

基于不可分小波和改进YOLOv8的交通标志检测算法

范阳旭

湖北大学计算机与信息工程学院, 湖北 武汉

收稿日期: 2024年3月19日; 录用日期: 2024年4月9日; 发布日期: 2024年4月18日

摘要

在现代智能交通系统中, 高效且准确的交通标志检测对于辅助驾驶和自动驾驶系统具有重要意义。针对背景复杂的道路场景中交通标志尺寸小导致的识别精度低, 漏检等问题, 提出了一种基于不可分小波和改进YOLOv8的交通标志检测算法。首先, 采用不可分小波处理输入图像, 有效提取高频信息以增强图像的细节表现, 提高模型的鲁棒性。其次, 引入针对小目标的检测层, 取代原始模型中的大目标检测层, 优化网络结构, 从而显著提升了小目标的检测性能。接着, 将网络中的跨步卷积替换成SPD-Conv, 有效减少特征信息的丢失。最后, 采用WIoU损失函数代替原有的损失函数。在TT100K数据集上进行训练, 实验结果显示, 改进后的算法相较于YOLOv8在精确率及mAP@0.5上, 分别提升了9.7%和11.5%, 性能明显优于原始算法。

关键词

交通标志检测, 不可分小波, YOLOv8, 小目标检测, SPD-Conv

Traffic Sign Detection Algorithm Based on Non-Separable Wavelets and Improved YOLOv8

Yangxu Fan

School of Computer and Information Engineering, Hubei University, Wuhan Hubei

Received: Mar. 19th, 2024; accepted: Apr. 9th, 2024; published: Apr. 18th, 2024

Abstract

In modern intelligent transportation systems, efficient and accurate traffic sign detection is of

great significance for assisted driving and autonomous driving systems. Aiming at the problems of low recognition accuracy and missed detection caused by the small size of traffic signs in road scenes with complex backgrounds, a traffic sign detection algorithm based on indivisible wavelets and improved YOLOv8 is proposed. Firstly, the input image is processed with Non-Separable wavelets to effectively extract high-frequency information in order to enhance the detail performance of the image and improve the robustness of the model. Second, a detection layer for small targets is introduced to replace the large target detection layer in the original model, and the network structure is optimized, thus significantly improving the detection performance of small targets. Next, the stepwise convolution in the network is replaced with SPD-Conv to effectively reduce the loss of feature information. Finally, the WIoU loss function is used instead of the original loss function. Trained on the TT100K dataset, the experimental results show that the improved algorithm improves 9.7% and 11.5% compared to YOLOv8 in terms of accuracy and mAP@0.5, respectively, and the performance is significantly better than the original algorithm.

Keywords

Traffic Sign Detection, Non-Separable Wavelet, YOLOv8, Small Object Detection, SPD-Conv

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

交通标志对于维护交通秩序和保障驾驶安全具有重要作用，而交通标志检测作为辅助驾驶系统的重要组成部分，能够为驾驶者提供重要的交通信息和指示，帮助驾驶员做出更为准确和迅速的决策，从而减少交通事故的发生。此外，随着无人驾驶技术的发展，交通标志检测对于实现完全自动驾驶具有重大意义。

交通标志检测算法可以分为传统和基于深度学习的两类。传统方法通过颜色、形状和边缘等特征进行识别，需人工提取特征，而且易受天气和光照等外部因素影响，因此，难以满足交通标志检测实时、准确的实际需求。相较而言，基于深度学习的算法在特征提取上优势明显，具备更强的泛化能力和检测性能，已成为交通标志检测的主流方法。其主要分为二阶段和单阶段两大类，二阶段算法如 R-CNN [1]、Fast R-CNN [2]、Faster R-CNN [3]等，通过区域生成网络来生成目标候选框，然后对目标候选框进行分类和回归。虽然其具有较高的检测精度，但检测速度较慢。单阶段算法代表如 SSD [4]和 YOLO [5] [6] [7] [8]系列，它可以直接从网络中提取特征来预测目标分类和位置，进行端到端检测，特点是速度快但精度略有不足。

近些年来，交通标志检测研究领域已取得显著成果。Wu 等人[9]在 YOLOv3 中引入了 DarkNet19 作为分类网络，提出了一种基于交通标志特征的数据增强方法，提高了检测性能，但对小型交通标志的检测效果较差。Zhang 等人[10]将空间金字塔池化(SPP)模块融合入 YOLOv3 并结合图像增强技术，利用底层细粒度特征，以精确定位目标。Yao 等人[11]针对实时交通标志检测，改进了 YOLOv4-Tiny 算法，引入自适应特征金字塔网络(AFPN)来融合不同尺度的特征层。在骨干网络中引入感受野块(RFB)，强化了两个特征层的特征提取效率，显著提升了交通标志的检测性能。Wang 等人[12]通过在 yolov5 中引入注意力模块和特征增强模块，提高了模型对交通标志的感知能力，有效缓解了特征通道减少引发的上下文信息损失。

