http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0713

融合注意力机制的飞机蒙皮损伤 LHN-YOLO 小目标检测算法

李耀华*，陈杰
(中国民航大学 交通科学与工程学院，天津 300300)

摘　　要：针对飞机蒙皮损伤小目标检测精度低、参数量大等问题，以飞机蒙皮铆钉损伤为研究对象，基于 YOLOv8n 算法提出了一种融合注意力机制的 LHN-YOLO 算法。为增强模型对小目标的检测能力，在骨干部分引入优化后的高效多尺度注意力 (EMA) 机制，通过轻量化卷积 LightConv 对 EMA 机制进行优化，形成 Light-EMA 模块，提高检测精度，降低计算复杂度；同时，引入 Haar 小波降采样模块和归一化高斯 Wasserstein 距离损失函数替换原模型的下采样操作和 CIoU 损失函数以提升预测性能。实验结果表明：LHN-YOLO 算法的准确率、召回率、平均精度均值分别为 92.7%、87.0%、91.7%，相较于改进前的原模型提升了 2.2%、3.5%、2.8%，参数量减少了 10%，有效提升了飞机蒙皮损伤小目标检测的效果。

关　键　词：YOLOv8n；注意力机制；小目标检测；铆钉损伤；蒙皮损伤检测
中图分类号：V267+.2; TP391.4
文献标志码：A　　**文章编号**：1001-5965(2026)11-3735-10

随着民航业的快速发展，飞机的设计和维护技术也有了很大进步，铆钉作为飞机结构中广泛使用的连接方式，若其松动脱落或腐蚀且不及时修理，将严重威胁飞机的飞行安全。飞机蒙皮采用了大量的铆钉进行连接，传统的飞机蒙皮缺陷检测主要依靠人工来进行，要求技术人员通过目视或光学仪器辅助的方式来识别飞机蒙皮上存在的缺陷。这种检测方法不仅存在耗时长、技术人员高空作业的风险，而且检测结果严重依赖技术人员的专业知识和主观判断[1-2]。因此，研究采用快速而准确的目标检测算法对飞机蒙皮进行小目标损伤检测意义重大。

目标检测技术体系当前主要包含两大研究分支：①基于传统视觉特征的检测算法，这类算法通过提取目标区域的边缘响应、亮度分布、直方图统计及光谱特征等人工设计特征完成缺陷识别与区域分割[3]，但在环境适应性、实时性及定位精度等关键性能上仍存在明显瓶颈；②基于深度学习的智能检测算法，该方向可进一步划分为单阶段与双阶段检测架构，单阶段架构以 SSD、YOLO 等端到端检测框架为代表，双阶段架构则以 Faster R-CNN 等网络为典型实现[4]。

相较于传统算法，基于深度学习的检测框架在系统稳定性、处理效率和识别精度等多个维度实现了显著改进。特别是在单阶段检测领域，YOLO 系列算法通过持续迭代优化，在检测实时性与准确度的平衡性方面表现突出。值得关注的是，该系列最新推出的 YOLOv8 开源版本(2023 年 1 月发布)，通过改进网络结构与训练策略，在保持毫秒级推理速度的同时，使平均精度均值(mean average precision, mAP)指标获得了显著提升。

原始 YOLOv8 模型用于飞机蒙皮损伤检测遇到的问题主要在于：小目标特征信息在提取过程中

收稿日期：2024-09-27；**录用日期**：2024-12-19；**网络出版时间**：2025-02-18 15:03
网络出版地址：link.cnki.net/urlid/11.2625.V.20250218.1111.001
基金项目：国家自然科学基金委员会-中国民航局民航联合研究基金重点支持项目 (U2033209)
*通信作者. E-mail：yh_li@cauc.edu.cn

引用格式：李耀华，陈杰. 融合注意力机制的飞机蒙皮损伤 LHN-YOLO 小目标检测算法 [J]. 北京航空航天大学学报，2026，52（11）：3735-3744. Li Y H，Chen J. Aircraft skin damage LHN-YOLO small target detection algorithm incorporating attention mechanisms[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（11）：3735-3744 （in Chinese）.

容易丢失[5-6], YOLOv8原模型适用于标准尺度的目标[7-8], 对于小目标检测效果不佳, 传统的损失函数对小目标对象偏差过于敏感, 不利于模型的收敛等[9]。针对以上问题, 现有研究人员提出的解决方法有: 对特征融合机制进行改进, 在网络中添加小目标检测层, 充分利用目标的浅层特征信息, 提高对小目标检测的准确性, 但因为增加了模型复杂度, 影响模型的计算效率[10]; 通过在模型的合适位置嵌入注意力机制模块, 增加对输入特征的有效信息利用率, 能够对特定尺寸的目标起到较好的检测效果[11-12]; 用轻量化的卷积替换颈部网络层和C2f模块的标准卷积, 大幅降低模型参数, 提升模型的计算效率, 但在面对大量的小目标对象时存在漏检现象[13]; 设计新的卷积块结构, 在减少特征图尺寸的同时捕获特征图的全局信息, 以取代池化和跨步卷积层, 从而提高小目标特征表示能力[14]; 考虑到检测对象的形状与分布, 提出更符合检测目标的损失函数和图像处理算法, 在特定形状和尺寸的对象上能够取得较好的检测效果[15]。

目前的目标检测算法在对飞机蒙皮损伤进行小目标检测时, 容易遇到检测精度不佳、模型收敛速度慢、模型参数量大的问题, 为解决这些问题, 本文引入了高效多尺度注意力(efficient multi-scale attention, EMA)模块; 为避免引入注意力机制模块带来的参数量增加, 将原模块中内部卷积核为1×1和3×3的卷积块替换为轻量化卷积LightConv, 形成Light-EMA模块, 减少参数量, 提升检测精度; 为防止特征信息丢失和更快地收敛模型, 引入Haar小波降采样(Haar wavelet downsample, HWD)模块[16]和归一化高斯Wasserstein距离(normalized Gaussian Wasserstein distance, NWD)[17]损失函数, 提升模型对小目标特征的提取能力和收敛速度。

1 飞机蒙皮损伤LHN-YOLO小目标检测算法

1.1 LHN-YOLO网络结构建立

YOLOv8n的总体结构包括输入、骨干网络、颈部网络、头部网络4部分。根据具体需要, 将模型按照网络模型的大小分为YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l和YOLOv8x等5种模型[18]。

