基于机器视觉的工业环境中粉尘检测技术研究

苏 宁[1] 冀俊豹[1] 贺伟夕[1] 赵立强[1]
SU Ning JI Junbao HE Weixi ZHAO Liqiang

摘 要 在特定工业环境中产生的粉尘不仅对设备的运行和维护构成损坏，更对工人的健康和环境产生威胁，因此，对粉尘进行实时有效的检测并及时采取相应的措施具有重要意义。文章基于 YOLOv8 目标检测基础模型，进行了多项改进。引入了动态上采样器 DySample，以增强对不同尺寸粉尘目标的检测能力，提升小目标粉尘的检测精度。另外，在工业环境中，粉尘检测通常存在复杂的背景，如机器设备、管道等，这些背景可能会干扰模型的判断，导致误检。为了改善复杂背景下的检测性能，将 Dynamic ATSS 标签分配策略应用于训练过程中的正负样本分配阶段，通过动态调整预测框和真实框重叠程度的指标（intersection over union，IoU）阈值，使得正样本的选择更加符合粉尘目标的分布特性，提高检测器的泛化能力。实验结果显示，改进后的模型相比传统 YOLOv8 模型，在粉尘目标检测的精度、召回率以及 mAP 等指标上均有提升，其中精度提升 10%，召回率提升 3%，mAP50 提升 5% 左右，且在应对复杂背景和多尺度目标时表现更加优异。这种方法提供了新的技术支持和思路，为粉尘检测在工业环境中实现自动化和智能化提供理论依据。

关键词 YOLOv8；目标检测；粉尘识别；深度学习

doi：10.3969/j.issn.1672-9528.2024.12.047

0 引言

钢铁产业作为国民经济的重要基础产业，在经济发展进程中扮演着举足轻重的角色，发挥着关键作用，其影响力广泛而深远，辐射至众多领域在基础设施建设领域，无论是纵横交错的铁路网线、车水马龙的公路干道，还是横跨天堑的桥梁、繁忙高效的机场以及吞吐量大的港口，无一不高度依赖钢材的大规模供应。正因此，钢铁产业的稳健前行，为这些大型基础设施项目筑牢根基，注入强劲动力，强有力地推动着国家现代化建设。

在钢铁厂的生产过程中，大宗原材料是基本的生产保证。而作为主要原料之一的铁矿石，在用于钢铁生产之前，通常需要经过一系列的加工步骤，其中包括将铁矿石破碎和磨成粉末。在相对封闭的物料堆场中，散装原料和工业废渣的装卸、转运会产生大量粉尘。解决粉尘扩散的有效方法是喷水降尘，对已产生的粉尘或堆料进行喷水，达到一定湿润度后表层形成硬壳，从而抑制内部粉尘的飞扬。料场大棚内的喷雾抑尘方式是在料堆大棚内部，采用雾炮机与粉尘浓度监测仪联动治理，当粉尘浓度监测仪检测到料棚内的粉尘浓度达到预设阈值上限时，则启动雾炮机进行治理；当粉尘浓度监测仪检测到料棚内的粉尘浓度低于预设阈值下限，则停止抑尘设备作业；这种联动治理方式通常在粉尘产生、扩散至整个料棚内并达到一定浓度才进行治理，其治理具有明显的滞后性。因此需要一种高效且准确的粉尘识别技术，在产尘源头及时发现并进行精准抑尘，减少粉尘扩散，降低治理成本，能够显著实现节能减排效果。基于上述情况，利用深度学习技术对粉尘进行识别的研究具有实际意义。

近年来，人工智能中的深度学习技术进步明显，特别是在目标检测领域。目标检测是一项计算机视觉中旨在识别图像或视频中目标对象所处的位置（通过边界框）的关键任务，并将目标对象识别出来。目标检测技术在深度学习快速发展的同时也有了显著的提高。