尽管现有检测算法在交通标志识别领域已取得显著成果，但由于实际应用场景中交通标志距离较远，中小目标占比大，仍面临检测精度不足与漏检的挑战。因此，本文提出一种改进的 YOLOv8 检测算法——SUW-YOLO，旨在保证实时性的同时，显著提高对小目标交通标志的检测精度。本文的主要贡献如下：

1) 随着网络加深，常常导致大量高频信息的损失，故本文提出在预处理阶段引入了不可分小波，通过提取图像的高频信息并将其与原始图像一同输入网络，从而强化图像的细节特征，提高检测性能。

2) 由于交通标志常占的像素低于 32×32 ，为避免特征随网络加深而消失，本文对模型结构进行优化，并用微小目标检测层代替原始的大目标检测层，增强小目标检测能力，同时显著减少了模型参数和计算复杂度。

3) 因为跨步卷积会导致细粒度信息的丢失和低效的特征表示，本文引入 SPD-Convolution 替换掉原模型的跨步卷积，以减少由此产生的特征信息损失，增强低分辨率小目标的特征提取能力。

4) 将原有损失函数 CIoU loss (Complete Intersection over Union) 替换为 WIoU loss (Wise Intersection over Union)，加快模型收敛速度。

2. YOLOv8 网络介绍

YOLO 系列经过不断的优化迭代，YOLOv8 作为 YOLO 系列的最新版本，有着更强的检测性能，可用于目标检测，实例分割，图像分类。基于放缩系数，提供了 YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, YOLOv8x 五个不同大小的模型，来满足不同场景的需求，本文选取体积小，精度较高的 YOLOv8s。

如图 1 所示，YOLOv8 的网络框架主要包括：输入端(Input)、主干网络(Backbone)、颈部网络(Neck)和输出端(head)。输入端使用了 Mosaic 数据增强，自适应的锚框计算，空间扰动等。在主干网络中包含 Conv、C2f、SPPF 等模块，主要负责从输入图像提取特征。其中，YOLOv8 设计了全新的 C2f 结构，通过更多的分支跨层连接，获得丰富的梯度流信息，实现了进一步的轻量化。SPPF 模块是空间金字塔池化，作用是融合多尺度特征，能够将不同尺度的特征图转换成统一尺度。颈部网络主要用于融合来自主干网络的多尺度特征，采用了 PANet (Path Aggregation Network)，使用自顶向下和自底向上的结构，从而精确地保留了空间信息，提高了网络对不同尺度特征的融合能力。

3. YOLOv8 改进策略

针对目前背景复杂的小目标交通标志检测存在的检测精度低，漏检等问题，本文提出一种基于 SUW-YOLO 的交通标志检测模型，模型结构图如图 2 所示。

3.1. 不可分小波

不可分小波的出现，旨在克服传统的二维张量积小波所面临的固有局限性。不同于张量积小波的分立式行和列处理，不可分小波采用不可分小波滤波器直接在多个方向同时进行小波变换，这种方法能够更全面地捕捉图像中的特征。因此，构造不可分小波的关键在于不可分小波滤波器的设计，Chen 等人[13]提出了一种具有高维紧支撑性以及正交性的非张量积小波滤波器，其具体形式为：

$$m_0(\xi) = \frac{1}{s} X(\xi) \prod_{j \in Z_N} (U_j D(A^T \xi)) \left(\prod_{j \in Z_N} U_{N-1-j}^T \right) V_0, \xi \in R^d \quad (1)$$

