文章编号:1007-2780(2026)03-0388-14

YOLO-DyMiF:一种面向低算力平台的动态多尺度交通标志检测网络

宋绍剑*,李　昊,李　刚,李国进
(广西大学 电气工程学院,广西 南宁 530004)

摘要:为了解决自动驾驶场景中交通标志目标体积小、易被环境干扰而导致检测精度低,以及车载平台算力和功耗有限、难以支撑复杂模型的问题,本文提出了一种改进的轻量化检测算法 YOLO-DyMiF(Dynamic Mixer and Feature Fusion)。该模型在 YOLOv10n 的基础上进行了两方面改进:首先,设计一种基于动态高效卷积(Adaptive Efficient Conv, AEConv)的高效动态混合器(Efficient Dynamic Mixer Structure, EDMS),并将其嵌入 C3k2 模块以构建 C3k2_EDMS 模块,用于替换 YOLOv10n 模型中的 C2f 模块,在保持主干网络特征表达能力的前提下有效压缩参数规模;其次,设计了以分层多尺度空间增强模块(Hierarchical Multi-scale Spatial Enhancement, HMSE)为核心的动态特征融合颈部网络,它通过跨层交互和自适应加权融合增强多尺度特征表征能力,在兼顾中、大目标检测性能的同时提升小目标交通标志检测精度。在 TT100K 数据集上的实验结果表明,与当前领先的 Mamba-YOLOt 相比,YOLO-DyMiF 算法的 mAP50 提高 1%,模型参数量下降了 58.3%,计算量下降了 42.3%。所提出的模型能够在确保高检测精度的同时显著降低计算成本,可以为自动驾驶场景中的交通标志检测提供可靠的技术支持。

关　键　词:目标检测;交通标志;自动驾驶;多尺度目标;边缘计算
中图分类号:TP391.4　**文献标识码**:A　**doi**:10.37188/CJLCD.2026-0021　**CSTR**:32172.14.CJLCD.2026-0021

YOLO-DyMiF: a dynamic multi-scale traffic sign detection network for low-computing-power platforms

SONG Shaojian*, LI Hao, LI Gang, LI Guojin

(*School of Electrical Engineering*, *Guangxi University*, *Nanning* 530004, *China*)

Abstract: To address the low detection accuracy caused by the small size of traffic signs in autonomous driving scenarios and their susceptibility to environmental interference, as well as the limited computing capability and power budget of onboard platforms that make complex models difficult to deploy, an improved lightweight detector named YOLO-DyMiF (Dynamic Mixer and Feature Fusion) is proposed. The proposed model, based on YOLOv10n, introduces two major improvements. Firstly, an Efficient Dynamic Mixer Structure (EDMS) based on Adaptive Efficient Convolution (AEConv) is designed and embedded into the C3k2 module to get a new module named C3k2_EDMS, which replaces the C2f module in YOLOv10n. This design effectively reduces the parameter scale while preserving the feature

收稿日期:2026-02-09;**修订日期**:2026-02-20.
基金项目:国家自然科学基金(No. 618630003)
Supported by National Natural Science Foundation of China (No. 618630003)
*通信联系人,E-mail:sjsong03@163.com

representation capability of the backbone network. Secondly, a dynamic feature fusion neck network is developed with the Hierarchical Multi-scale Spatial Enhancement (HMSE) module. Through cross-layer interactions and adaptive weighted fusion, the neck enhances multi-scale feature representation, improving the detection accuracy of small traffic signs while keeping detection performance on medium and large objects. Experimental results on the TT100K dataset show that, in comparison with the state-of-the-art Mamba-YOLOt, YOLO-DyMiF improves mAP50 by 1.0%, reduces the number of parameters by 58.3%, and decreases computational cost by 42.3%. The proposed model significantly reduces the computational cost while ensuring high detection accuracy, which provides reliable technical support for traffic sign detection in autonomous driving applications.

Key words: object detection; traffic signs; autonomous driving; multi-scale objects; edge computing

1　引　　言

交通标志检测是自动驾驶视觉感知中的基础任务之一，直接关系到车辆对限速、禁行、警告与指示信息的理解与执行，从而影响决策规划的安全性与合规性[1]。在高级驾驶辅助系统（Advanced Driver Assistance System，ADAS）与智能交通系统中，需要稳定、快速且准确地检测与识别交通标志[2]。然而，现实的交通环境往往是复杂多变的，给交通标志检测带来了诸多问题[3-5]：(1)动态多形态多尺度目标实时检测任务中小目标检测精度偏低。在真实交通场景中，交通标志易被树木、广告牌、车辆、行人、施工围挡等遮挡，部分老旧标志存在破损、褪色、涂鸦覆盖等问题，导致关键特征丢失。此外，车辆行驶中，交通标志会呈现从远小近大的剧烈尺度变化，而且不同行驶角度（斜向、俯仰）会导致标志几何形变，目标的尺度与视角差异变化剧烈，这些因素严重影响自动驾驶的交通标志实时检测精度、速度和鲁棒性。(2)复杂室外环境干扰强，严重影响交通标志实时检测的结果。在极端光照条件下，如强光直射、逆光、夜间无路灯、隧道明暗交替、雨天反光/眩光，均会造成图像中的标志过曝、欠曝、特征模糊，严重影响了目标检测的准确性。(3)自动驾驶系统的端侧平台算力相对受限，实时性与检测精度之间存在平衡矛盾。

