图像和尺度变化较大的样本图像,YOLOv8 的检测效果并不理想,分别出现漏检一把刀和误检一个扳手并漏检一个剪刀的情况,而 LEM-YOLO 则全部正确检出。这进一步验证了本文所提方法能够有效提高模型对 X 射线安检图像的检测效果。

4 结 论

所提算法 LEM-YOLO 针对 X 射线违禁品检测中存在背景信息复杂、尺度变化剧烈、小目标不易检测、部署困难等问题,在 YOLOv8s 的基础上进行改进。在骨干网络中设计轻量化边缘特征增强模块 LEFE,强化了边缘信息的表征,从而提高模型的检测精度;在颈部网络中提出了高效多级特征融合金字塔网络 EM-FPN,实现上下文信息的高效融合,在轻量化的同时提升模型的多尺度感知能力;采用 Shape-IoU 作为回归损失函数,削弱尺寸变化带来的影响,进一步提高了模型的定位效果。实验结果表明,与原算法相比,LEM-YOLO 的 mAP 提高了 2.56%,模型体积减少了 50.67%,且在检测难度较大的多种违禁品时表现更佳。与其它主流算法相比,LEM-YOLO 具有更高的检测精度和更小的模型体积,便于部署在安检设备中,为智能化安检的进步做出了积极贡献。由于 LEM-YOLO 算法的帧率有所下降,后续工作将针对这一问题进行优化。

参考文献

- [1] 梁添汾,张南峰,张艳喜,等.违禁品 X 光图像检测技术应用研究进展综述[J].计算机工程与应用,2021,57(16):74-82.
LIANG T F, ZHANG N F, ZHANG Y X, et al. Summary of research progress on application of prohibited item detection in X-Ray images [J]. Computer Engineering and Applications, 2021,

- 57(16): 74-82.
[2] WANG AO, CHEN H, LIN Z J, et al. Repvit: Revisiting mobile cnn from vit perspective[C]. IEEE/ CVF Conference on Computer Vision and Pattern Recognition, 2024: 15909-15920.
[3] 董乙杉,郭靖圆,李明泽,等.基于反向瓶颈和 LCBAM 设计的 X 光违禁品检测[J].计算机科学与探索,2024,18(5):1259-1270.
DONG Y SH, GUO J Y, LI M Z, et al. X-ray prohibited items detection based on inverted bottleneck and light convolution block attention module [J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(5): 1259-1270.
[4] 张良,薛志诚.基于自适应多尺度特征融合的 X 光违禁品检测[J].信号处理,2024,40(4):789-800.
ZHANG L, XUE ZH CH. X-Ray prohibited-item detection based on adaptive multi-scale feature fusion[J]. Journal of Signal Processing, 2024, 40(4): 789-800.
[5] LIU D M, LIU J CH, YUAN P X, et al. Lightweight prohibited item detection method based on YOLOv4 for X-RAY security inspection[J]. Applied Optics, 2022, 61(28): 8454-8461.
[6] 李松,亚森江·木沙.改进 YOLOv7 的 X 射线图像违禁品实时检测[J].计算机工程与应用,2023,59(12): 193-200.
LI S, YASENJIANG M SH. Improved YOLOv7 X-Ray image real-time detection of prohibited items[J]. Computer Engineering and Applications, 2023, 59(12): 193-200.
[7] 曹洋,张莉,孟俊熙,等.针对 X 光安检场景的多目标违禁品识别算法[J].激光与光电子学进展,2022,59(10):324-332.
CAO Y, ZHANG L, MENG J X, et al. Multi-target prohibited item recognition algorithm for X-Ray