在这个基础上，Liu 等人[14]提出了不可分小波的构造方法，如果设小波变换的伸缩矩阵 $A = [2, 0; 0, 2]$ ，则滤波器组可表示为：

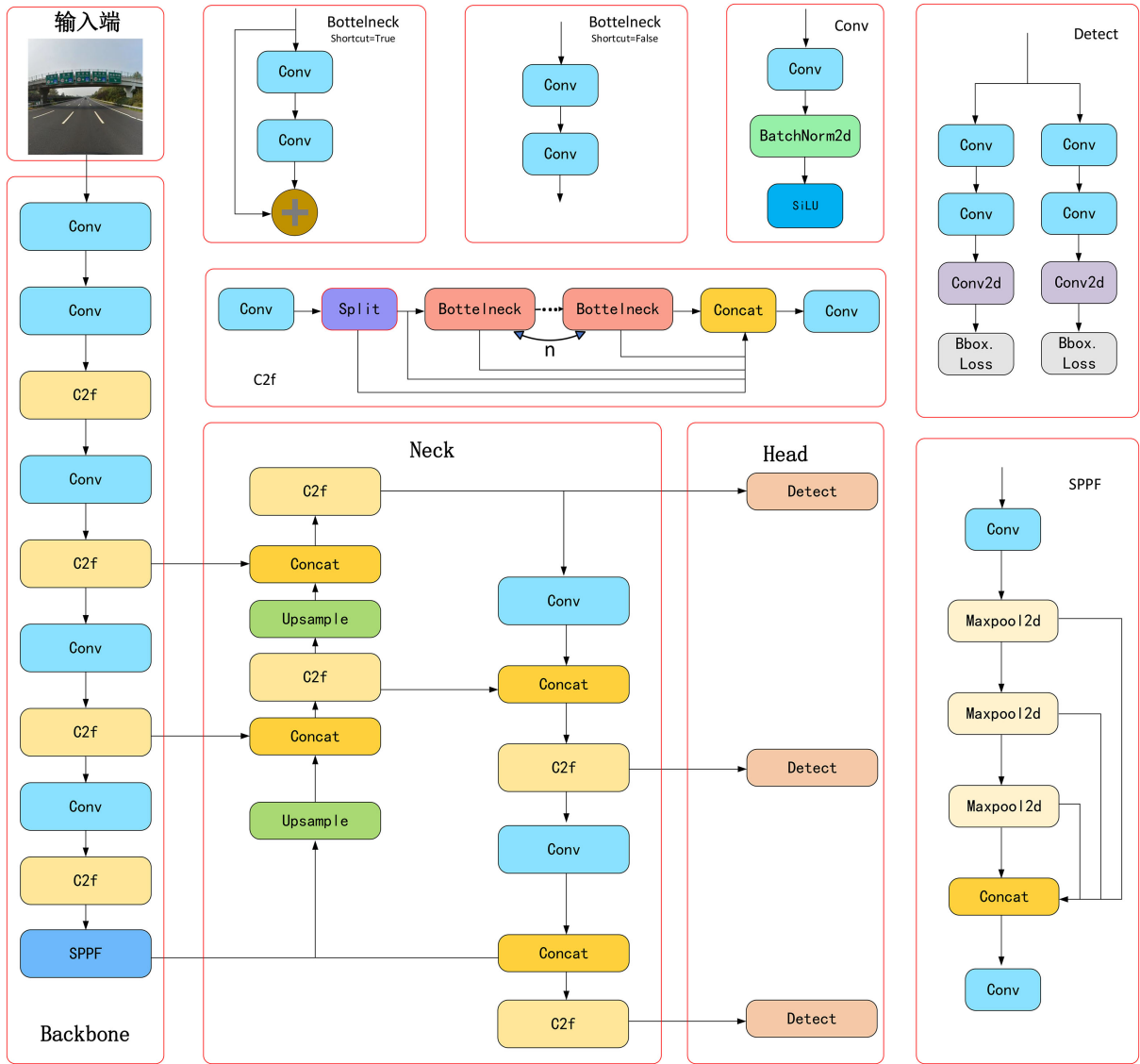


Figure 1. YOLOv8 network architecture diagram
图 1. YOLOv8 网络结构图

$$[m_0(x, y), m_1(x, y), m_2(x, y), m_3(x, y)] = \frac{1}{4} X(1, x, y, xy) \prod_{j=1}^k [U_j D(x^2, y^2) U_j^T V] \quad (2)$$

其中, $x = e^{-iw_1}$, $y = e^{-iw_2}$, $U_j (j=1,2)$ 和 $V/2 = (V_0, V_1, V_2, V_3)/2$ 都是正交阵, V_1, V_2, V_3 都是 4×1 的向量, $V_0 = (1, 1, 1, 1)^T$ 。

为构造四通道且中心对称的 6×6 滤波器组, 取 $K = 2$,

$$U_j = \frac{1}{2} \begin{pmatrix} \cos(a_j) + \cos(b_j) & -\sin(a_j) - \sin(b_j) & \sin(b_j) - \sin(a_j) & \cos(a_j) - \cos(b_j) \\ \sin(a_j) + \sin(b_j) & \cos(a_j) + \cos(b_j) & \cos(a_j) - \cos(b_j) & \sin(a_j) - \sin(b_j) \\ \sin(a_j) - \sin(b_j) & \cos(a_j) - \cos(b_j) & \cos(a_j) + \cos(b_j) & \sin(a_j) + \sin(b_j) \\ \cos(a_j) - \cos(b_j) & \sin(b_j) - \sin(a_j) & -\sin(a_j) - \sin(b_j) & \cos(a_j) + \cos(b_j) \end{pmatrix} (j=1,2) \quad (3)$$