从传统的计算机视觉方法到现代的基于深度学习的方法，目标检测技术经历了显著的进化。早期主要技术包括：HaarLike 特征 +AdaBoost（adaptive boosting）[1]、HOG（histogram of oriented gradients）[2]+SVM（support vector machine）[3]、DMP（dynamic movement primitives）[4] 等，这些传统方法依赖于手工设计特征，容易受到光照、视角、尺度等因素的影响，泛化能力不足，侦测速度较慢，难以适应现实复杂场景。

1. 河北科技师范学院数学与信息科技学院 河北秦皇岛 066004

在以卷积神经网络为基础的目标检测中，主要分为两个阶段和一个阶段的探测器。以 RCNN（region-based convolutional neural network）[5] 为代表的 TwoStage 探测器首先产生候选区域，通过有选择的搜索，然后对各个候选区域进行特征提取和分类。虽然 RCNN 提升了检测精度，但计算效率较低，需对每个候选框单独运行 CNN（convolutional neural network）[6]。而以 YOLO（you only look once）[7] 为代表的 OneStage 探测器，则首次将目标侦测任务视为单一回归问题，直接透过侦测速度极快、适用于实时应用的单一网络，对物体类别、边框位置进行预测。

YOLOv8[8] 相对于最初的 YOLO 版本进行了多方面的改进，使其在精度、速度和适应性上有了显著的提升。YOLOv8 虽然是一个相当复杂的结构，但各部分的职责相对清晰，更容易根据实际应用场景进行改进，以适应不同尺度和复杂背景条件下的粉尘检测挑战。本研究的目标，是在保持模型轻量化特点的前提下，优化其对外轮廓并不明显的粉尘的特征提取能力，提高模型的抗干扰能力。

1 YOLOv8 网络结构

YOLO 系列模型由 Joseph Redmon 在 2015 年首次提出，最初的目标是通过一个简单的神经网络结构实现目标检测。随着版本的更新，YOLO 系列逐渐成为业界经典的实时检测方法之一。YOLOv8 是该系列的较新版本，发布于 2023 年，由 Ultralytics 开发。

YOLOv8 的网络结构在 Backbone（主干网络）、Neck（颈部网络）和 Head（检测头）三大模块上进行了全面的优化设计，特别注重提取与融合多尺度特征。

YOLOv8 的 Backbone 负责从输入图像中提取特征。相比早期版本，YOLOv8 使用了更深的卷积层和现代网络设计，结合了 CSPNet（cross stage partial network）[9] 的思想，确保信息流高效传递并减少冗余计算。

YOLOv8 的 Neck 负责整合不同尺度的特征图，实现多尺度检测。YOLOv8 采用了 FPN（feature pyramid network）[10] 和 PANet（path aggregation network）[11] 的结构，保证探测多尺度物体的能力。

YOLOv8 的 Head 负责最终的目标定位与分类预测。相较于 YOLOv1，YOLOv8 的检测头采用了更灵活的设计，通过 Anchor-Free[12] 机制，直接预测目标的中心位置和边界框，不依赖于预设的锚框。