- security scene[J]. *Laser & Optoelectronics Progress*, 2022, 59(10): 324-332.
- [8] 穆思奇, 林进健, 汪海泉, 等. 基于改进 YOLOv4 的 X 射线图像违禁品检测算法[J]. *兵工学报*, 2021, 42(12): 2675-2683.
- MU S Q, LIN J J, WANG H Q, et al. An algorithm for detection of prohibited items in X-ray images based on improved YOLOv4 [J]. *Acta Armamentarii*, 2021, 42(12): 2675-2683.
- [9] SIMA H, CHEN B L, TANG CH SH, et al. Multi-scale feature attention-detection transformer: Multi-scale feature attention for security check object detection[J]. *IET Computer Vision*, 2024, 18(5): 613-625.
- [10] 粟兴旺, 王晓明, 黄金波, 等. 基于可变形卷积与注意力机制的 X 光安检违禁品检测[J]. *电子测量技术*, 2023, 46(10): 98-108.
- SU X W, WANG X M, HUANG J B, et al. X-ray security contraband detection based on deformable convolution and attention mechanism [J]. *Electronic Measurement Technology*, 2023, 46(10): 98-108.
- [11] 王海群, 魏培旭. 基于改进 YOLOv8 的 X 光图像违禁品检测[J]. *无线电工程*, 2024, 54(10): 2288-2295.
- WANG H Q, WEI P X. X-Ray image contraband detection based on improved YOLOv8 [J]. *Radio Engineering*, 2024, 54(10): 2288-2295.
- [12] 颜志明, 李新伟, 杨艺. 基于改进 YOLOv8s 的 X 射线图像违禁品检测算法[J/OL]. *计算机工程与应用*, 1-9 [2024-07-10]. <http://kns.cnki.net/kcms/detail/11.2127.tp.20240705.1641.010.html>.
- YAN ZH M, LI X W, YANG Y. X-ray image contraband detection based on improved YOLOv8s [J/OL]. *Computer Engineering and Applications*, 1-9 [2024-07-10]. <http://kns.cnki.net/kcms/detail/11.2127.tp.20240705.1641.010.html>.
- [13] SONG B, LI R Q, PAN X G, et al. Improved YOLOv5 detection algorithm of contraband in X-ray security inspection image[C]. 2022 5th International Conference on Pattern Recognition and Artificial Intelligence(PRAI). IEEE, 2022: 169-174.
- [14] 左景, 石洋宇, 卢树华. 基于轻量化卷积和 SCAM 改进的 X 光违禁品检测[J/OL]. *计算机科学与探索*, 1-17 [2024-07-14]. <http://kns.cnki.net/kcms/detail/11.5602.TP.20240711.1543.007.html>.
- ZUO J, SHI Y Y, LU SH H. Improved X-ray prohibited items detection method based on lightweight convolution blocks and SCAM attention mechanism[J/OL]. *Journal of Frontiers of Computer Science and Technology*, 1-17 [2024-07-14]. <http://kns.cnki.net/kcms/detail/11.5602.TP.20240711.1543.007.html>.
- [15] CHEN J R, KAO SH H, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [16] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample[C]. IEEE/CVF International Conference on Computer Vision, 2023: 6027-6037.
- [17] ZHANG H, ZHANG S, Shape-IoU: More accurate metric considering bounding box shape and scale[J]. *ArXiv preprint arXiv:2312.17663*, 2023.
- [18] CHEN Y F, ZHANG CH Y, CHEN B, et al. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases[J]. *Computers in Biology and Medicine*, 2024, 170: 107917.
- [19] 曾泓翔, 文志诚. 基于改进 YOLOv5 的 X 光违禁品检测算法[J]. *计算机工程与应用*, 2024, 60(16): 217-227.
- ZENG H X, WEN ZH CH. X-Ray contraband detection algorithm based on improved YOLOv5[J]. *Computer Engineering and Applications*, 2024, 60(16): 217-227.
- [20] LI CH, ZHOU AO J, YAO AN B, Omni-dimensional dynamic convolution[J]. *ArXiv preprint arXiv:2209.07947*, 2022.
- [21] MIAO C J, XIE L X, W F, et al. Sixray: A large-scale security inspection X-RAY benchmark for prohibited item discovery in overlapping images[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 2119-2128.
- [22] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer vision, Cham: Springer Nature Switzerland, 2024: 1-21.
- [23] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. *ArXiv preprint arXiv:2405.14458*, 2024.
- [24] ZHAO Y AN, LYU W Y, XU SH L, et al. Detrs beat YOLOs on real-time object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [25] 马昌嵩, 裴晓芳, 周磊, 等. 反向加权融合多尺度特征的 X 射线图像违禁品检测[J]. *国外电子测量技术*,

2024,43(4):170-180.

MA CH S, PEI X F, ZHOU L, et al. Reverse weighted fusion of multi-scale features for prohibited object detection in X-Ray images [J]. Foreign Electronic Measurement Technology, 2024,43(4):170-18.

- [26] 孙嘉傲, 董乙杉, 郭靖圆, 等. 自适应与多尺度特征融合的 X 光违禁品检测 [J]. 计算机工程与应用, 2024, 60(2): 96-102.

SUN J AO, DONG Y SH, GUO J Y, et al. Detection

of X-Ray contraband by adaptive and multi-scale feature fusion [J]. Computer Engineering and Applications, 2024, 60(2): 96-102.

作者简介

赵小涛,硕士研究生,主要研究方向为图像处理。

E-mail:1776190785@qq.com

李新伟(通信作者),博士,副教授,主要研究方向为机器学习和计算机视觉。

E-mail:lixinwei@hpu.edu.cn