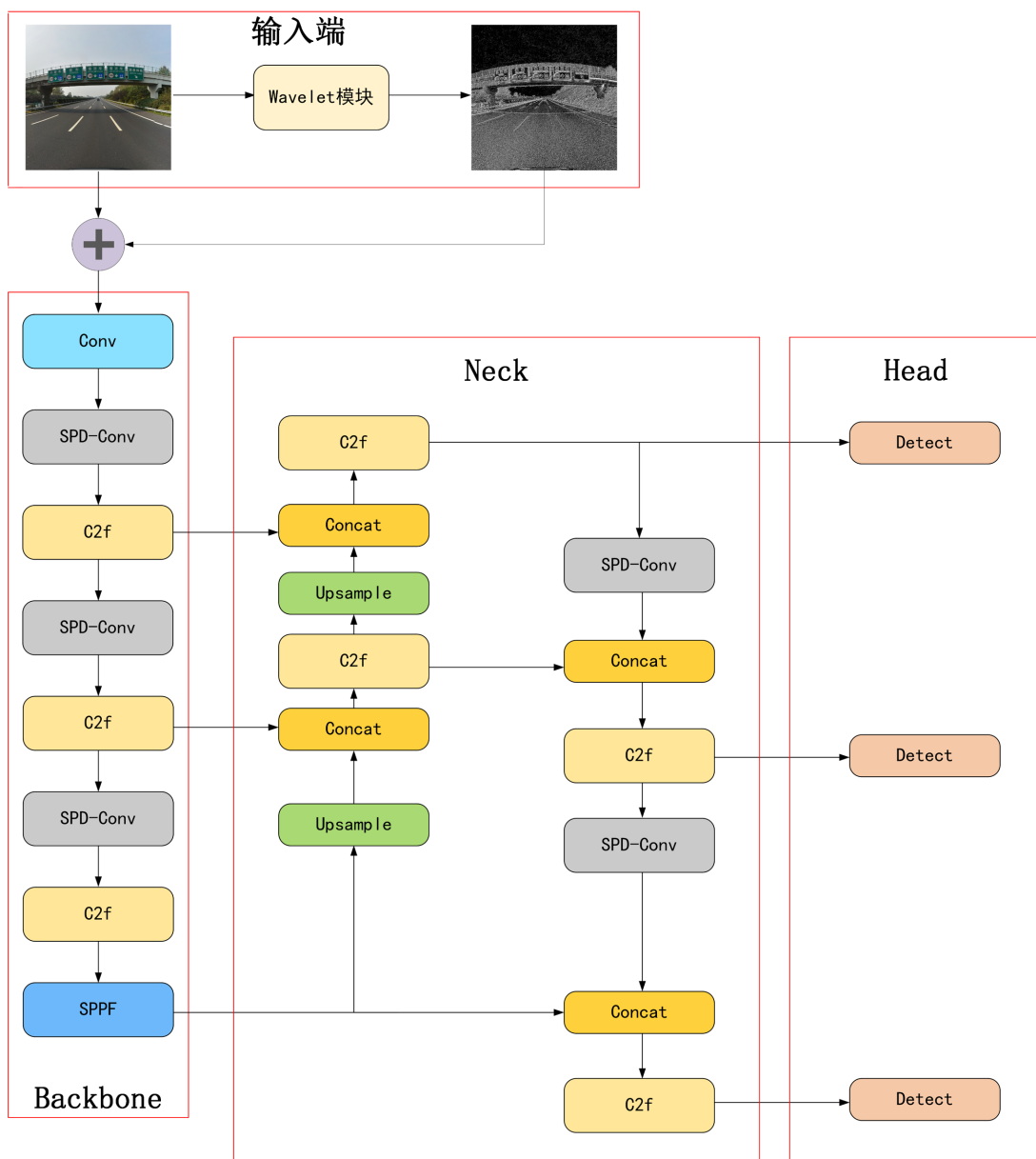


Figure 2. Improve the model structure diagram
图 2. 改进模型结构图

$$V = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix} \quad (4)$$

可验证 $U_j(\alpha_j, \beta_j)$ 为中心对称的正交矩阵，于是我们可以构造一组具备对称性、紧支撑和正交特性的滤波器组 H_0, H_1, H_2, H_3 ，其中， H_0 为低通滤波器，而 H_1, H_2 和 H_3 是三个不同的高通滤波器。然后，我们取 $a_1 = \pi/4$ ， $a_2 = \pi/5$ ， $b_1 = \pi/7$ ， $b_2 = -\pi/3$ ，可得到四通道不可分小波低通滤波器的时域形式：

