DOI：10.20271/j.cnki.1673-9833.2026.5010

基于改进 YOLOv8 无人机目标检测算法

王子傲[1]，孙 晓[1]，罗剑宾[2]

（1. 湖南工业大学 机械工程学院，湖南 株洲 412007；2. 株洲南方阀门股份有限公司，湖南 株洲 412007）

摘 要：针对无人机图像检测中目标多尺度变化和小目标检测困难的问题，提出了一种基于改进 YOLOv8 的目标检测算法。为提高检测精度并降低计算量，设计了轻量化可加权深度分离卷积模块，减少模型参数和计算复杂度的同时，保持较强的特征提取能力。为增强多尺度特征融合，提出强化特征金字塔网络结构，通过动态权重调整机制，提升对小目标和远距离目标的检测能力。此外，采用归一化 Wasserstein 距离损失函数，并增加辅助检测头，进一步提升不同尺度目标检测的鲁棒性和精度。实验结果表明，改进算法在 Drone 数据集上的平均精度（mAP）达 64.9%，较原始 YOLOv8 提升了 1.76%，检测精度提高了 3%，验证了所提方法在无人机目标检测中的有效性。

关键词：YOLOv8；无人机；多尺度变换；小目标；NWD 损失函数

中图分类号：TP273 **文献标志码**：A **文章编号**：1673-9833(2026)05-0074-08

引文格式：王子傲，孙 晓，罗剑宾 . 基于改进 YOLOv8 无人机目标检测算法 [J]. 湖南工业大学学报，2026，40(5)：74-81.

UAV Target Detection Algorithm Based on the Improved YOLOv8

WANG Zi’ao[1]，SUN Xiao[1]，LUO Jianbin[2]

（1. School of Mechanical Engineering，Hunan University of Technology，Zhuzhou Hunan 412007，China；
2. Zhuzhou Nanfang Valve Co., Ltd.，Zhuzhou Hunan 412007，China）

Abstract：In view of the challenges of multi-scale variations of targets and the difficulty in detecting small targets in UAV image detection, an object detection algorithm has thus been proposed based on an improved YOLOv8. To improve detection accuracy and reduce computational load, a lightweight weighted depth-wise separable convolution module is designed, which reduces model parameters and computational complexity while maintaining strong feature extraction capability. To enhance multi-scale feature fusion, a strengthened feature pyramid network structure is proposed, which utilizes a dynamic weight adjustment mechanism to improve the detection capability for small and distant targets. In addition, a normalized Wasserstein distance loss function is adopted, with an auxiliary detection head added, further enhancing the robustness and accuracy of target detection at different scales. The experimental results show that the improved algorithm can achieve an average precision (mAP) of 64.9% on the Drone dataset, an increase of 1.76% compared to the original YOLOv8, with a 3% improvement in detection accuracy, thus confirming the effectiveness of the method in drone object detection.

Keywords：YOLOv8；UAV；multiscale transformation；small object；NWD loss function

收稿日期：2025-02-19

基金项目：湖南省重点领域研发计划基金资助项目（2022GK2068）；湖南省自然科学基金省市联合基金资助项目（2021JJ50053）

作者简介：王子傲，男，湖南工业大学硕士生，主要研究方向为无人机集群视觉控制和运动感知，E-mail：1481004061@qq.com

1　研究背景

近年来，无人机集群技术在多个领域展现出广泛的应用前景，包括环境监测、农业管理、集群对抗等。无人机集群的识别与运动感知能力在确保集群协同飞行和任务执行中起着至关重要的作用，尤其是在复杂环境下的目标识别与动态跟踪 [1]。无人机集群图像通常存在多目标、尺度变化大、背景复杂、遮挡与动态模糊等特点。此外，无人机在飞行过程中的快速移动，导致图像中的目标呈现出不同的运动模态和形变，使得图像处理更加复杂。然而，传统目标检测算法通常需要较高的计算资源，难以满足无人机集群平台对低功耗和高实时性的严格要求。近年来，深度学习技术的迅猛发展为无人机集群的目标检测与运动感知提供了新的解决思路。在目标检测领域，尤其是图像处理方面，深度学习技术通常分为两阶段（Two-Stage）算法和一阶段（One-Stage）算法 [2]。

两阶段算法通常包含两个主要步骤，它们会生成一系列候选区域，这些区域可能包含目标对象。然后，在第二阶段，算法会对这些候选区域进行分类，确定它们是否包含目标，并调整这些候选框以更准确地定位目标对象位置。R-CNN[3]（region-based convolutional neural networks）是两阶段算法的代表性方法。然而，由于它们需要处理大量的候选框，因此通常速度较慢，并不适用于无人机图像实时检测。

一阶段算法则直接对图像中的目标进行定位和分类，无需像两阶段算法那样先生成候选区域。它们通常将目标检测任务视为一个回归问题，直接预测目标对象的位置和类别。YOLO[4]（You Only Look Once）和 SSD[5]（Single Shot Multibox Detector）是一阶段算法的代表性方法。它们直接在图像的多个位置预测目标对象的存在和类别，特别是 YOLO 系列算法凭借其端到端的快速检测能力，成为了目标检测中的主流方法。