LHN-YOLO是在YOLOv8n网络的基础上进行改进后得到的, 为增强网络对小目标的检测能力, 通过使用LightConv对EMA机制进行优化, 形成了Light-EMA模块, 将Light-EMA模块应用于骨干网络部分; 同时, 引入HWD模块和NWD损失函数替换原模型的下采样操作和CIoU损失函数, 分别用于骨干网络、颈部网络和头部网络。改进后的LHN-YOLO网络结构如图1所示。

LHN-YOLO网络在骨干网络部分的4个C2f模块后加入Light-EMA模块。骨干网络用于提取输入信息的特征, 通过卷积块、快速空间金字塔池化(spatial pyramid pooling-fast, SPPF)和C2f模块组合实现了不同尺度特征的高效融合, 加强了浅层与深层特征间的信息传递。Light-EMA模块通过使用LightConv优化了EMA机制, 使其能够在减少参数量和计算复杂度的基础上提升对小目标对象的检测精度, 将其置于骨干网络的C2f模块后有助于融合不同尺度的特征信息, 提高模型的检测精度。颈部网络部分借助路径聚合网络(path aggregation network, PAN)、特征金字塔网络(feature pyramid network, FPN)增强了局部与全局特征的融合, 丰富了目标特征信息。本文在颈部网络和骨干网络部分使用HWD模块替换了原模型的卷积块, 减少了下采样过程中小目标特征的丢失。在头部网络部分使用NWD损失函数替换原模型的CIoU损失函数, 提升模型对小目标的检测性能。

1.2 融合轻量化卷积的Light-EMA模块

由于数据集中的铆钉损伤绝大部分属于小目标对象, 在使用原模型进行预测时容易出现错检、漏检的情况, 本文引入了EMA机制, 提高模型对小目标对象的检测能力, 同时利用LightConv对注意力机制进行改进, 形成了Light-EMA机制, 在不增加参数量的基础上提高检测精度。

Ouyang等[19]提出的新型EMA模块主要通过3个关键机制提高特征提取的效率和性能: ①将输入特征图沿通道维度划分为多个子特征组, 以学习不同语义; ②采用并行的1×1和3×3分支, 同时捕获深层和浅层的特征依赖关系, 并利用3条并行路径提取注意力权重描述符; ③建立通道和空间位置之间的相互依赖关系, 提供跨空间信息聚合的方法, 融合1×1和3×3分支的输出以获得最终的特征表示。EMA模块通过这些设计有效捕获了多尺度的语义特征, 从而提高了模型的性能。但在模型中加入EMA模块会导致参数量增加, 其中的二维卷积Conv2d参数量较大、计算复杂度较高, 存在提升空间。

为解决该问题, 本文引入LightConv对EMA机制进行优化。LightConv是一种轻量化卷积操作, 用于减少模型的计算量和参数量[20]。

LightConv通过权重共享操作减少模块的参数量。在其使用的深度卷积中, 对通道的参数进行绑定, 使其相比于普通正则卷积, 参数量减少一半。例如, 对于特征维度d=1 024, 卷积核大小k=7, 正则

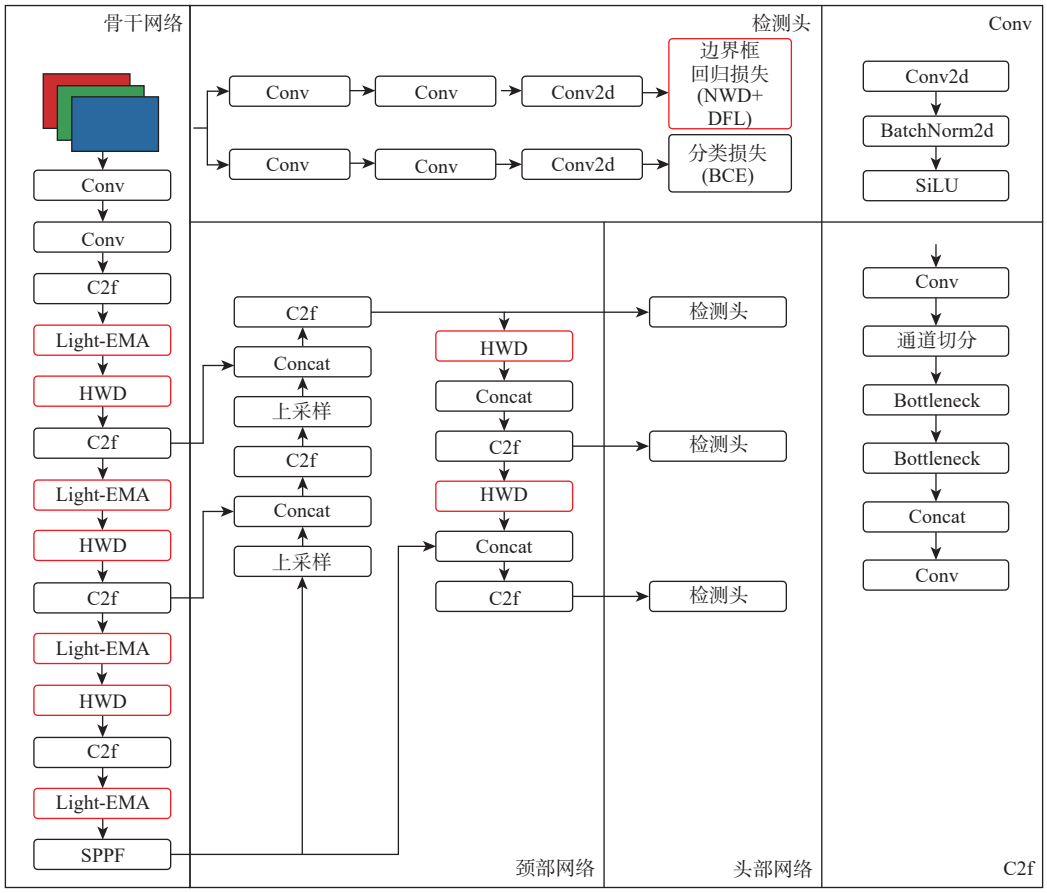


图 1　LHN-YOLO 网络结构

Fig. 1　LHN-YOLO network structure

卷积需要 7 340 032 权重 ($d^2\times k$), 深度可分离卷积需要 7 168 权重 ($d\times k$), 而在权重共享的情况下, 权重共享分组数 $H=16$, 只有 112 权重 ($H\times k$)。通过权重共享的方式, LightConv 实现了计算复杂度降低和浮点运算数减少。