$$H_0 = \begin{bmatrix} 0.1491 & 0.0253 & 0.2806 & 0.2111 & 0.0304 & 0.1789 \\ 0.1058 & -0.0180 & 0.1585 & -0.1498 & -0.0216 & -0.1269 \\ 0.2532 & 0.0430 & 0.1736 & 0.0135 & -0.0432 & -0.0459 \\ -0.0459 & -0.0431 & 0.0135 & 0.1736 & 0.0430 & 0.2533 \\ -0.1269 & -0.0216 & -0.1497 & 0.1585 & -0.0180 & 0.1058 \\ 0.1789 & 0.0304 & 0.2111 & 0.2806 & 0.0253 & 0.1491 \end{bmatrix} \quad (5)$$

在卷积神经网络(CNN)中,多次下采样和池化操作主要会导致高频信息的损失,牺牲图像的局部细节。针对这个问题,基于加性小波分解原理,我们通过上面构造的低通滤波器对图像进行卷积,从而得到图像的低频子图,再用原图减去低频子图可以获得图像的高频部分,随后,将获得的高频部分与原图像相加,从而得到增强细节信息的图像,作为输入数据输入到YOLOv8网络中进行目标检测。这不仅强化了图像的细节特征,还保留了原始图像的全局信息,从而提高了网络对小目标交通标志的检测准确率。

3.2. 小目标检测层

YOLOv8原始模型的输出端,采用 20×20 , 40×40 , 80×80 三种不同尺度的检测层,分别用来检测大、中、小三种不同尺寸的目标物体,正常来说,深层特征图感受野更大,具备更加丰富的语义信息,但位置信息较少,底层特征图则相反,而交通标志由于拍摄距离较远,一般只占图像很小的一部分, 20×20 的感受野对交通标志来说过大,加上部分特征会随网络加深而丢失,导致YOLOv8的检测效果不佳,故本文添加了一个 160×160 的检测层,替换掉 20×20 的检测层,同时,对网络结构进行优化,裁剪掉骨干网络用于大目标检测的相关卷积层,以减少不必要的计算量。

3.3. SPD-Conv

卷积神经网络(CNN)在计算机视觉任务中,如图像分类和目标检测,都表现的非常出色,但当处理低分辨率图像或小物体时,其性能会显著下降,这很大程度是因为CNN中的跨步卷积和池化层会导致大量的细粒度丢失和低效的特征学习。在参考文献[15]中,作者提出了一种无跨步卷积——SPD-Conv,他是由空间到深度(SPD)和非跨步卷积组成的创新模块,可以代替传统的跨步卷积和池化层。

SPD主要是对特征图进行下采样的同时,把空间特征信息保留到通道维度,避免了特征信息的丢失。再由非跨步卷积层调整通道维度,并确保特征的丰富性。对于尺度为 $S \times S \times C_1$ 的中间特性图 X ,可以裁剪出多个子特征图。原理公式如下(6):

$$\begin{aligned} f_{0,0} &= X[0:S:scale, 0:S:scale], \\ f_{1,0} &= X[1:S:scale, 0:S:scale], \dots, \\ f_{0,1} &= X[0:S:scale, 1:S:scale], \dots, \\ f_{scale-1,0} &= X[scale-1:S:scale, 0:S:scale], \dots, \\ f_{0,scale-1} &= X[0:S:scale, scale-1:S:scale], \dots, \\ f_{scale-1,scale-1} &= X[scale-1:S:scale, scale-1:S:scale] \end{aligned} \quad (6)$$

当比例因子($scale$)是2时,对于宽为 S ,高为 S ,通道为 C_1 的特征图 X 可得到4个子图 $f_{0,0}, f_{0,1}, f_{1,0}, f_{1,1}$ 。将这些子图在通道位图进行拼接,得到新特征图 X' 的通道数是输入特征图的4倍,即 $4C_1$,然后再进行一次非跨步卷积来调整通道维度。如图3所示。总的来说,SPD-Conv解决了传统卷积自带的局限性,对计算机视觉任务是一个更加有效的特征提取方法,特别是处理低分辨率和小目标物图像时。