为了提高 YOLO 算法在无人机图像上的检测能力，研究人员做出了一系列的改进。Bao Z.[6] 提出了一种基于改进 YOLOv8 的无人机图像检测方法，该方法通过结合 ParNet 注意力机制和 RepVGG 重参数化模块，提高了目标检测的准确性和推理效率。Chang Y. L. 等 [7] 提出了一种基于 YOLO 模型的改进小目标检测算法，将网络框架中原来的卷积换成 SPD-Convolution 模块，以消除池化操作对特征信息的影响，并增强模型从低分辨率和小目标中提取特征的能力，但增加了模型的参数量。Tan L. 等 [8] 提出了一种改进的 YOLO 算法，通过空心卷积重采样提高特征提取性能、采用超轻量级子空间注意力机制（ULSAM）增强多尺度特征表示，并引入软非极大值抑制（Soft-NMS）减少目标遮挡带来的漏检问题，但速度有所降低。Dong Y. 等 [9] 提出了一种轻量级无人机检测模型，通过结合深度学习目标检测与超分辨率特征增强，改进 YOLO 算法并应用通道和图层修剪，提升了小目标检测能力，但缺少对多尺度目标的检测能力。 Luo X. D. 等 [10] 对 YOLO 进行了改进，针对无人机航拍图像中的小目标和复杂背景，采用不同激活函数、EIoU 损失、IECA 模块、SPP 替代金字塔池化，并引入自适应空间特征融合（ASFF）模块，提升了目标检测性能。Hui Y. 等 [11] 提出了一种创新的 STRCN 卷积结构，结合了 SwinTransformer 和 CNN，设计了轻量级分类器 CNeB、改进的 CWDHead 检测头，以及基于 SimAM 和新型数据增强方法的优化策略，但对小目标的检测能力较差。

综上所述，尽管现有研究在特征提取、轻量化优化、损失函数改进等方面做出了贡献，但仍存在计算复杂度高、小目标检测能力有限、多尺度目标适应性不足等问题。因此，本文提出了一种基于改进 YOLOv8 的多尺度特征融合无人机目标检测模型 LFN-YOLO，通过引入学习型权重分配机制，在特征融合过程中动态调整不同尺度、不同区域的权重，从而增强了特征图的适应性；通过定制的特征聚焦模块与特征扩散机制，能让每个尺度的特征都具有详细的上下文信息，更有利于后续目标的检测与分类；通过 Wasserstein 距离过滤掉相似的候选框，优化了边界框合并策略，提高了多目标检测的精度。

2　网络模型

2.1　YOLOv8 模型

YOLOv8 模型主要由 3 个部分组成：Backbone、Neck 和 Head，其结构如图 1 所示。YOLOv8 采用了先进的骨干网络和颈部架构，它采用了无锚点的分割 Ultralytics 头部设计，这有助于提高准确性并使检测过程更加高效。YOLOv8 还专注于在准确性和速度之间保持最佳平衡，适合于不同应用领域的实时目标检测任务。Backbone 部分负责特征提取，采用了一系列卷积和反卷积层，同时使用了残差连接和瓶颈结构来减小网络的大小并提高网络性能。该部分采用 C2f 模块作为基本构成单元，与 C3 模块相比，C2f 模块通过更有效的结构设计，减少了冗余参数，提高了计算效率。Neck 部分负责多尺度特征融合，通过将来自 Backbone 不同阶段的特征图进行融合，增强特征表达能力。Head 部分负责最终的目标检测

和分类任务，包括一个检测头和一个分类头，检测头包含一系列卷积层和反卷积层，用于生成检测结果。分类头采用全局平均池化（global average pooling）对每个特征图进行分类，通过减少特征图的维度，输出每个类别的概率分布。

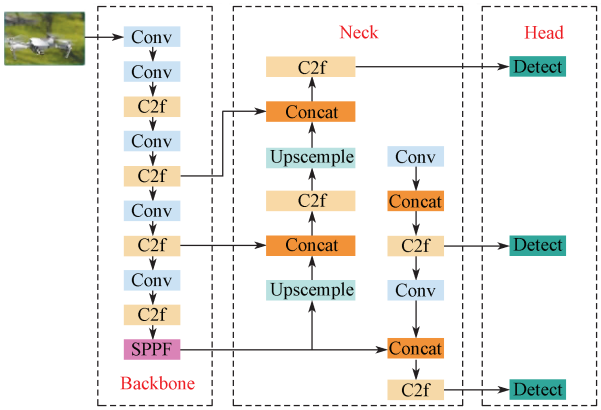


图 1 YOLOv8 网络模结构型图

Fig. 1 YOLOv8 network model diagram

2.2 轻量化提取卷积 LAWDS

在传统无人机目标检测模型中，增强模型的检测精度往往伴随着计算量增加，尤其在多尺度融合和注意力机制的应用中，计算复杂度会显著增加。为了在保持较高检测性能的同时确保实时性，并适应边缘计算和移动设备上的部署需求，本文借鉴了 RFAConv 卷积[12]的思想，提出一种轻量化的卷积模块 LAWDS，用于代替原有的 Conv 模块。LAWDS 模块结构如图 2 所示。