LightConv 计算列表中, 第 i 个元素和输出通道 c 的计算过程如下:

$$\text{LightConv}\left(\boldsymbol{X},\boldsymbol{W}_{\left[\frac{cH}{d}\right],:},i,c\right)=\text{DepthwiseConv}\left(\boldsymbol{X},\text{Softmax}\left(\boldsymbol{W}_{\left[\frac{cH}{d}\right],:}\right),i,c\right)\tag{1}$$

深度卷积在每个通道上独立地执行卷积, 参数个数可以从 d^2k 减少到 dk。式(1)中: $\boldsymbol{X}$ 为输入序列特征张量; LightConv() 表示轻量化卷积操作; DepthwiseConv() 表示深度卷积操作; Softmax() 表示 Softmax 操作; k 为卷积核大小。对于元素 i 和输出维数 c, 权重为 $\boldsymbol{W}\in\mathbf{R}^{d\times k}$ 的深度卷积的输出定义为

$$\text{DepthwiseConv}(\boldsymbol{X},\boldsymbol{W}_{c,:},i,c)=\sum_{j=1}^{k}W_{c,j}X_{\left(i+j-\left\lfloor\frac{k+1}{2}\right\rfloor\right),c}\tag{2}$$

LightConv 使用 Softmax 操作在时间维度 k 上归一化权重 $\boldsymbol{W}\in\mathbf{R}^{H\times k}$:

$$\text{Softmax}(\boldsymbol{W})_{h,j}=\exp\left(W_{h,j}\right)\Big/\sum_{j'=1}^{k}\exp\left(W_{h,j'}\right)\tag{3}$$

LightConv 结构如图 2 所示。图中: LConv 表示轻量级卷积。

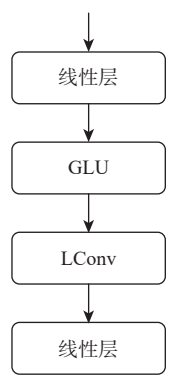


图 2　LightConv 结构

Fig. 2　LightConv structure

门控线性单元 (gated linear unit, GLU) 的计算公式如下:

$$U_l(\boldsymbol{X})=(\boldsymbol{X}*\boldsymbol{W}+\boldsymbol{b}_1)\otimes\sigma(\boldsymbol{X}*\boldsymbol{V}+\boldsymbol{b}_2)\tag{4}$$

式中: σ 表示 Sigmoid 函数; $\otimes$ 表示逐元素相乘; $\boldsymbol{b}_1$、$\boldsymbol{b}_2$ 为偏置项; $\boldsymbol{V}$ 为可学习权重。

综上所述，LightConv以较低的计算复杂度使得模型能够更快速地进行目标检测和推理，并且拥有不错的泛化性。观察到，在EMA机制中，1×1和3×3分支使用了卷积核分别为1×1和3×3的卷积，在网络中加入EMA模块后，参数量有所增加，这不利于提升模型的预测性能，并且会影响模型的运行速度，因此，本文使用LightConv模块对EMA模块中的Conv(1×1)和Conv(3×3)进行替换，减少了参数量、计算复杂度的同时提升了检测精度。

改进后的Light-EMA模块网络结构如图3所示。

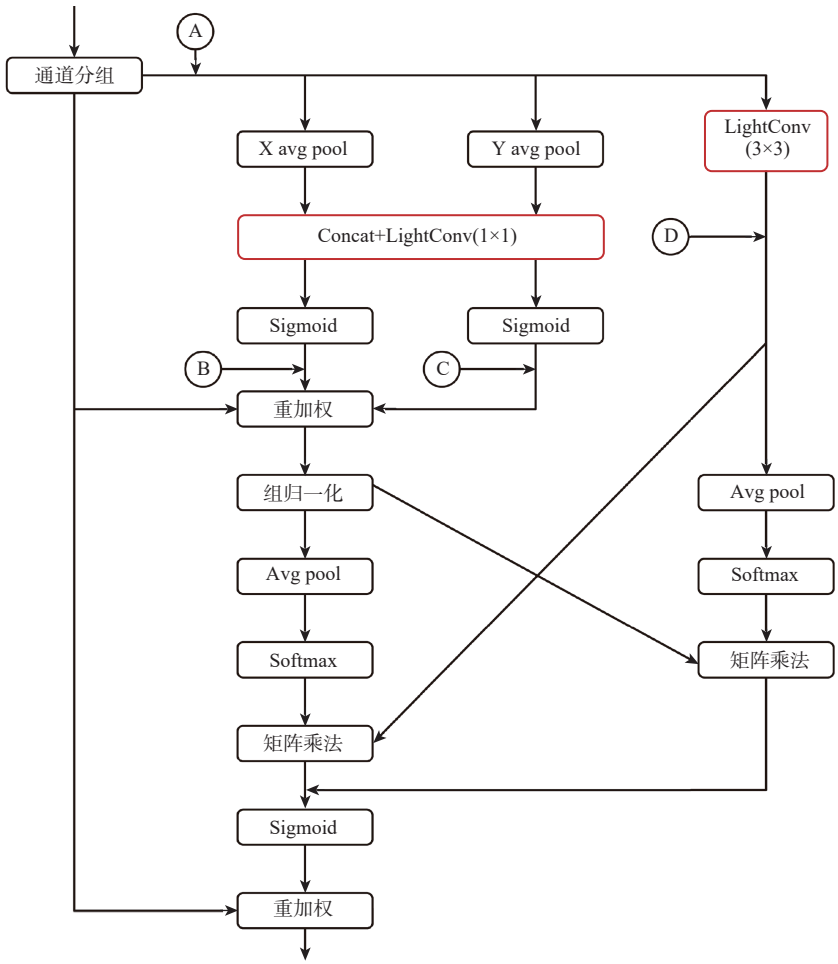


图3 Light-EMA结构

Fig. 3 Light-EMA structure

在Light-EMA运算过程中：

$$\boldsymbol{F}_{\mathrm{B}}=\boldsymbol{F}_{\mathrm{C}}=\sigma\left(f_{\mathrm{LightConv}}^{1\times1}\left(\mathrm{cat}\left(\mathrm{XAvg}(\boldsymbol{F}_{\mathrm{A}}),\mathrm{YAvg}(\boldsymbol{F}_{\mathrm{A}})\right)\right)\right) \tag{5}$$

$$\boldsymbol{F}_{\mathrm{D}}=f_{\mathrm{LightConv}}^{3\times3}(\boldsymbol{F}_{\mathrm{A}}) \tag{6}$$