本研究采用SPD-Conv作为下采样机制,降低下采样过程中的特征信息损失,以提升了在复杂环境下小物体检测的能力,增强了模型的泛化能力和鲁棒性。

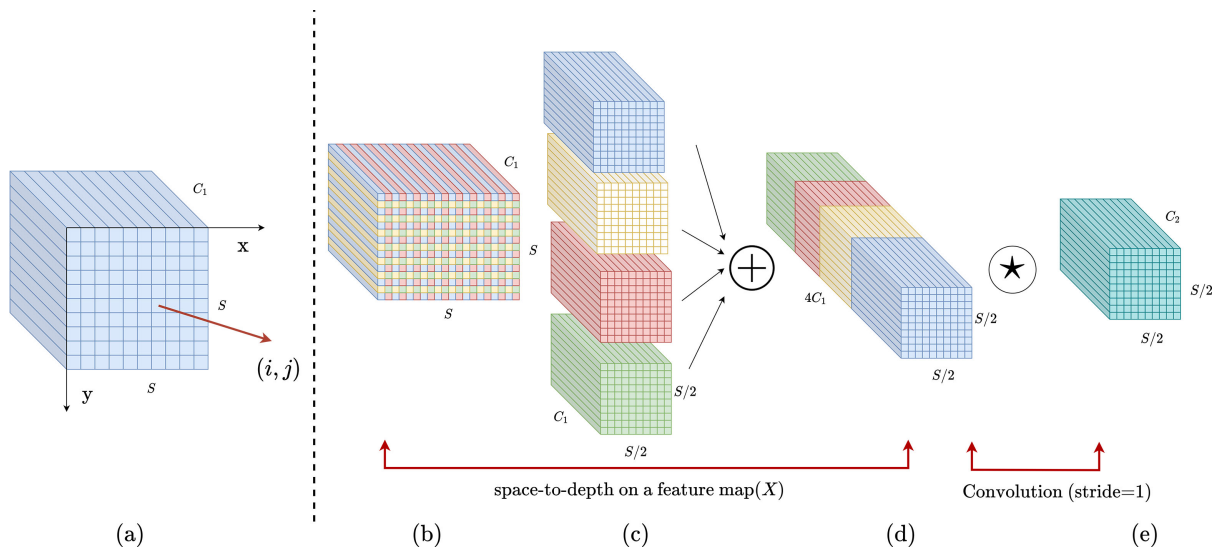


Figure 3. SPD-Conv at scale = 2
图 3. Scale = 2 时 SPD-Conv

3.4. 改进损失函数

在 YOLOv8 的初始架构中，边界框回归损失函数采用 CIoU，其公式详见式(7)：

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{(c_w)^2 + (c_h)^2} + \frac{4}{\pi^2} \left(\tan^{-1} \frac{w^{gt}}{h^{gt}} - \tan^{-1} \frac{w}{h} \right) \quad (7)$$

尽管 CIoU 考虑到了边界框回归的距离、交集面积和中心点偏移等因素，但它没有考虑真实框与预测框之间方向的不匹配，这导致质量差的回归样本对损失函数影响显著，进而减缓模型训练收敛速度并降低预测框精度。Tong 等人[16]提出了 Wise-IoU (WIoU)损失函数，具备动态非单调聚焦机制，它共有三个版本，WIoUv1 设计了基于注意力的预测框损失计算，而 WIoU v2 与 WIoU v3 在其基础上进一步融合了聚焦系数以增强模型性能。本文采用 WIoU v3 优化边界框回归损失，它通过降低高质量锚框的竞争性和减少低质量数据的有害梯度，使模型能够更好地专注于普通质量的锚框，从而提高了模型的整体性能。其表达式如下：

$$L_{WIoU v1} = R_{WIoU} L_{IoU} \quad (8)$$

$$R_{WIoU} = \exp \left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*} \right) \quad (9)$$

$$L_{WIoU v3} = r L_{WIoU v1}, r = \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (10)$$

超参数 α 和 δ 负责调节非单调聚焦系数 r 及离群值 β 。

在交通标志检测应用中，WIoU v3 可以动态优化小目标交通标志的损失权重，优化模型预测框以更精确贴合真实边界，进而增强模型的泛化能力及检测性能。

4. 实验与分析

4.1. 实验数据集

为了验证本文提出模型的有效性，选取由清华大学和腾讯联合实验室整理公开的 TT100K 数据集，

其中包含了 10 万张 5 个中国城市的腾讯街景图，涵盖了不同照明和天气的 120 多类交通标志，但由于某些类别的样本数据分布不均衡，故本文选取了 42 个实例都超过 100 张的交通类别，其中训练集有 6598 张，测试集有 1889 张。

4.2. 评估指标

目标检测的主要测评指标分为模型检测精度和模型复杂度，本实验采用精确率 P (Precision)、召回率 R (Recall)、平均准确率 mAP、每秒帧数(FPS)还有参数来对模型进行评估。