YOLOv8 的工作流程可以分为输入预处理、特征提取、特征融合、目标检测和后处理五个主要阶段。从输入图像到目标检测输出，这些步骤共同完成了的整个过程。总结如下：

（1）输入图像经过预处理，缩放、归一化以及数据增强（训练时），并调整为模型输入的格式。

（2）Backbone 网络负责对特征进行提取并生成多尺度特征图。

（3）Neck 网络融合了多尺度特征，以确保检测大目标和小目标的能力。

（4）Head 网络进行目标检测和分类，生成目标的位置和类别预测，使用 Anchor-Free 机制提高检测精度。

（5）后处理步骤过滤掉低置信度的检测结果，并通过 NMS 去除掉重复的边界框，最后输出检测结果。

YOLOv8 的网络结构图，如图 1 所示。

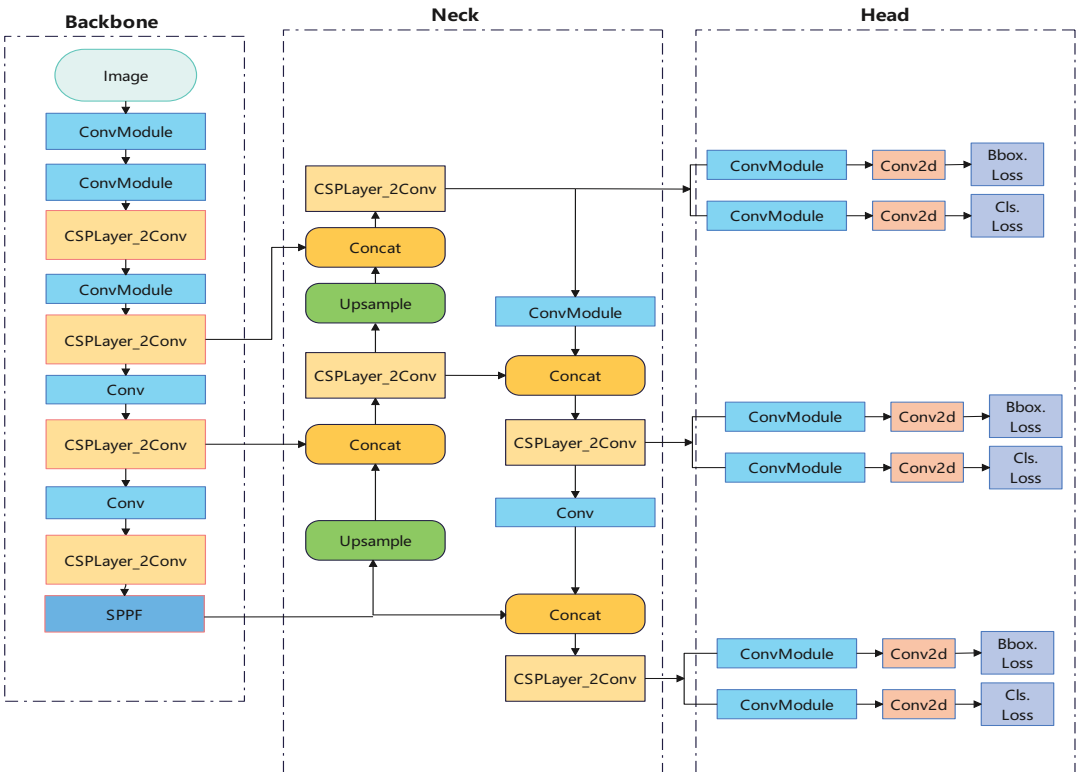

图 1 YOLOv8 的网络结构图

2 模型优化与改进

2.1 上采样优化 DySample

粉尘通常呈现出细微、弥散和不规则的形态，尤其是在图像的边缘区域，这些细微的特征容易因为规则的固定而模糊化，无法像传统的双线性插值和最近的邻接插值那样准确地捕捉烟尘的边界和细节。

为了提高图像中边缘区域的采样质量，确保在粉尘密集区域和边缘处获得清晰的特征图，有助于更精确地识别和定位，本文引入了 DySample[13]，一种轻量级、高效且动态的上采样方法，通过动态生成权重来自适应调整采样过程。

DySample 主要基于动态采样权重的计算。设输入特征图为 $\boldsymbol{F}$，输出的上采样特征图为 $\boldsymbol{F}_{\text{up}}$，对于特征图中的每个位置 (i,j)，DySample 根据以下公式动态计算该位置的采样结果来生成动态权重，设 $\boldsymbol{W}(i,j)$ 为位置 (i,j) 的动态采样权重，该权重通过轻量级网络 G 生成：

$$\boldsymbol{W}(i,j) = G(\boldsymbol{F}(i,j)) \tag{1}$$

式中：G 是一个轻量级的卷积网络，它根据输入特征 $\boldsymbol{F}(i,j)$ 动态生成每个像素位置的权重。

动态采样，DySample 的上采样过程通过加权和邻域像素点进行：

$$\boldsymbol{F}_{\text{up}}(i,j) = \sum_{k\in N(i,j)} W_k(i,j)\cdot \boldsymbol{F}(k) \tag{2}$$