式中：$\boldsymbol{F}_{\mathrm{A}}\in\mathbf{R}^{C'\times H'\times W'}$，$C'$、$H'$、$W'$分别为输入数据的通道数、高度、宽度；$f_{\mathrm{LightConv}}^{1\times1}$表示卷积核为1×1的LightConv卷积运算；$f_{\mathrm{LightConv}}^{3\times3}$表示卷积核为3×3的LightConv卷积运算；cat表示拼接操作；$\mathrm{XAvg}(\boldsymbol{F}_{\mathrm{A}})$表示水平维度平均池化，$\mathrm{YAvg}(\boldsymbol{F}_{\mathrm{A}})$表示垂直维度平均池化，$\mathrm{XAvg}(\boldsymbol{F}_{\mathrm{A}})\in\mathbf{R}^{C'\times H'\times1}$，$\mathrm{YAvg}(\boldsymbol{F}_{\mathrm{A}})\in\mathbf{R}^{C'\times1\times W'}$。

1.3 HWD模块

在提取铆钉等小目标对象的特征信息时，往往需要进行下采样操作，但原模型中的下采样操作在缩放图像的同时容易丢失小目标对象的特征信息，导致模型对小目标对象预测性能不高，为此本文引入HWD模块。

Xu等[16]提出基于HWD模型。HWD模块在深度神经网络的编码器部分能够尽可能地保留图像的信息，更容易提取出更具判别性的特征，从而有助于提高分割性能。因此，针对传统下采样方法直接降低特征图的空间分辨率，容易造成信息丢失的情况，本文引入HWD模块替换骨干网络层和颈部网络层中的部分卷积进行下采样操作。HWD模块包括无损特征编码块和特征学习块。首先，在不损失信息的前提下，利用无损特征编码块中的Haar小波变换先进行近似信息和高频信息提取，再进行下采样操作；然后，在特征学习块（由标准卷积层、批量归一化和ReLU激活层组成）中使用卷积操作进行特征学习，以过滤冗余信息。

一维 Haar 变换的小波基函数ϕ_1和尺度函数ψ_1可以定义如下:

$$\begin{cases}\phi_1 = \dfrac{1}{\sqrt{2}}\phi_{1,0}(x)+\dfrac{1}{\sqrt{2}}\phi_{1,1}(x)\\ \psi_1 = \dfrac{1}{\sqrt{2}}\phi_{1,0}(x)-\dfrac{1}{\sqrt{2}}\phi_{1,1}(x)\end{cases} \tag{7}$$

式中: $\phi_{j,k}(x)$定义为

$$\phi_{j,k}(x) = \sqrt{2^j}\phi(2^j x-k) \tag{8}$$

其中: $k=0,1,2,\cdots,2^j-1$, j和k为基函数的阶数。

通过以上操作, HWD 模块能在有效减少计算开销的前提下, 提升训练的速度和精度。HWD 模块的结构示意图如图 4 所示。

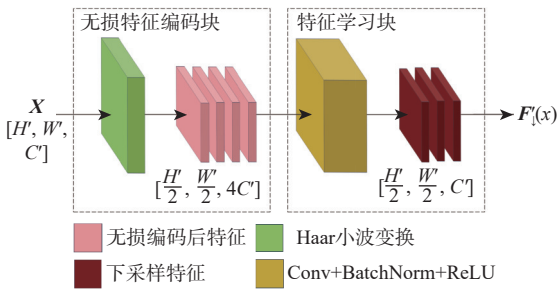


图 4　HWD 模块结构示意图

Fig. 4　HWD module structure

1.4　NWD 损失函数

在对飞机蒙皮铆钉损伤进行预测时, 由于数据集中铆钉的形状多为正圆形或椭圆形, 有效信息集中于预测框中心部位, CIoU 损失函数在对这类目标进行评估时, 往往将整个矩形预测框内的像素都作为目标对象, 对于不同尺度的对象敏感度差异较大, 尤其在进行小目标检测时, CIoU 较大的位置偏差敏感性不利于模型的收敛。

为解决该问题, 本文引入了 NWD 损失函数改进原模型的 CIoU 损失函数。Xu 等[17] 提出的 NWD 损失函数基于 Wasserstein 距离设计了一个更好的小目标度量, 针对小目标检测任务中普遍存在的几何特征适配问题, 实际场景中目标物体普遍存在非规则的几何形态特征, 导致矩形检测框难以避免地包含非目标区域像素。通过系统性分析可知, 目标有效像素与背景干扰像素在检测框内呈现显著空间分布规律: 前者主要集聚于检测框中心区域, 后者则密集分布于边界过渡区域。为精确量化检测框内各像素的空间贡献度差异, 可采用基于概率密度函数的建模方法, 将检测框表征为二维高斯分布模型。在该模型框架下, 中心像素对应概率密度峰值, 其响应强度随空间位置向边界移动呈现径向衰减特性, 这种机制可有效反映目标像素的空间梯度变化规律。

对于转化为二维高斯分布的边界框$N_a=(c_{x_a},c_{y_a},w_a,h_a)$和$N_b=(c_{x_b},c_{y_b},w_b,h_b)$之间的二阶 Wasserstein 距离, 可以简化为

$$W_2^2(N_a,N_b) = \left\| \left(\left[c_{x_a},c_{y_a},\frac{w_a}{2},\frac{h_a}{2}\right]^{\mathrm{T}}, \left[c_{x_b},c_{y_b},\frac{w_b}{2},\frac{h_b}{2}\right]^{\mathrm{T}} \right) \right\|_2^2 \tag{9}$$

式中: c_{x_a}、c_{y_a}、c_{x_b}、c_{y_b}分别表示对应边界框的中心点的横纵坐标; w_a、h_a、w_b、h_b分别表示边界框N_a、N_b的宽度和高; $W_2^2(N_a,N_b)$表示取值范围为$[0,+\infty)$且不能直接作为分配标签的相似性度量。为得到与 IoU(使其取值在 0～1 之间) 相似的取值范围, 并使度量值平滑到位置偏差, NWD 选择了一个指数非线性变换函数, 将高斯 Wasserstein 距离重新映射到另一个空间, 从而将其取值范围归一化得到了归一化高斯 Wasserstein 距离:

$$\mathrm{NWD}(N_a,N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a,N_b)}}{C'}\right) \tag{10}$$