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

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

精确率和召回率公式如上所示，其中 TP 是正确将正样本预测为的样本数量，FP 是错误将负样本预测为正的样本数量，FN 是错误将正样本预测为负的样本数量。

平均精度(AP)是对某一类别的精确率 - 召回率曲线下的面积计算，mAP是所有类别的AP的平均值，可以用来评估模型对所有类别的检测性能，mAP约接近 1，表示模型的检测性能越好。公式如下：其中， n 是类别数； $AP(j)$ 是第 j 个类别的 AP。

$$AP = \int_0^1 P(R) dR \quad (13)$$

$$mAP = \frac{1}{n} \sum_{j=1}^n AP(j) \quad (14)$$

FPS 表示模型每秒可以处理的图像帧数，值越大，实时性越好。而模型参数通常和模型复杂度成正比。

4.3. 实验环境与参数配置

本实验的硬件平台为 Intel(R) Xeon(R) Silver 4210 CPU @ 2.20GHz，操作系统为 window10，GPU 为 16G 显存的 Quadro RTX 5000。实验基于 pytorch 的深度学习框架，开发环境是 python3.8、PyTorch 1.90、CUDA 10.2。初始学习率为 0.01，权重衰减系数为 0.0005，动量设置为 0.637，优化算法是随机梯度下降 (SGD)，输入图像大小为 640×640 ，批量大小设为 16，num_workers 设为 8，close_mosaic 设为 10，即在最后十个 epochs 在训练时关闭 mosaic 以提高训练效果，训练次数为 150 个 epoch，训练时不使用预训练权重。

4.4. 实验结果与分析

4.4.1. 消融实验

为了更好的验证改进模型的有效性，以 YOLOv8s 为基线模型，将 Precision、Recall、mAP@0.5、mAP@0.5:0.95、FPS、Parameters 作为性能评估指标，通过组合不同的模块配置进行消融研究，A 代表引入无跨步卷积，B 代表替换大目标检测层为微小目标检测层，C 是引入 WIoU，D 代表加入不可分加性小波变换，其中，“√”表示集成了该模块，而“×”则表示未包含该模块，结果如下表 1 示：

通过表 1 可知，随着各个模块的加入，模型的精确度、召回率和 mAP 等指标都有不同程度的提升，相较于 YOLOv8s，改进的 SUW-YOLO 算法在精确度上提升至 87.5%，提升了 9.7 个百分点，mAP@0.5 提升至 88.6%，提高了 11.5 个百分点，虽然 FPS 有所下降，但仍然满足目标检测的实时性需求。

Table 1. Ablation experiment**表 1.** 消融实验

A	B	C	D	Precision/%	Recall/%	mAP@0.5/%	mAP@.5:.95/%	FPS	Parameters/M
×	×	×	×	77.8	68.7	77.1	60.4	204	11.15
√	×	×	×	83.1	72.8	82	64.5	161	12.69
√	√	×	×	85.2	77.3	86.9	67.9	123	4.51
√	√	√	×	86.6	78.9	88.1	68.6	125	4.51
√	√	√	√	87.5	80.4	88.6	69.1	120	4.52

4.4.2. 对比实验

为了进一步验证本文改进的 SUW-YOLO 算法在交通标志检测领域相较于现有主流模型的性能优势, 选取 Precision、Recall、mAP@0.5、mAP@0.5:0.95、FPS、Parameters 作为性能评估指标, 在统一配置环境和相同数据集的条件下, 与 Faster R-CNN、YOLOv3、YOLOv5、YOLOv6、YOLOv8 等算法进行了实验对比。

通过表 2 数据可知, 改进后的 SUW-YOLO 算法在性能检测方面全面领先于其他算法, 尽管 FPS 稍有降低, 但模型参数量更少, 同时实现了检测性能的显著提高。

**Figure 4.** Comparison of traffic sign detection effect**图 4.** 交通标志检测效果对比图

Table 2. Comparative experiments of different algorithms in TT100K
表 2. 不同算法在 TT100K 中的对比实验

Models	Precision/%	Recall/%	mAP@0.5/%	mAP@.5:.95/%	FPS	Parameters/M
Faster R-CNN	61.4	51.3	53.1	42.7	14	138.68
YOLOv3-tiny	66	45.8	52.7	40.6	256	12.15
YOLOv5s	70.3	61.1	65.5	51.2	212	9.14
YOLOv6s	65	60.1	65.7	51.1	208	16.31
YOLOv8s	77.8	68.7	77.1	60.4	204	11.15
Ours	87.5	80.4	88.6	69.1	120	4.52

为了更直观地展示本文改进方法的有效性, 将 YOLOv5s, YOLOv8s 以及本文的 SUW-YOLO 算法做可视化对比, 如图 4 所示。通过对比三个模型的检测结果可发现, 在远距离小型交通标志检测方面, 改进后的 SUW-YOLO 算法展现出更高的检测精度和性能。在 A 组中 YOLOv5s 和 YOLOv8s 都将 p130 错检成 p180, 而 SUW-YOLO 网络准确的完成检测。通过 B, C 组可看到, YOLOv5s 和 YOLOv8s 都有不同程度的漏检情况, 而 SUW-YOLO 算法有效地缓解了这个问题, 这进一步证明了本研究算法在复杂环境对远距离小目标交通标志的检测具有显著优势。