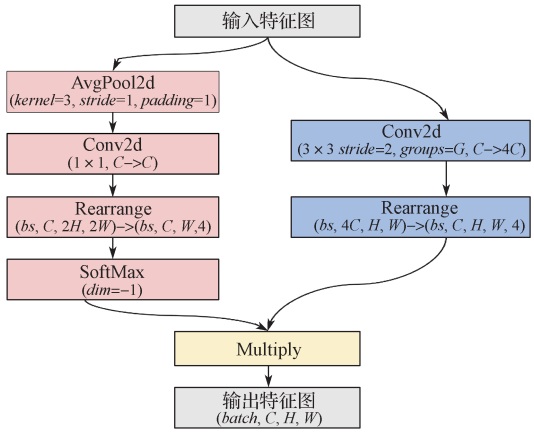


图 2 LAWDS 结构图

Fig. 2 LAWDS structure diagram

由图 2 可知，LAWDS 模块分为注意力分支和下采样分支。注意力分支为了提取特征图的局部上下文信息，同时保留输入的通道数，对输入特征图执行 3 × 3 平均池化操作，步幅为 1，填充为 1。使用 1 × 1 卷积对池化后的特征图进行处理，通道数保持不变。为了将每个子区域的注意力特征归一化为概率分布，确保每个像素点的重要性可以动态加权调整，将注意力特征图按子区域（2 × 2patch）的方式划分，并在每个子区域内部执行 SoftMax 操作。LAWDS 模块利用 SoftMax 生成的动态注意力权重，可以根据输入特征图的局部信息分配资源，确保模型更加关注关键区域。同时，下采样分支为了在降低分辨率的同时保留特征信息，并为后续的加权操作提供足够的表达能力，对输入特征图进行 3 × 3 的分组卷积，步幅为 2，分辨率减小一半，输出的通道数扩展为 4 × channels，并且进行重新排列与注意力权重对齐。在下采样过程中，模块通过权重分配有效整合低级特征和高级特征，减少信息丢失，保留目标的多尺度信息。

在传统 YOLOv8 中，特征融合是通过简单的加权求和或拼接实现的，这种固定的融合方式在无人机目标检测下无法充分利用不同尺度的优势，尤其是对于多尺度的无人机目标类型，固定的融合策略可能会导致性能瓶颈。LAWDS 卷积模块结合了分组卷积、动态注意力机制和下采样策略，在保持较高计算效率的同时，能够有效提高对小目标和多尺度目标的检测能力。它通过逐元素加权的方式调整特征图，使得模型能够更加关注目标区域，同时抑制背景噪声，显著提升了目标检测的精度和性能。经过实验对比，LAWDS 在替换原有的卷积后，在计算量和参数量降低的同时，精确度有所提升。

2.3 特征聚焦扩散金字塔网络 FDPN

特征金字塔网络[13]（feature pyramid network，FPN）是一种深度学习的网络结构，主要用于提取图像的多尺度特征。通过构建不同层次的特征图，FPN 能够有效捕捉从细粒度到粗粒度的信息，适用于多种计算机视觉任务，如目标检测、语义分割等。

传统的 YOLO 系列（包括 YOLOv8）采用了标准的 FPN 来处理多尺度目标。但 FPN 主要通过自上而下的特征融合进行上采样与加权融合，依赖于每一层的单一尺度特征。在处理复杂场景，尤其是多尺度目标时，可能无法充分利用低、中、高分辨率特征的协同作用。因此为了能让每个尺度的特征都具有详细的上下文信息，更有利于后续目标的检测与分类，设计了 FDPN 模块，FDPN 结构如图 3 所示。

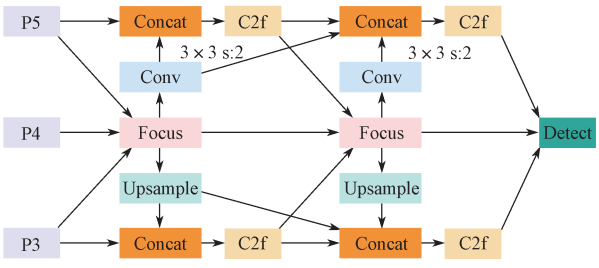


图 3 FDPN 结构图

Fig. 3 FDPN structure diagram

通过改进的特征融合机制和定制的特征聚焦模块与特征扩散机制，在传统 FPN 结构的基础上增强了不同尺度特征之间的相互联系。

在改进 FDPN 模块中加入了 Focus Feature 模块，结构如图 4 所示。通过下采样对齐多尺度特征的分辨率，同时减少不必要的计算开销，对浅层特征 P3 使用 ADown 模块进行下采样，将其分辨率调整为与 P4 和 P5 一致。将经过处理的 P3、P4 和 P5 特征沿通道维度（*dim*=1）拼接，得到一个融合后的特征图。将拼接后的特征图分别通过不同感受野的深度可分离卷积，为了保留原始信息，避免多次卷积处理可能导致的信息损失，拼接后的特征图通过一个 Identity 路径，保持原始特征不变。将多尺度深度卷积（5×5、7×7、9×9、11×11）的输出特征与恒等映射路径的特征进行逐元素加法。最后通过逐点卷积和残存连接，减少通道数、降低计算复杂度和增强最终特征表达能力。