式中：$N(i,j)$ 表示位置 (i,j) 的采样邻域，$W_k(i,j)$ 是第 k 个邻域位置的动态权重。

特征融合，最终上采样特征图通过将输入特征图和动态采样结果结合，实现高质量的特征提升。

避免了复杂性和高成本的同时，又结合了动态上采样的优点。DySample 采用简单而高效的方法来生成内容感知的上采样结果，无需额外的高分辨率特征输入。这使得 DySample 在保持高性能的同时，还降低了模型的复杂性和计算成本。

2.2 标签分配策略

在目标检测模型中，输入图像经过特征提取网络得到一系列特征图，这些特征图与预定义的 Anchor 框结合，用来预测目标的类别和位置。标签分配在目标检测模型中具有决定性作用。

标签分配是指：设输入特征图为 $\boldsymbol{F}$，Anchor 集合为 $A=\{a_i\}_{i=1}^{N}$，其中每个 Anchor 框 a_i 具有其坐标和尺寸。目标是为每个 Anchor 框分配正负标签，即确定哪些 Anchor 是正样本，哪些是负样本。不同的标签分配策略会显著影响检测模型的性能表现。

对于基于 Anchor 的检测模型，Anchor 和对应 Ground Truth 之间的 IoU 阈值是关键因素，因为正负样本的划分取决于这个阈值。而本文采用的算法基于 IoU 与 Ground Truth 的分布，自适应地生成阈值。

为了更好地关注到复杂背景下的粉尘，使用了一种简单且有效的标签分配方法，Dynamic ATSS 是 ATSS（adaptive training sample selection）的动态版本，该方法根据模型在训练和测试阶段的状态动态调整标签分配。ATSS 中的关键部分是通过计算每个 Ground Truth g_j 与其所有对应 Anchor 框的 IoU 值的分布，动态生成正负样本的自适应 IoU 阈值。对于 Ground Truth g_j，计算与其相关的所有 Anchor 框的 IoU 值集合 $\{\text{IoU}(a_i,g_j)\}$，并通过以下公式得到一个自适应的阈值 $\tau(g_j)$：

$$\tau(g_j) = \mu(\{\text{IoU}(a_i,g_j)\}) + \lambda\cdot\sigma(\{\text{IoU}(a_i,g_j)\}) \tag{3}$$

式中：$\mu(\cdot)$ 表示 IoU 的均值；$\sigma(\cdot)$ 表示 IoU 的标准差；λ 是一个超参数，控制 IoU 阈值的动态调整幅度。此阈值 $\tau(g_j)$ 确定了哪些 Anchor 框可以被认为是正样本。在 ATSS 中，IoU 值大于或等于自适应阈值 $\tau(g_j)$ 且 Anchor 的中心点落在 Ground Truth 框的范围内， 数学表示为：

$$a_i \in \text{Positive Samples if IoU}(a_i,g_j) \geq \tau(g_j) \text{ and } a_i \text{ center inside } g_j \tag{4}$$

Dynamic ATSS 相较于传统 ATSS 的核心改进在于引入了模型预测信息。即，除 IoU 阈值外，还使用模型的预测分数来动态调整正负样本的选择。假设 Anchor a_i 的分类分数为 $s(a_i)$，那么可以为每个 Ground Truth g_j 生成一个自适应的分数阈值 $\tau_s(g_j)$。

在 Dynamic ATSS 中，正样本的选择条件不仅取决于 IoU，还要满足：

$$a_i \in \text{Positive Samples if IoU}(a_i,g_j) \geq \tau(g_j) \text{ and } s(a_i) \geq \tau_s(g_j) \tag{5}$$