2　实验验证分析

2.1　数据集处理

数据集来自中国民航大学机坪拍摄和某航空公司提供的飞机蒙皮图像, 去除因模糊、拍摄角度不佳等而不能使用的图像后, 得到共 556 张飞机蒙皮铆钉损伤图像。为增加模型的鲁棒性和泛化能力, 减少过拟合的影响, 对该数据集进行数据增强, 使用包括水平翻转、对比度变化、高斯噪声等 9 种数据增强方法, 扩充比例为 1:9, 扩充后的数据集包括铆钉损伤图像 5 004 张。扩充后的数据集效果如图 5 所示。对扩充后的数据集按照训练集、验证集和测试集以 8:1:1 的比例进行划分。图 6 为数据集目标分布统计结果。

2.2　对比实验

2.2.1　评价指标

实验硬件配置包括 NVIDIA GeForce RTX 4060

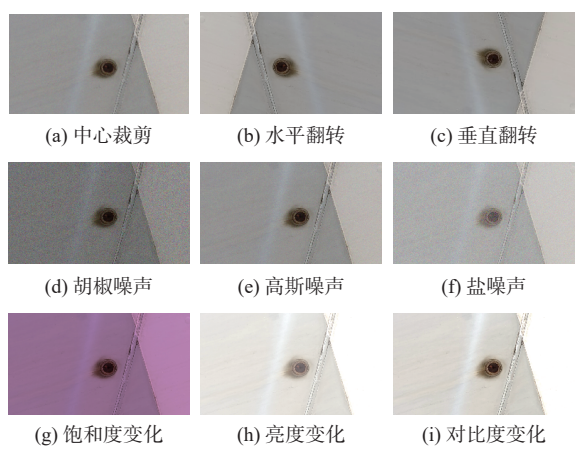


图 5　数据增强后的数据集效果

Fig. 5　Dataset after data enhancement

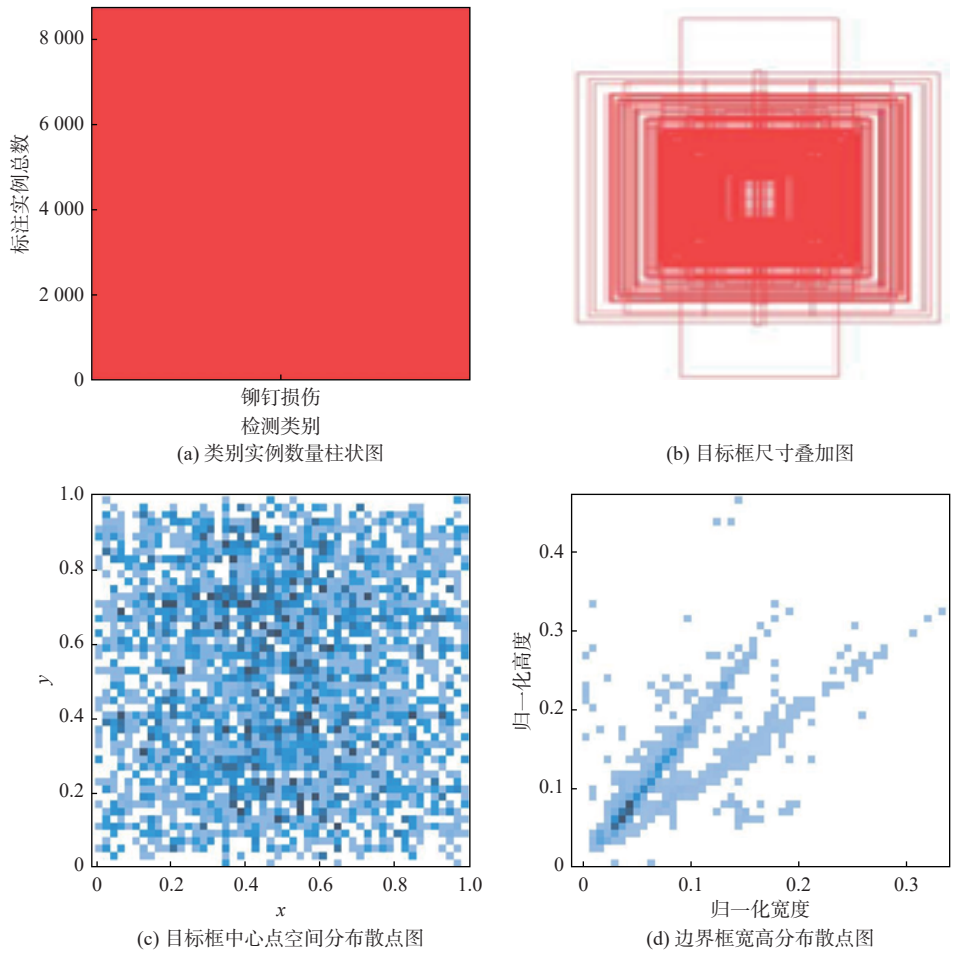


图6 数据集的目标尺寸分布

Fig. 6 Target sizes distribution of dataset

Ti GPU和Intel(R) Core(TM) i5-13400F 2.50 GHz, 其中, GPU的内存为8 GB RAM。软件方面, 本文使用Python3.11.9、PyTorch2.2.1和Cuda12.4深度学习框架。实验中, 使用SGD优化器, 初始学习率设置为0.01, 批处理大小为16, 输入图像尺寸为640像素×640像素。

采用精确度P、召回率R、mAP@0.5、参数量、浮点运算数和检测速度等6个指标对网络性能进行评估。

P是指所有样本中预测正确样本所占比例, R是指预测正确样本占所有正样本的比例, 两者具体计算公式如下:

$$P=\frac{T_P}{T_P+F_P}\times 100\% \tag{11}$$

$$R=\frac{T_P}{T_P+F_N}\times 100\% \tag{12}$$

式中: T_P为正样本中预测正确的数量; F_P为负样本中预测错误的数量; F_N为将正样本预测为负样本的数量。

mAP@0.5表示所有类别的平均精度(average precision, AP)在交并比(intersection over union, IoU)阈值为0.5时的平均值, 计算公式如下:

$$S_{\mathrm{mAP}}=\frac{\sum_{i=1}^{N}S_{\mathrm{AP}_i}}{N}\times 100\% \tag{13}$$