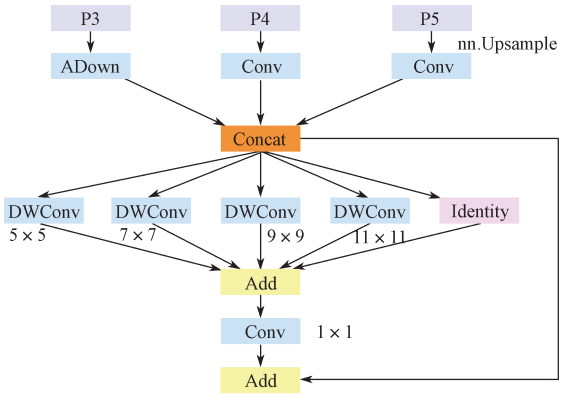


图 4　Focus Feature 结构图

Fig. 4　Focus Feature structure diagram

FDPN 模块不仅引入了加权融合，还通过多尺度特征融合（Fused P3、P4、P5）进一步提升了低级、中级和高级特征的互补性，从而改善了模型的全局感知能力和多尺度目标的检测能力。模块更加充分地利用各层次特征中的关键信息，避免了传统 FPN 可能忽略细节的问题，尤其在小目标和大目标检测任务上表现更好。通过融合不同尺度的信息，在多尺度物体上的检测精度显著提升，尤其适用于复杂背景下的无人机目标检测。

在 YOLO 系列中，计算效率和推理速度是关键指标。传统的 FPN 虽然能够提高检测精度，但计算开销较大，尤其是在多尺度融合时，可能导致推理速度下降。FDPN 模块设计时充分考虑了计算效率，通过高效的特征融合方法，在保证精度提升的同时，最大程度地减少计算负担。FDPN 模块在融合过程中采用了加权融合和高效的上采样策略，这使得特征融合操作不仅更加精确，而且计算代价较低。相较于传统的 FPN 模块，FDPN 模块在提升多尺度特征融合能力的同时，保持了较低的计算成本和更高的推理速度，达到了实时检测任务的需求。

2.4　NWD 损失函数

在 YOLOv8 中，CIoU（complete intersection over union）是一种常用的边界框回归损失函数，用于衡量预测框与目标框之间的差异。

CIoU 损失函数的定义：

$$L_{\mathrm{CIoU}}=1-IoU+\frac{\rho^2\left(b,b^{\mathrm{g}}\right)}{c^2}+\alpha v\ 。\tag{1}$$

式中：*IoU*（intersection over union）为预测框和目标框之间的交并比；$\rho^2\left(b,b^{\mathrm{g}}\right)$ 为预测框中心点和目标框中心点之间的欧几里得距离，且

$$\rho^2\left(b,b^{\mathrm{g}}\right)=\left(x-x^{\mathrm{g}}\right)^2+\left(y-y^{\mathrm{g}}\right)^2,\tag{2}$$

其中，(x,y) 和 $\left(x^{\mathrm{g}},y^{\mathrm{g}}\right)$ 分别为预测框和目标框的中心点坐标；c^2 为包围框的对角线距离的平方（包围框是包围预测框和目标框的最小矩形）；αv 为用于衡量预测框和目标框宽高比的相似性，其中，

$$v=\frac{4}{\pi^2}\left(\arctan\frac{\omega^{\mathrm{g}}}{h^{\mathrm{g}}}-\arctan\frac{\omega}{h}\right)^2,\tag{3}$$

$$\alpha=\frac{v}{\left(1-IoU\right)+v}\ 。\tag{4}$$

其中，ω、h 分别为预测框的宽度和高度，ω^{g}、h^{g} 分别为目标框的宽度和高度。

CIoU 在衡量预测框和目标框相似性时，主要集中于几何特性。然而，该方法忽略了目标形状、分布特性以及对多尺度目标的一致性优化。为此，在 YOLOv8 中提出了一种基于 Wasserstein 距离的归一化高斯分布匹配损失函数（NWD）[14]，作为边界框回归的核心优化指标。

NWD 损失通过将预测框和目标框视为二维高斯分布，计算其 Wasserstein 距离，并进行归一化处理。

对于两个多元高斯分布 $P_1=N\left(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1\right)$ 和 $P_2=N\left(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2\right)$，二阶 Wasserstein 距离的平方形式为

$$W_2^2\left(P_1,P_2\right)=\left\|\boldsymbol{\mu}_1-\boldsymbol{\mu}_2\right\|^2+Tr\left[\left(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2\right)-2\left(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2}\right)^{1/2}\right]。\tag{5}$$

式中：$\left\|\boldsymbol{\mu}_1-\boldsymbol{\mu}_2\right\|^2$ 为均值之间的欧式距离；*Tr* 为矩阵的迹；$\left(\boldsymbol{\Sigma}_1^{1/2}\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{1/2}\right)^{1/2}$ 为协方差矩阵之间的相关性部分。

为了使 Wasserstein 距离更加适合作为损失函数，常用归一化版本。NWD 的损失函数可以定义为

$$L_{\mathrm{NWD}}(P_1,P_2)=\frac{W_2^2(P_1,P_2)}{1+\|\boldsymbol{\mu}_1-\boldsymbol{\mu}_2\|^2+Tr(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)}。\quad(6)$$