式中：$\tau(g_j)$ 是自适应 IoU 阈值，$\tau_s(g_j)$ 是分类分数的阈值。

这种机制确保选择具有较高 IoU 和分类分数的高质量 Anchor 框作为正样本，有效减少低质量样本的干扰。通过引入预测结果，选择与 Ground Truth 拥有较高 IoU 的高质量样本作为正样本，这样可以减少分类得分与 IoU 之间的差异，从而生成更多的高质量边界框。

另外通过自适应标签分配算法，动态选择高质量正样本，不仅提升了模型的检测性能，还减少了这些正样本的边界框损失。对于正样本，边界框损失通常采用 L1 损失或 GloU 损失来优化 Anchor 框的定位：

L1 损失公式：

$$\mathcal{L}_{l1} = \sum_{a_i\in\text{Positive Samples}} \| b_i - g_j \|_1 \tag{6}$$

GloU 损失公式：

$$\mathcal{L}_{\text{GloU}} = 1 - \frac{\text{Area of}(a_i\cap g_j)}{\text{Area of}(a_i\cup g_j)} + \frac{\text{Area of the smallest enclosing box}}{\text{Area of}(a_i\cup g_j)} \tag{7}$$

通过正样本的自适应分配，Dynamic ATSS 能确保高质量 Anchor 框在训练中占据主导地位，从而提高边界框预测的精度。

2.3 改进后的网络结构

在原先 Yolov8 的 Neck 网络中的 UpSample 上采样方法替换为 DySample，又使用 Dynamic ATSS，改进了 Head 网络中标签分配策略。优化后的网络结构如图 2 所示。