式中: N为检测总类别数; S_{AP}表示以P为横轴、R为纵轴所围成的P-R曲线下的面积; S_{mAP}表示mAP的值, 为AP的平均值。

2.2.2 Light-EMA模块对比实验

在YOLOv8结构中添加EMA模块时, 本文对比了将EMA模块置于颈部网络的C2f模块后和置于骨干网络的C2f模块后的网络预测性能, 对比实验结果如表1所示。通过对比实验可知, 与将EMA模块置于颈部C2f模块之后相比, 将其置于骨干C2f模块之后时R值、mAP@0.5值略有下降, 但P值提升了1.8%, 因此, 本文采用将EMA模块置于骨干部分的C2f模块后。

为验证加入EMA模块能增强模型对小目标的识别能力, 使用梯度加权类激活映射(gradient-weighted class activation mapping, Grad-CAM)视觉解释方法输出网络在预测过程中的热力图, 在目标检测领域, 小目标通常被定义为尺寸小于图像总尺寸

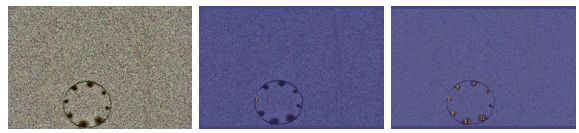
表 1　在不同位置应用 EMA 模块的网络效果对比

Table 1　Comparison of network effects of applying EMA modules at different locations

模型	*P*/%	*R*/%	mAP@0.5/%	参数量	检测速度/(帧·s^{-1})	浮点运算数/10^{9}次
YOLOv8n	90.5	83.5	88.9	3.0×10^6	159	8.1
EMA-颈部	89.2	85.2	89.9	2.9×10^6	140	8.4
EMA-骨干	91.0	85.0	89.4	3.1×10^6	123	8.4

0.5% 的物体, 其对比度通常低于 0.15[14], 本文选取数据集中符合小目标特征定义的图像进行可视化, 热力图如图 7 所示。可知, 加入 EMA模块后, 模型对小目标特征的检测能力显著提高。

观察到, 在 EMA 模块中, 1×1分支和3×3分支分别使用了卷积核为1×1和3×3的卷积, 相比于 LightConv, 参数量大、浮点运算数大, 从而影响了整体算法的运算效率。因此, 本文使用 LightConv 对 EMA 模块中的 Conv1×1 的 Conv3×3 进行替换, 使用一维 LightConv 替换1×1卷积、使用三维 LightConv 替换3×3卷积。在进行对比实验时, 将 Conv1×1的位置标记为 1 位置、Conv3×3的位置标记为 2 位置, 分别对 1 位置进行替换、对 2 位置进行替换和对 1、2 位置同时进行替换, 并且分别命名为 EMA (L1)、EMA(L2)、EMA(L12)。对比实验结果如表 2 所示。可见, 对 1、2 位置同时进行替换, 在 *P* 值、*R* 值和 mAP@0.5 值上均得到了提升, 同时减少了模型的参数量, 最终选择对 1、2 位置均使用 LightConv 进行替换形成 Light-EMA 模块。



(a) 原图　(b) YOLOv8n　(c) YOLOv8n+EMA

图 7　通过 Grad-CAM 可视化添加 EMA 机制后预测热力图

Fig. 7　Predicted heat map after adding EMA mechanism through Grad-CAM visualization

表 2　改进 EMA 模块对比实验结果

Table 2　Comparative experimental results of improved EMA module

模型	*P*/%	*R*/%	mAP@0.5/%	参数量	检测速度/(帧·s^{-1})	浮点运算数/10^{9}次
YOLOv8n	90.5	83.5	88.9	3.0×10^6	159	8.1
EMA(L1)	90.8	84.7	90.2	3.0×10^6	120	8.6
EMA(L2)	91.0	85.5	90.1	3.0×10^6	123	8.5
EMA(L12)	91.3	87.5	90.4	2.9×10^6	117	8.7

2.2.3　消融实验

为验证本文改进的有效性, 对比了改进过程中依次加入 Light-EMA 模块、HWD 模块和 NWD 损失函数的网络模型预测性能。消融实验结果如表 3 所示。

表 3　LHN-YOLO 的消融实验结果

Table 3　Results of ablation experiments with LHN-YOLO

模型	*P*/%	*R*/%	mAP@0.5/%	参数量	检测速度/(帧·s^{-1})	浮点运算数/10^{9}次
YOLOv8n	90.5	83.5	88.9	3.0×10^6	159	8.1
YOLOv8n+Light-EMA	91.3	87.5	90.4	2.9×10^6	117	8.7
YOLOv8n+Light-EMA+HWD	92.0	84.5	90.0	2.7×10^6	126	8.0
YOLOv8n+Light-EMA+HWD+NWD	92.7	87.0	91.7	2.7×10^6	119	8.2

从消融实验结果可知, 在原模型的骨干部分中加入 Light-EMA 模块, *P* 值、*R* 值和 mAP@0.5 值较 YOLOv8n 模型分别提升了 0.8%、4.0% 和 1.5%, 并且减少了 0.1×10^6 的参数量。表明 Light-EMA 模块在嵌入模型后, 减少参数量的同时提升了模型的检测精度; 将骨干部分和颈部中的卷积块替换为 HWD 模块, 形成 YOLOv8n+Light-EMA+HWD 模型, 较改进前 YOLOv8n 模型的 *P* 值、*R* 值和 mAP@0.5 值分别提升了 1.5%、1.0% 和 1.1%, 同时参数量下降了 0.3×10^6, 有效达到了减少模型参数量的同时提升检测精度的目标; 引入 NWD 模块替换原模型中的 CIoU 损失函数, 最终的 YOLOv8n+Light-EMA+HWD+NWD 模型相比于改进前 YOLOv8n 模型的 *P* 值、*R* 值和 mAP@0.5 值分别提升了 2.2%、3.5% 和 2.8%, 加入 NWD 模块使原模型更加聚焦于小目标特征, 在小目标特征占比较大时能更快收敛, 稳定性更好。最终模型的 *P* 值、*R* 值、mAP@0.5 值分别为 92.7%、87.0% 和 91.7%, 能有效地进行飞机蒙皮小目标检测任务。