归一化是为了避免因为分布的尺度（scale）或方差（variance）过大造成损失的不稳定。NWD 在 Wasserstein 距离的基础上进行了归一化，CIoU 在高重叠或低重叠情况下可能数值不稳定，而 NWD 通过归一化约束范围，提高了梯度计算的稳定性。不同目标的大小差异对训练影响减小，使得小目标和大目标的权重更均衡。归一化的 Wasserstein 距离解决了不同目标尺度下损失不均衡的问题，提升了模型对多尺度目标的检测能力。

经过上述改进后的 LFN-YOLO 模型结构如图 5 所示。

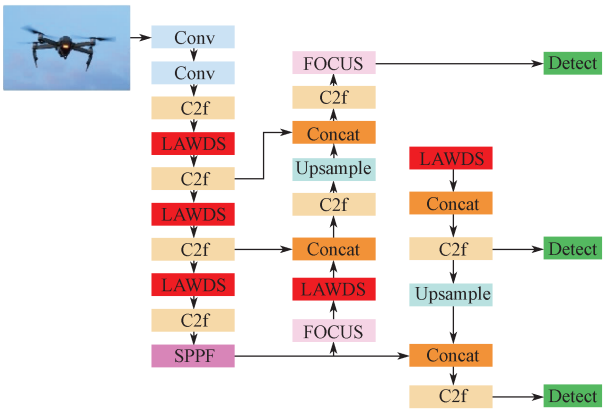


图 5 LFN-YOLO 模型结构图

Fig. 5 LFN-YOLO model structure diagram

3 实验

3.1 数据集

本研究使用的无人机目标检测数据集来源于 Kaggle，专门用于无人机检测任务，旨在训练无人机检测模型，实现对其他无人机的识别和跟踪。数据集包含 1 400 张无人机图像，该数据集涵盖不同型号的无人机，并具有显著的多尺度变化、复杂背景（如天空、建筑物、树木、车辆）及多种飞行姿态等特征，为目标检测任务提供了多样化的样本，如图 6 所示。



图 6 无人机视角数据集图

Fig. 6 Drone perspective dataset image

数据标注采用 YOLO 格式，每个标注文件包含目标类别及其归一化边界框坐标，即中心点坐标（x, y），宽度 ω，高度 h，确保模型能够高效学习无人机的空间分布特征。该数据集适用于无人机集群检测、无人机追踪及反无人机系统等应用场景，为提高无人机自主识别和智能监测能力提供了重要的数据支持。

3.2 实验环境及评价标准

本实验中，YOLOv8 所使用的是 Ultralytics 软件包，图像大小设置为 640，epochs 设置为 150，batch 设置为 4，优化器选用随机梯度下降（SGD）。

在图像目标检测中，评价指标对衡量模型性能至关重要。平均精度 AP（average precision）、精确度 P（precision）和召回率 R（recall）是最常见的评价指标，从不同的角度衡量模型的检测效果。

精确度衡量的是检测结果中正类别（positive）样本的准确率。具体来说，它是真正例（true positives，TP）除以所有被预测为正类别的样本数（TP+ 假正例 false positives，FP）。高精确度意味着模型的预测结果中，正类别的比例较高。

$$P=\frac{TP}{TP+FP}\times 100\%。\quad(7)$$

召回率衡量的是模型检测出所有实际正类别样本的能力。它是真正例（TP）除以所有实际正类别样本数（TP+ 假负例 false negatives，FN）。高召回率意味着模型能够检测出大部分实际存在的正类别样本。

$$R=\frac{TP}{TP+FN}\times 100\%。\quad(8)$$

平均精度是评估目标检测模型性能的一个重要指标，特别是在二分类问题中。它通过计算不同召回率水平上的精确度的平均值来得出。在目标检测中，AP 通常通过绘制 precision-recall 曲线并计算曲线下的面积来获得，这可以提供一个综合的模型性能度量。

$$AP=\int_0^1 P(r)\mathrm{d}r。\quad(9)$$

3.3 模型对比实验

为了评估 LADWS 卷积的有效性，选用了常见的下采样卷积进行比较。例如 ADown[15]、DSconv[16] 等。为了保证实验结果的可比性，本文所有的卷积改动都在相同位置。在 Drone 数据集进行训练和测试后，结果如表 1 所示。ADown 卷积在 $mAP50$ 没有提升，DSconv 在平均精度上提升较少，可以看出本文使用的卷积模块在检测任务精度和 mAP 上都有明显提升，实验结果表明，尽管 recall 指标存在微小波动，但其在减少漏检、平衡计算资源方面，远超传统模块，验证了改进方案的有效性。综合对比证明了 LADWS

卷积的有效性。

表 1　基于 Drone 数据集的不同卷积模块对比实验结果

Table 1　Comparative experiment results of different convolution modules based on the Drone dataset　%

卷积模块	*P*	*R*	*mAP*50	*mAP*50-95
Conv	93.5	93.0	95.8	63.2
ADown	94.3	92.5	95.8	64.0
DSconv	93.9	92.5	96.1	63.8
CARAFE	93.8	93.4	96.2	64.0
dysample	95.2	88.1	95.0	63.3
our	93.9	92.4	96.1	64.1