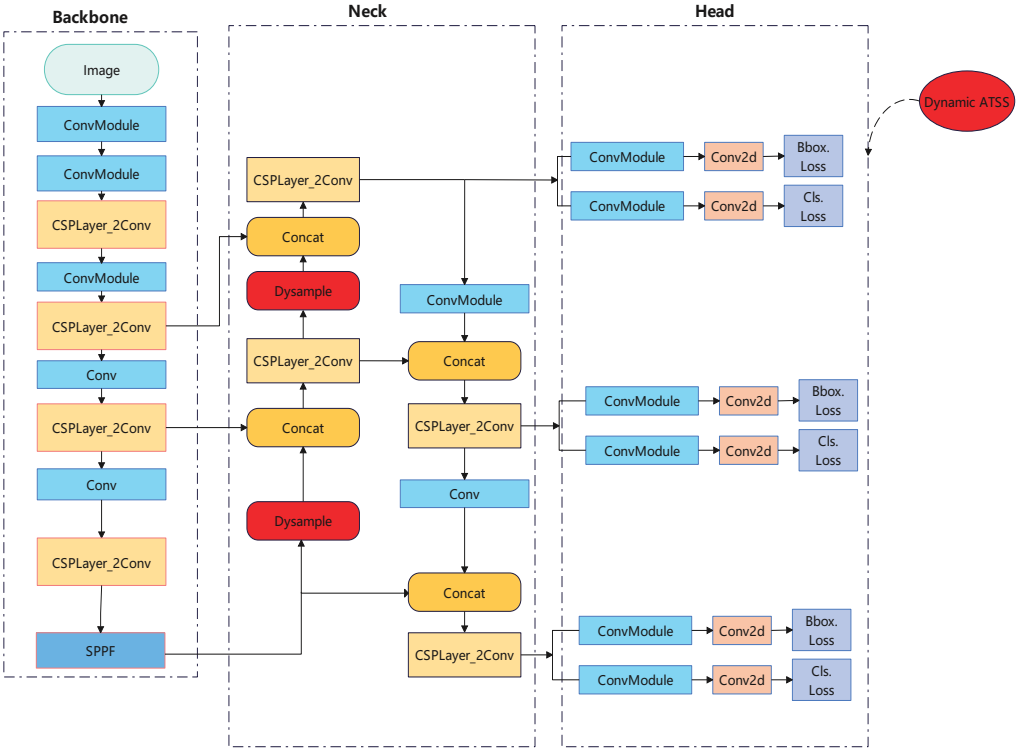

图 2 优化后的网络结构图

经实验证明，优化后的网络，增强了特征提取能力，提高了细节恢复能力，DySample 动态上采样方法能够根据输入特征自适应地调整采样权重，避免了传统上采样带来的细节丢失问题。提高了正样本选择的质量，Dynamic ATSS 通过自适应调整 IoU 阈值，选择与 Ground Truth 高 IoU 的正样本，提升模型在训练过程中正负样本的划分准确性，有助于减少误检和漏检，生成了更高质量的边界框。

3 实验及结果分析

3.1 实验环境

本文的训练平台配置如表 1 所示。

表 1 训练平台配置表

设备名称	LAPTOP-C2FRGPJ6
操作系统	Windows11
处理器	Intel(R) Core(TM) i9-14900 HX 2.20 GHz
基带 RAM	32.0 GB
显存	NVIDIA GeForce RTX4060
硬盘容量	1 TB
编程环境	Python3.8.10
GPU 加速	CUDA12.4

3.2 数据集及参数设置

本文采集了某钢铁厂矿粉装卸、转运时的监控视频数据。通过写程序将视频数据转换为单帧图片，然后利用帧差分法缩减图片数量，最终通过人工筛选出带有粉尘的图片。从中筛选出 1200 张图片，按照 8:1:1 的比例划分训练集、验证集、测试集。使用 LabelImg 软件对采集的数据进行人工标注，将图像中的粉尘位置标注出来，保存为 YOLO 需要的 TXT 格式。实验中，输入图像 imgsz 为 640，epochs 迭代次数 300，初始学习率 lr 等于 0.002，动量 momentum 等于 0.9，训练的批次大小 batch 为 8。

3.3 评价指标

本文使用精确度、召回率和 mAP50，mAP50-95 作为检测模型性能的评价指标。

精确度（Precision）是指在所有被模型预测为正样本的检测结果中，真正为正样本的比例，其公式为：

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)}+\text{False Positives (FP)}} \quad (8)$$

式中：True Positives（TP）指模型正确检测出的目标；False Positives（FP）指模型错误检测为目标的无关区域（即误报）。精确度主要反映模型的准确性，即模型预测的目标中有多少是真正的目标。

召回率（Recall）是指在所有实际为正样本的目标中，模型能够正确检测出来的比例，其公式为：

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)}+\text{False Negatives (FN)}} \quad (9)$$

False Negatives（FN）指模型未能检测出的实际目标（即漏检）。召回率反映了模型的全面性，即模型在所有真实存在的目标中检测出了多少。

mAP 是指平均精度（AP）的均值。AP 是通过计算不同阈值下的 Precision 和 Recall 关系曲线，并计算该曲线下面积得到的值，其定义为：

$$\text{AP} = \int_0^1 P(r)\,\mathrm{d}r \quad (10)$$

式中：$P(r)$ 是在召回率 r 处的精确度。一般通过不连续的点

来计算这个积分，即在不同召回率点的精确度做平均。

mAP（mean average precision）对于多类别的检测任务，mAP 是所有类别 AP 的均值：

$$\mathrm{mAP} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{AP}_i \tag{11}$$

式中：N 是类别数；AP_i 是类别 i 的平均精度。

mAP50 是在 IoU ≥ 0.5 h 的平均精度 (AP) 的均值，其公式为：

$$\mathrm{mAP50} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{AP}_i \quad \text{at IoU} = 0.5 \tag{12}$$

mAP50-95 是在 IoU 阈值步长为 0.05 的平均精度均值，区间为 0.5~0.95，其公式为：

$$\mathrm{mAP50\text{-}95} = \frac{1}{10}\sum_{\mathrm{IoU}=0.5}^{0.95} \mathrm{mAP}_{\mathrm{IoU}} \tag{13}$$