2.2.4　不同模型实验及预测效果对比分析

为验证本文算法的优越性, 在使用相同实验条件、相同数据集的情况下, 将改进后的算法与其他算法进行对比。在对比实验中选取参数量与 YOLOv8n 相近的模型进行比较, 包括 YOLOv5n、YOLOv6n、YOLOv7-tiny、YOLOv8n、YOLOv9c。

分析表 4 可知, 与其他 YOLO 系列算法相比, LHN-YOLO 算法在参数量上要高于 YOLOv5n, 但在 *P* 值、*R* 值和 mAP@0.5 值上的提升达到了 2.5%、2% 和 2.8%; 对比 YOLOv8n 算法, LHN-YOLO 的检

表4 不同模型的网络预测性能对比

Table 4 Comparison of network prediction performance of different models

模型	P/%	R/%	mAP@0.5/%	参数量	检测速度/(帧·s^{-1})	浮点运算数/10^9次
YOLOv5n	90.2	85.0	88.9	2.5×10^6	145	7.1
YOLOv6n	88.5	83.2	87.9	4.7×10^6	137	11.8
YOLOv7-tiny	78.3	77.5	77.3	6.0×10^6	127	13.2
YOLOv8n	90.5	83.5	88.9	3.0×10^6	159	8.1
YOLOv9c	89.8	82.6	88.0	25.3×10^6	94	102.3
LHN-YOLO	92.7	87.0	91.7	2.7×10^6	119	8.2

测速度提升了25.2%，但在P值、R值和mAP@0.5值方面提升了2.2%、3.5%和2.8%。对比其他算法，LHN-YOLO也在P值、R值和mAP@0.5值上更有优势，同时参数量少于大部分模型，证明模型对比其他算法具有一定的优越性。

2.2.5 LEVIR-ship数据集下的模型性能验证

为验证本文算法在小目标检测性能方面相较于YOLOv8n算法有所提升，使用由卫星拍摄的小目标数据集LEVIR-ship进行验证[21]。

LEVIR-ship数据集使用了高分一号和高分六号卫星的多光谱相机所拍摄的图像，只使用R、G和B波段，在像素分辨率为10 000像素×10 000像素～50 000像素×20 000像素的数据集中采集了85个场景。对原始图像进行裁剪，最终得到1 973个正样本和1 903个负样本，大小为512像素×512像素。数据集包含了一类对象：船舶类对象，且船舶类对象像素尺寸几乎都在20像素×20像素以下，集中在10像素×10像素左右。因此，该数据集能够很好地验证改进后算法相较于基准算法在小目标检测领域的性能提升，对比实验结果如表5所示。

表5 改进前后模型在LEVIR-ship数据集上的精度对比

Table 5 Comparison of accuracy of model before and after improvement on LEVIR-ship dataset

模型	P/%	R/%	mAP@0.5/%	参数量	检测速度/(帧·s^{-1})	浮点运算数/10^9次
YOLOv8n	80.4	78.3	81.7	3.0×10^6	294	8.1
LHN-YOLO	85.4	77.6	82.4	2.7×10^6	232	8.2

由表5可得，在LEVIR-ship数据集中，LHN-YOLO算法在R值略微下降的情况下，P值得到了显著提升，增加了5.0%，mAP@0.5值也提升了0.7%。通过对比实验验证了算法在小目标检测性能方面相较于YOLOv8n算法有所提升。

3 结 论

1）为了对飞机蒙皮存在的小目标特征进行有效检测，同时尽可能地降低网络的参数量，针对飞机蒙皮铆钉损伤这一典型缺陷，提出融合注意力机制的轻量化飞机蒙皮损伤小目标检测算法LHN-YOLO。

2）设计了Light-EMA机制，使用轻量化卷积LightConv对EMA机制中的二维卷积进行优化，并将其应用于骨干网络部分，在减少模型参数量和计算复杂度的情况下，提升了网络对小目标特征的检测能力；在骨干网络和颈部网络中使用HWD模块代替原模型的下采样操作，能在下采样操作的同时尽可能地保留小目标的特征信息；在损失函数中使用NWD损失函数代替CIoU损失函数，对小目标特征进行检测时更稳定，有助于模型进一步收敛。

3）在自制数据集下，LHN-YOLO算法在较YOLOv8n网络参数量减少了10%的情况下，准确率、召回率、平均精度均值分别提升了2.2%、3.5%和2.8%；在公开小目标船舶数据集LEVIR-ship下，准确率较YOLOv8n网络提升了5.0%，验证了本文算法在小目标检测领域的有效性。

LHN-YOLO算法虽然在小目标检测性能和参数量上对比原模型和其他模型都有所优化，但在检测速度和网络参数量上仍有优化空间，未来改进后可以更好地应用于飞机蒙皮损伤检测任务中。

参考文献（References）

[1] 章东平, 王杼涛, 夏岳键, 等. 基于改进YOLOv8的飞机蒙皮缺陷检测算法[J]. 北京航空航天大学学报, 2026, 52(1): 38-48.
Zhang D P, Wang Z T, Xia Y J, et al. Aircraft skin defect detection algorithm based on enhanced YOLOv8[J]. Journal of Beijing University of Aeronautics and Astronautics, 2026, 52(1): 38-48(in Chinese).

[2] Wang P, Luo Y, Zhu Z L. FDI-YOLO: feature disentanglement and interaction network based on YOLO for SAR object detection[J]. Expert Systems with Applications, 2025, 260: 125442.

[3] Tao X, Zhang D P, Ma W Z, et al. Automatic metallic surface defect detection and recognition with convolutional neural networks [J]. Applied Sciences, 2018, 8(9): 1575.

[4] Ding M, Wu B E, Xu J, et al. Visual inspection of aircraft skin: automated pixel-level defect detection by instance segmentation[J]. Chinese Journal of Aeronautics, 2022, 35(10): 254-264.

[5] 刘振江, 张会娟, 姬淼鑫, 等. 轻量级多尺度特征融合增强的空间非合作小目标检测算法[J/OL]. 北京航空航天大学学报, 2024 [2024-09-20]. https://doi.org/10.13700/j.bh.1001-5965.2024.0509.
Liu Z J, Zhang H J, Ji M X, et al. Lightweight multi-scale feature fusion enhanced spatial non-cooperative small target detection algorithm[J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2024[2024-09-20]. https://doi.org/10.13700/j.bh.1001-5965.2024.0509(in Chinese).