为了验证特征聚焦扩散金字塔网络 FDPN 的优势，使用了其他常用的小目标检测金字塔网络进行对比实验，例如 AFPN-P345[17]、GFPN[18] 等网络结构，结果如表 2 所示。从表中可以看出，本文改进的金字塔网络在精度和 mAP 上都有明显的优势，更加说明 FDPN 模块能有效地提升原模型的性能。

表 2　基于 Drone 数据集不同特征金字塔网络对比实验结果

Table 2　Comparative experiment results of different feature pyramid networks based on the Drone dataset　%

网络结构	*P*	*R*	*mAP*50	*mAP*50-95
FPN	93.5	93.0	95.8	63.2
AFPN-P345	94.3	90.5	95.1	63.1
GFPN	94.7	91.6	95.8	64.1
EfficientRepBiPAN	95.3	89.0	94.9	63.1
CA-HSFPN	93.7	92.0	95.4	63.1
our	94.8	92.0	96.1	64.7

YOLOv8 模型使用的是 CIoU 损失函数，CIoU 主要集中于几何特性，忽略了目标形状、分布特性以及对多尺度目标的一致性优化。为了解决这一局限性，本文采用了 NWD 损失函数来替代 CIoU 损失函数，并且同时选取了一些常用的损失函数进行比较，例如 WiseIoU[19] 、InnerWiseIoU[20]，实验结果如表 3 所示。可以看出，NWD 损失函数在精度和 *mAP*50-95 方面都要优于其他损失函数的。

表 3　基于 Drone 数据集不同损失函数对比实验结果

Table 3　Comparative experiment results of different loss functions based on the Drone dataset　%

函数	*P*	*R*	*mAP*50	*mAP*50-95
CIoU	94.8	92.0	96.1	64.7
WiseIoU	92.1	92.8	96.4	64.6
InnerWiseIoU	94.1	91.7	96.3	64.1
NWD	94.9	91.0	96.1	64.9

3.4　模型消融实验

基于本文提出的改进策略，为了验证不同改进策略的检测性能优劣，对不同的改进方案基于 Drone 数据集进行了消融实验，结果如表 4 所示。实验均是基于 Drone 数据集进行的，讨论不同改进策略对 YOLOv8 模型性能的影响。以 YOLOv8n 作为基础模型，原始 YOLOv8n 模型在精度、*mAP*50、*mAP*50-95 的表现分别为 93.5%, 95.8%, 63.2%。对 YOLOv8n 上的 Conv 卷积替换成 LAWDS 之后，模型的精确度提高了 0.4%，模型的 *mAP*50 提升了 0.3%，*mAP*50-95 提升了 0.9%，这表明改进的 LAWDS 卷积对模型的整体性能有一定的提升。同时引入 LAWDS 卷积和 FPND 特征金字塔网络后，模型的精度提升了 0.6%，模型的 *mAP*50 提升了 0.3%，*mAP*50-95 提升了 0.6%，这表明在加入了 FPND 模块后对模型整体的性能有一定的提升，尤其在 mAP 方面。同时加入 LAWDS 模块、FPND 模块与更换 NWD 损失函数后，模型的精度提升了 3.0%，模型的 *mAP*50 提升了 0.9%，*mAP*50-95 提升了 1.7%，这表明综合的改进策略能有效地提升模型的性能，表明了模型在对无人机目标检测能力上得到了显著增强。综合分析，对卷积、特征聚焦扩散金字塔网络和损失函数的改进方案是可行的。特别是当三者同时改进时，模型的精度和 mAP 取得了最佳性能，这表明卷积的改进有助于特征的提取，金字塔网络的改进有助于检测多尺度目标，损失函数的改进则更加准确地对目标进行定位与分类。

表 4　基于 Drone 数据集的消融实验结果

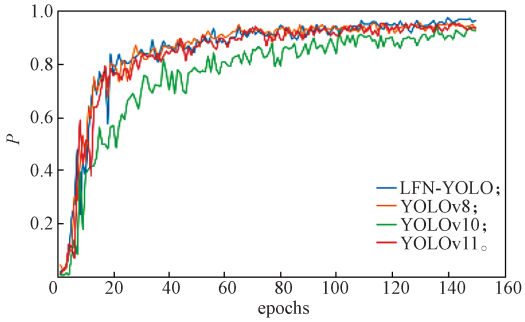
Table 4　Ablation experiment based on the Drone dataset

YOLOv8	LAWDS	FPND	NWD	*P*/%	*mAP*50/%	*mAP*50-95/%
√				93.5	95.8	63.2
√	√			93.9	96.1	64.1
√		√		94.8	96.1	64.7
√			√	95.1	96.0	64.1
√	√	√		94.1	96.1	63.8
√	√		√	95.2	95.9	64.6
√	√	√	√	96.5	96.7	64.9

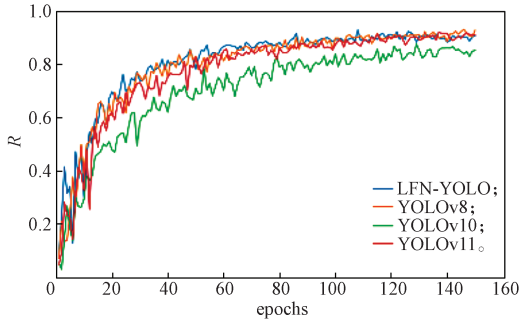
3.5　模型可视化分析

为能更直观地分析改进后模型的性能，本文将训练完成的模型数据进行可视化展示，结果如图 7 所示。从图 7 中可见，改进后模型的精度和 AP 均优于原 YOLOv8 模型以及 YOLOv10、YOLOv11 模型；尽管 YOLOv10、YOLOv11 这类新版本模型在通用目标检测任务上有性能提升，但其针对无人机检测特殊场景的优化空间已十分有限，而 YOLOv8 为本文的算法改进提供了更优的基础，由此可明确本文改进后