mAP 是衡量目标检测模型整体性能的一个重要指标，反映了模型在不同类别和阈值下的检测精度。mAP 的大小代表模型在准确识别和定位目标方面的综合表现。

3.4 实验分析

为了更好地比较各个模块对实验结果的影响，本文通过消融实验进行比较，分别将改进点加入 YOLOv8 的网络结构中，进行相互对比实验，其结果如表 2 所示。

表 2 对比实验结果

单位：%

算法	Precious	Recall	mAP50	mAP50-95
YOLOv8	80.2	83.1	85.2	71.6
YOLOv8+Dysample	80.6	84.4	86.3	75.3
YOLOv8+Dynamic ATSS	88.5	86,2	86.8	76.2
YOLOv8+Dysample+ Dynamic ATSS	90.3	86.4	**89.8**	80.3

作为基准模型，YOLOv8 在精度、召回率以及 mAP50 和 mAP50-95 上都表现出了较高的性能。它是一个性能优越的目标检测模型，适合实时检测任务。Precious 为 80.2%，Recall 为 83.1%，代表了较好的分类和定位能力。mAP50 和 mAP50-95 分别为 85.2% 和 71.6%，表现相对均衡。尽管基础性能不错，但在复杂场景、小目标和边界不清晰的物体上，可能表现稍逊，尤其是在更高要求的精度和广泛的 IoU 范围内（mAP50-95 较低）。DySample 提供了一种更精确的上采样方法，特别适用于处理小目标和模糊边界的目标。相比基准模型，Recall 提升到 84.4%，mAP50 和 mAP50-95 分别提升到 86.3% 和 75.3%，特别是细化了边界预测，提升了对小目标的检测效果。尽管上采样精度提高，但对整体分类能力（Precious）的提升较小（仅从 80.2% 提升到 80.6%），需要与其他优化策略结合以进一步提升精度。Dynamic ATSS 提供了自适应的标签分配策略，大幅提升了正样本选择的合理性，尤其对分类准确度有明显提升。Precious 显著提升到 88.5%，表现出强大的分类能力，召回率也保持在 86.2% 的较高水平。mAP50 和 mAP50-95 均有所提升，分别达到 86.8% 和 76.2%，表明在更高 IoU 阈值下模型的表现更加稳健。动态标签分配策略虽然有效，但对提升上采样细节的贡献有限，需要结合其他机制进一步优化检测边界。结合 DySample 和 Dynamic ATSS 后，该模型在所有指标上表现最优。Precious 达到 90.3%，是所有方案中最高的，表明其分类能力极强。Recall 达到 86.4%，说明该模型可以捕捉到更多的目标。mAP50 和 mAP50-95 分别提升至 89.8% 和 80.3%，这表明无论在较低或较高的 IoU 阈值下，模型均表现优异。这组合的优点在于：DySample 通过更精确的上采样提升边界处理能力、Dynamic ATSS 通过动态分配标签优化正负样本选择。

同时为了直观地感受到实际模型在检测过程中的改进效果，选取了测试集中的部分图片执行预测，然后进行模型改进前后对比，对比图如图 3 所示。

图 3 模型改进前后对比图

图 3 中 (a) 和 (b) 分别对应了 YOLOv8 原始网络和本文改进的 YOLOv8 网络，可以看出图 (a) 第一张图中，对于明显的粉尘团识别置信度较低，仅为 0.59，这样对于更小的粉尘目标更容易造成漏检，而经过改进后的模型，对于同一张图片进行预测后，置信度提升至 0.85。在其他的预测结果中，改进之后的网络结构在粉尘图像的识别能力上均有不同程度的提升，因此改进后的网络在感知细节和特征表达能力方面更加强大，能够更准确地捕捉到粉尘的形状和特征，检测精度提升明显，模型性能更佳。