[6] 郭艳, 王智文, 赵润星. YOLO-POD: 基于多维注意力机制的高精度PCB微小缺陷检测算法[J]. 电子学报, 2024, 52(7): 2515-2528.
Guo Y, Wang Z W, Zhao R X. YOLO-POD: high-precision PCB tiny-defect detection algorithm based on multi-dimensional atten-

tion mechanism[J]. Acta Electronica Sinica, 2024, 52(7): 2515-2528 (in Chinese).

[7] 吴一全, 童康. 基于深度学习的无人机航拍图像小目标检测研究进展[J]. 航空学报, 2025, 46(3): 130755.
Wu Y Q, Tong K. Research advances on deep learning-based small object detection in UAV aerial images[J]. Acta Aeronautica et Astronautica Sinica, 2025, 46(3): 130755(in Chinese).

[8] 刘延芳, 佘佳宇, 袁秋帆, 等. 无人机遥感图像实时小目标检测方法[J]. 航空学报, 2024, 45(14): 53-72.
Liu Y F, She J Y, Yuan Q F, et al. Real-time small target detection networks for UAV remote sensing[J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(14): 53-72(in Chinese).

[9] 张光华, 李聪发, 李钢硬, 等. 基于改进 YOLOv7-tiny 的无人机航拍图像小目标检测算法[J]. 工程科学与技术, 2025, 57(3): 235-246.
ZHANG G H, LI C F, LI G Y, et al. Small target detection algorithm for UAV aerial images based on improved YOLOv7-tiny[J]. Advanced Engineering Sciences, 2025,57(3): 235-246(in Chinese).

[10] 桑嘉更, 张志佳, 肖传民, 等. 基于 YOLOv8s 模型改进的道路交通目标检测方法研究[J]. 红外与激光工程, 2024, 53(11): 20240256.
Sang J G, Zhang Z J, Xiao C M, et al. An improved YOLOv8s method and its application in road traffic target detection[J]. Infrared and Laser Engineering, 2024, 53(11): 20240256(in Chinese).

[11] 李子豪, 王正平, 贺云涛. 基于自适应协同注意力机制的航拍密集小目标检测算法[J]. 航空学报, 2023, 44(13): 239-249.
Li Z H, Wang Z P, He Y T. Aerial-photography dense small target detection algorithm based on adaptive cooperative attention mechanism[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(13): 239-249(in Chinese).

[12] 李耀华, 刘传召. 融合注意力机制的轻量化航空发动机表面缺陷检测算法[J]. 航空动力学报, 2025, 40(6): 20240096.
Li Y H, Liu C Z. Lightweight aero-engine surface defect detection algorithm fused with attention mechanism[J]. Journal of Aerospace Power, 2025, 40(6): 20240096(in Chinese).

[13] Wang J X, Liu M, Du Y R, et al. PG-YOLO: an efficient detection algorithm for pomegranate before fruit thinning[J]. Engineering Applications of Artificial Intelligence, 2024, 134: 108700.

[14] Zhu J X, Qin C, Choi D. YOLO-SDLUWD: YOLOv7-based small target detection network for infrared images in complex backgrounds[J]. Digital Communications and Networks, 2025, 11(2): 269-279.

[15] 王海燕. 基于图像识别的飞机蒙皮铆接质量检测技术研究[D]. 沈阳: 沈阳航空航天大学, 2019.
Wang H Y. Research on quality detection technology of aircraft skin riveting based on image recognition[D]. Shenyang: Shenyang Aerospace University, 2019(in Chinese).

[16] Xu G P, Liao W T, Zhang X, et al. Haar wavelet downsampling: a simple but effective downsampling module for semantic segmentation[J]. Pattern Recognition, 2023, 143: 109819.

[17] Xu C, Wang J W, Yang W, et al. Detecting tiny objects in aerial images: a normalized Wasserstein distance and a new benchmark[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 190: 79-93.

[18] JOCHERG.UltralyticsYOLOv8[EB/OL].[2023-01-10].https://github.com/ultralytics/ultralytics.

[19] Ouyang D L, He S, Zhang G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2023: 1-5.

[20] Wu F, Fan A, Baevski A, et al. Pay less attention with lightweight and dynamic convolutions[PP/OL]. arXiv (2019-02-22)[2024-09-01]. https://doi.org/10.48550/arxiv.1901.10430.

[21] Chen J Q, Chen K Y, Chen H, et al. A degraded reconstruction enhancement-based method for tiny ship detection in remote sensing images with a new large-scale dataset[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5625014.

Aircraft skin damage LHN-YOLO small target detection algorithm incorporating attention mechanisms

LI Yaohua*, CHEN Jie

(School of Transportation Science and Engineering, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Aiming at the problems of low detection accuracy and a large number of parameters for small targets in aircraft skin damage detection, an LHN-YOLO algorithm incorporating the attention mechanism is proposed based on the YOLOv8n algorithm with aircraft skin rivet damage as the research object. To augment the model's capacity for detecting small targets, an optimized efficient multi-scale attention (EMA) mechanism is integrated into the backbone, with the EMA mechanism refined through the lightweight convolution LightConv, resulting in the Light-EMA module that enhances detection precision while minimizing computational complexity; concurrently, the Haar wavelet downsampling module and normalized Gaussian Wasserstein are incorporated to supplant the original model's downsampling module. Distance is introduced to replace the downsampling operation and CIoU loss function of the original model to improve the prediction performance. The results of the experiment reveal that the LHN-YOLO algorithm's average accuracy, recall, and precision are 92.7%, 87.0%, and 91.7%, respectively, which are improvements of 2.2%, 3.5%, and 2.8% over the original model, alongside a 10% reduction in parameters, significantly boosting the detection of small target damage on aircraft surfaces.

Keywords: YOLOv8n; attention mechanism; small target detection; rivet damage; skin damage detection

Received: 2024-09-27; **Accepted**: 2024-12-19; **Published Online**: 2025-02-18 15:03

URL: link.cnki.net/urlid/11.2625.V.20250218.1111.001

Foundation item: National Natural Science Foundation of China-Civil Aviation Joint Research Fund of Civil Aviation Administration of China (CAAC) Key Supporting Projects (U2033209)

* **Corresponding author**. E-mail: yh_li@cauc.edu.cn