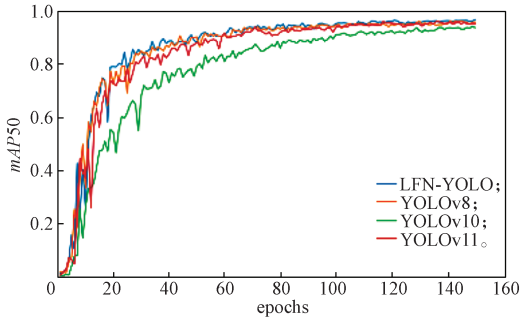
的模型性能优于上述对比模型。



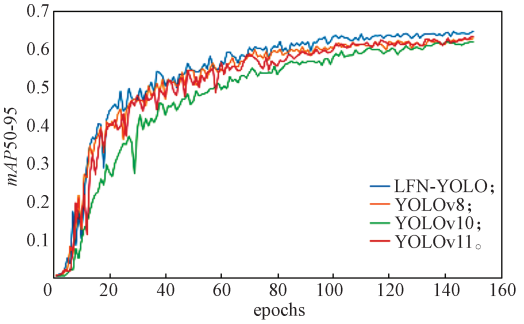
a）*P*



b）*R*



c）*mAP*50

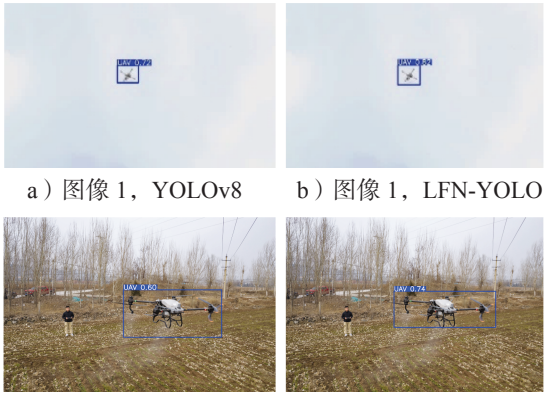


d）*mAP*50-95

图 7 不同模型测试结果对比图

Fig. 7 Comparison chart of results from different models

为了进一步验证改进前后模型的性能对比以及改进后的模型是否解决小目标和多尺度目标漏检的问题，本文用训练前后的模型对相同场景下的无人机图片进行检测。实验分为两组：第一组场景主要针对小目标的检测，结果如图 8 所示；第二组场景主要针对多尺度目标的检测，结果如图 9 所示。



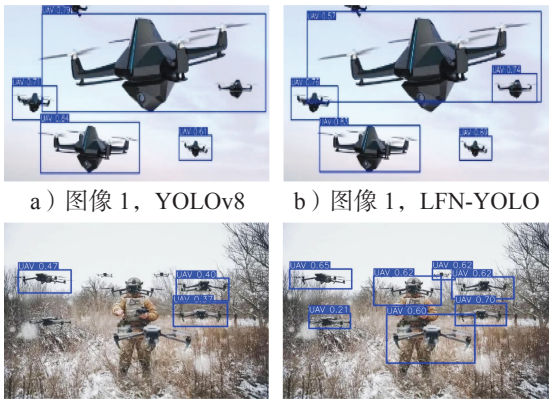
a）图像 1，YOLOv8　b）图像 1，LFN-YOLO

c）图像 2，YOLOv8　d）图像 2，LFN-YOLO

图 8 小目标检测效果对比

Fig. 8 Comparison of small target detection effects

由图 8 可知，在原模型检测下针对无人机小目标检测精度分别为 72% 和 60%，在改进后的模型检测下分别为 82% 和 74%。改进后对于不同场景的小目标检测精度都比原模型要高，证明改进后的模型对小目标的检测更加准确。



a）图像 1，YOLOv8　b）图像 1，LFN-YOLO

c）图像 2，YOLOv8　d）图像 2，LFN-YOLO

图 9 多尺度变换检测效果对比

Fig. 9 Comparison of multi-scale transformation detection effects

如图 9 所示，相比于图 8 场景更加复杂，图中无人机的尺度变化大、目标数量多。在使用原模型进行检测时，因为多尺度目标的问题导致了个别小目标出现了漏检的情况，在使用原模型检测图中无人机目标漏检率为 62.5%，在使用改进后的模型检测可以看到漏检率降低了 50%，漏检情况明显改善。

通过两组实验结果，验证了本文的算法能够有效地提高小目标的检测精度，同时检测多尺度目标时也能减少无人机漏检的情况。

4 结语

在对图像中的无人机检测中，存在目标尺寸变化大、远距离检测困难、小尺寸目标多等问题。为了在无人机检测中提高对无人机目标的检测精度，降低目标检测的错检与漏检率，便于满足无人机集群的控