5. 结论

本文提出了一种基于不可分小波和改进 YOLOv8 的交通标志检测模型 SUW-YOLO, 可以有效解决复杂环境下小型目标交通标志检测的检测精度低, 漏检等问题。实验结果表明, SPD-Conv 模块的引入可以减少因跨步卷积导致的特征信息损失, 增强对小目标的特征提取能力; 将用微小目标检测层替换大目标检测层, 不仅增强小目标检测性能, 也显著减少了模型参数和计算复杂度; 使用 WIoU 作为损失函数, 增强模型的泛化能力及检测效率; 不可分小波变换的使用则有效提高了模型对细节信息的捕获能力。在 TT100K 数据集中, 改进的 SUW-YOLO 模型相比较于原始模型在精确率及 mAP@0.5 上分别实现了 9.7% 和 11.5% 的显著提升, 大幅提高了网络对复杂路况下小目标交通标志检测性能。鉴于交通标志检测环境的复杂性, 未来工作将致力于探索在雾、雪、及低光照等恶劣天气条件下维持或提升检测效果的方法。

参考文献

- [1] Girshick, R., Donahue, J., Darrell, T., et al. (2014) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 580-587. <https://doi.org/10.1109/CVPR.2014.81>
- [2] Girshick, R. (2015) Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, 7-13 December 2015, 1440-1448. <https://doi.org/10.1109/ICCV.2015.169>
- [3] Ren, S., He, K., Girshick, R., et al. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [4] Liu, W., Anguelov, D., Erhan, D., et al. (2016) Ssd: Single Shot Multibox Detector. In: Leibe, B., Matas, J., Sebe, N. and Welling, M., Eds., *Computer Vision—ECCV 2016, Lecture Notes in Computer Science*, Vol. 9905, Springer, Cham, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [5] Redmon, J., Divvala, S., Girshick, R., et al. (2016) You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [6] Redmon, J. and Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 7263-7271.

<https://doi.org/10.1109/CVPR.2017.690>

- [7] Redmon, J. and Farhadi, A. (2018) Yolov3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.
- [8] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) Yolov4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.
- [9] Wu, Y., Li, Z., Chen, Y., *et al.* (2020) Real-Time Traffic Sign Detection and Classification towards Real Traffic Scene. *Multimedia Tools and Applications*, **79**, 18201-18219. <https://doi.org/10.1007/s11042-020-08722-y>
- [10] Zhang, H., Qin, L., Li, J., *et al.* (2020) Real-Time Detection Method for Small Traffic Signs Based on Yolov3. *Ieee Access*, **8**, 64145-64156. <https://doi.org/10.1109/ACCESS.2020.2984554>
- [11] Yao, Y., Han, L., Du, C., *et al.* (2022) Traffic Sign Detection Algorithm Based on Improved YOLOv4-Tiny. *Signal Processing: Image Communication*, **107**, Article 116783. <https://doi.org/10.1016/j.image.2022.116783>
- [12] Wang, J., Chen, Y., Dong, Z., *et al.* (2023) Improved YOLOv5 Network for Real-Time Multi-Scale Traffic Sign Detection. *Neural Computing and Applications*, **35**, 7853-7865. <https://doi.org/10.1007/s00521-022-08077-5>
- [13] Chen, Q., Micchelli, C.A., Peng, S., *et al.* (2003) Multivariate Filter Banks Having Matrix Factorizations. *SIAM Journal on Matrix Analysis and Applications*, **25**, 517-531. <https://doi.org/10.1137/S0895479802412735>
- [14] 刘斌, 彭嘉雄. 基于四通道不可分加性小波的多光谱图像融合[J]. 计算机学报, 2009, 32(2): 350-356.
- [15] Sunkara, R. and Luo, T. (2022) No More Strided Convolutions or Pooling: A New CNN Building Block for Low-Resolution Images and Small Objects. In: Amini, M.R., Canu, S., Fischer, A., Guns, T., Kralj Novak, P. and Tsoumakas, G., Eds., *Machine Learning and Knowledge Discovery in Databases, Lecture Notes in Computer Science*, Vol. 13715, Springer, Cham, 443-459. https://doi.org/10.1007/978-3-031-26409-2_27
- [16] Tong, Z., Chen, Y., Xu, Z., *et al.* (2023) Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. arXiv preprint arXiv:2301.10051.