3.5 主流算法性能对比

为了进一步证明本文算法的有效改进，本文同时将其与其他经典网络模型进行横向比较，比较结果如表 3 所示。

表 3 主流算法性能对比

单位：%

算法	Precious	Recall	mAP50	mAP50-95
Faster R-CNN	86.50	83.20	85.10	69.70
SSD	81.80	79.40	80.20	63.50
YOLOv3	82.30	78.90	82.70	65.80
YOLOv5	87.20	84.10	86.90	71.00
Ours	90.3	86.4	**89.8**	80.3

由表3可以看出，相比于经典Faster R-CNN、SSD等网络，改进后的YOLOv8网络在精度和召回率方法具有很大的检测优势，模式识别准确率更加高效，满足对粉尘进行实时有效的检测的要求的同时，实现了高精度的检测。

4 结语

因为由于粉尘作为检测目标具有无规则形状、目标大小不一、颜色与背景容易混淆等难以捕捉的特点，本文提出了一种基于改进YOLOv8基础模型的工业环境中粉尘检测算法。添加了DySample提升了特征上采样的精度和细节保留，而添加Dynamic ATSS后改进了标签分配的合理性，使得模型在各种复杂场景中的检测性能显著提升。上述改进方法在实时的工业环境粉尘检测中具有一定的实用价值，为工业环境中粉尘检测的自动化与智能化提供了新的思路和技术支持，具有一定的参考意义。

参考文献：

[1] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features[C/OL]//Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway:IEEE, 2001[2024-06-19].https://ieeexplore.ieee.org/document/990517.

[2] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C/OL]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).Piscataway:IEEE, 2005[2024-02-19].https://ieeexplore.ieee.org/abstract/document/1467360.

[3] CORTES C, VAPNIK V. Support-vector networks[J]. Machine learning, 1995, 20(9): 273-297.

[4] IJSPEERT A J, NAKANISHI J, SCHAAL S. Learning attractor landscapes for learning motor primitives[C]//Advances in Neural Information Processing Systems (NeurIPS). MA: MIT Press, 2002:1547-1554.

[5] 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 (CVPR). NewYork: ACM, 2014: 580-587.

[6] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015(5): 436-444.

[7] 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). Piscataway:IEEE, 2016: 779-788.

[8] Ultralytics. YOLOv8: State-of-the-art object detection model[EB/OL]. (2024-02-12)[2024-07-23]. https://github.com/ultralytics/ultralytics.

[9] WANG C Y, WU Y H, CHEN P Y, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C/OL]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).Piscataway: IEEE, 2020[2024-04-22].https://ieeexplore.ieee.org/document/9150780.

[10] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C/OL]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).Piscataway: IEEE, 2017[2024-05-11].https://ieeexplore.ieee.org/document/8099589.

[11] LIU S, LU Q, QIN H F, et al. Path aggregation network for instance segmentation[C/OL]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition.Piscataway: IEEE, 2018[2024-04-24].https://ieeexplore.ieee.org/document/8579011.

[12] LUO X, DENG J. CornerNet: detecting objects as paired keypoints[J]. Lecture notes in computer science, 2019, 11218(10): 165-781.

[13] LIU W Z, FU H T, CAO Z G, et al. Learning to upsample by learning to sample[C/OL]//2023 IEEE/CVF International Conference on Computer Vision (ICCV).Piscataway: IEEE, 2023[2024-1-09].https://ieeexplore.ieee.org/document/10377871.

【作者简介】

苏宁（1986—），男，安徽宿州人，硕士研究生，研究方向：计算机视觉、图像识别。

冀俊豹（2001—），男，河北邯郸人，硕士研究生，研究方向：智能系统、智慧农业。

贺伟夕（1999—），女，河北石家庄人，硕士研究生，研究方向：计算机视觉、智慧农业与遥感影像处理。

赵立强（1968—），通信作者（email: zhao_liqiang@126.com），男，河北秦皇岛人，博士，教授、硕士生导师，研究方向：机器视觉、模式识别、信息认证与安全。

（收稿日期：2024-09-20）