制要求，本文提出了一种基于改进的 YOLOv8 无人机图像检测模型。借鉴 RFAConv 卷积的思想，本文设计了 LAWDS 模块来代替原有的 Conv 模块，有效降低了模型参数量和计算复杂度，同时保持较强的特征提取能力。为了增强模型的多尺度融合能力，提出了基于强化特征金字塔网络 FDPN 的多尺度特征融合结构，通过动态权重调整增强了对小目标和远距离目标的检测能力。采用归一化高斯 Wasserstein 距离 NWD 作为边界框回归损失函数，进一步提高了模型在不同尺度目标检测中的精度。本文将改进的算法模型与其他目标检测模型进行了对比，得知改进后的算法性能要优于其他已有模型。

尽管本文提出的基于 YOLOv8 改进的无人机图像中无人机目标检测算法在无人机目标检测中获得了性能提升，且解决了无人机多尺度变化的问题，但本文的研究有一定的局限性，当前的研究主要集中在无人机图像中无人机的检测，目标类别过于单一，且对场景的多样性和复杂性考虑不足。后续将以如何提高多场景检测水平，如夜间、强光、密集目标等特色场景的检测，作为未来的研究方向，并且对模型进行轻量化处理，便于后续的无人机集群控制研究。

参考文献：

[1] LO L Y，YIU C H，TANG Y，et al. Dynamic Object Tracking on Autonomous UAV System for Surveillance Applications[J]. Sensors，2021，21(23)：7888.

[2] MITTAL P，SINGH R，SHARMA A. Deep Learning-Based Object Detection in Low-Altitude UAV Datasets：A Survey[J]. Image and Vision Computing，2020，104：104046.

[3] GIRSHICK R，DONAHUE J，DARRELL T，et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus：IEEE，2014：580−587.

[4] REDMON J，DIVVALA S，GIRSHICK R，et al . You Only Look Once：Unified，Real-Time Object Detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vagas：IEEE，2016：779−788.

[5] LIU W，ANGUELOV D，ERHAN D，et al. SSD：Single Shot Multibox Detector[C]//Computer Vision-ECCV 2016：14th European Conference. Amsterdam：Springer International Publishing，2016：21−37.

[6] BAO Z. The UAV Target Detection Algorithm Based on Improved YOLOv8[C]//Proceedings of the International Conference on Image Processing，Machine Learning and Pattern Recognition. New York：ACM，2024：264−269.

[7] CHANG Y L，LI D，GAO Y，et al. An Improved YOLO Model for UAV Fuzzy Small Target Image Detection[J]. Applied Sciences，2023，13(9)：5409.

[8] TAN L，LÜ X，LIAN X，et al. YOLOv4_Drone：UAV Image Target Detection Based on an Improved YOLOv4 Algorithm[J]. Computers & Electrical Engineering，2021，93：107261.

[9] DONG Y，MA Y，LI Y，et al. High-Precision Real-Time UAV Target Recognition Based on Improved YOLOv4[J]. Computer Communications，2023，206：124−132.

[10] LUO X D，WU Y Q，ZHAO L Y. YOLOD：A Target Detection Method for UAV Aerial Imagery[J]. Remote Sensing，2022，14(14)：3240.

[11] HUI Y，WANG J，LI B. STF-YOLO：A Small Target Detection Algorithm for UAV Remote Sensing Images Based on Improved SwinTransformer and Class Weighted Classification Decoupling Head[J]. Measurement，2024，224：113936.

[12] ZHANG X，LIU C，YANG D，et al. RFAConv：Innovating Spatial Attention and Standard Convolutional Operation[J]. ArXiv Preprint，2023，ArXiv：2304.03198.

[13] LIN T Y，DOLLÁR P，GIRSHICK R，et al. Feature Pyramid Networks for Object Detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu：IEEE，2017：2117−2125.

[14] WANG J，XU C，YANG W，et al. A Normalized Gaussian Wasserstein Distance for Tiny Object Detection[J]. ArXiv Preprint，2021，ArXiv：2110.13389.

[15] WANG C Y，YEH I H，LIAO H Y M. YOLOv9：Learning What You Want to Learn Using Programmable Gradient Information[C]// Computer Vision-ECCV 2024. Milan：ECCV，2024：1−21.

[16] WANG C Y，BOCHKOVSKIY A，LIAO H Y M. YOLOv7：Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver：IEEE，2023：7464−7475.

[17] YANG G，LEI J，ZHU Z，et al. AFPN：Asymptotic Feature Pyramid Network for Object Detection[J]. ArXiv Preprint，2023，ArXiv：2306.15988.

[18] XU X，JIANG Y，CHEN W，et al. Damo-Yolo：A Report on Real-Time Object Detection Design[J]. ArXiv Preprint，2022，ArXiv：2211.15444.

[19] TONG Z，CHEN Y，XU Z，et al. Wise-IoU：Bounding Box Regression Loss with Dynamic Focusing Mechanism. [J]. ArXiv Preprint，2023，ArXiv：2301.10051.

[20] ZHANG H，XU C，ZHANG S. Inner-Iou：More Effective Intersection Over Union Loss with Auxiliary Bounding Box[J]. ArXiv Preprint，2023，ArXiv：2311.02877.

（责任编辑：申　剑）