与检测精度同等关键的是视觉检测模型在车载平台上的实时性与资源约束问题。自动驾驶车辆通常使用功耗和体积受限的嵌入式计算平台来运行交通标志检测算法，其算力与显存容量远低于数据中心GPU。例如，NVIDIA Jetson Xavier NX系列嵌入式设备仅有384个CUDA核心以及8 GB显存容量，峰值算力为21 TOPS，在部署深度学习模型时容易受到实时帧率的制约[6]。在工程实践中，为兼顾速度与可部署性，检测系统常采用如640像素×640像素的固定且较低的输入分辨率进行训练与推理。但这会进一步压缩小目标的有效像素，导致小尺寸交通标志特征更难保留，从而使精度与速度之间的矛盾更加突出。虽然动态推理[7]、早退出与边云协同等思路可以缓解部分实时性压力，但在端侧独立运行的场景下，仍需要从模型网络结构设计层面开展轻量化改进，使检测器更适宜小目标与低算力部署。

本文主要聚焦交通标志检测的两个关键挑战：(1)在输入分辨率较低时，小目标交通标志的判别特征易在下采样与复杂背景干扰中损失，导致检测精度受限。(2)在车载端实时算力约束下，需要在尽量控制参数量与计算量的同时，提升跨尺度信息交互与小目标表征能力，实现精度与速度的有效平衡。为了解决上述挑战，本文提出了一种针对交通标志检测的改进方法，其大体思路如下：(1)提出基于EDMS的C3k2_EDMS结构，通过分组动态分支与各向异性带状感受野的自适应选择，在不显著增加计算开销的前提下增强浅、中层细粒度与方向性信息建模，从而提升小目标交通标志的特征提取能力。(2)构建以HMSE为核心的增强型动态多尺度融合颈部，在P3～P5多尺度检测层之间建立更致密的跨层交互与一致融合，使各尺度特征在保持轻量的条件下获得更充分的互补增强，从而提升小目标检测的召回与整体性能。

2 相关工作

为了解决复杂交通环境下交通标志检测任务中小目标检测精度不高、难以有效地实现精度与实时性之间的平衡等问题，现有研究主要集中在3方面：小目标检测策略、动态多尺度特征融合方法与模型轻量化设计。

2.1 小目标检测

在实际道路场景中，交通标志在视频图像中通常以小尺寸目标形式出现，在整幅图像中交通标志所占像素面积有限，与车辆、建筑等中、大目标存在显著差异，交通标志在全图中的面积占比通常不足0.2%，并常伴随远距离成像导致的模糊、压缩失真和遮挡现象，使目标边缘与内部纹理易被淹没在背景中[8]。交通标志像素信息的不足往往导致特征信息稀缺，影响检测与定位的准确性。同时，道路场景的背景成分比较复杂，当小目标出现在边缘区域或多目标重叠区域时，检测难度进一步增大，在目标尺度小、对比度低时更容易引发误检与漏检。为了提高小目标检测精度，包括数据增强[9]、多尺度融合[10]、利用上下文信息[11]以及超分辨率技术[12]等方法被广泛应用。Chen等人[13]提出了YOLO-TS算法，通过优化多尺度特征图的感受野，使其与各种数据集中的交通标志尺寸分布更加匹配，显著提升了检测精度。Xu等人[14]设计了渐进式尺度提取器，在交通标志检测任务中取得了较好的精度表现。Li等人[15]提出了YOLOv8-NTS算法，通过引入轻量混合注意力变换器模块，提高了模糊交通标志在复杂环境下的检测精度。这些方法旨在增强模型对小目标的特征提取能力，从而提升检测性能。

2.2 多尺度特征融合

提升小目标检测性能的关键之一在于有效的多尺度特征融合，使不同层级特征在语义信息与空间细节之间达到合理匹配。若跨尺度信息传递不足，浅层特征可能因缺少语义支撑而产生误检，深层特征也可能由于空间分辨率较低而导致小目标漏检与定位偏差。此外，不恰当的融合还可能引入背景冗余与尺度冲突，并造成定位偏差，从而导致特征污染并降低训练稳定性。因此，为提高小目标交通标志的检测精度，设计高效且有选择性的多尺度特征融合机制是提升小目标检测效果的关键策略。常见的研究方法主要通过多尺度特征融合和上下文建模来增强对小尺寸目标的检测性能[16]。近年来，相关研究重点逐渐从应用传统的特征金字塔[17]转向自适应、双向交互与细节增强的融合范式，以更精细地建模不同尺度特征间的关联关系。Sun等人[18]提出自适应序列融合，通过动态协调多源特征之间的互补性，提高定位信息与语义信息的融合精度。Yin等人[19]设计了面向小目标的双向特征金字塔，强化不同尺度特征层之间的信息传递，使细节与语义能够更充分交换。Wang等人[20]提出了难点注意力感知网络，融合多尺度难点特征并加入位置信息，提升细粒度边界与误报抑制能力。Yang等人[21]在端到端Transformer检测器中引入多尺度边缘感知特征提取网络，通过跨尺度提取特征并增强边缘信息，显著提升小目标的识别与定位能力。此外，多尺度感受野[22]、动态平滑特征融合[23]等方法可依据目标尺度与分布自适应分配不同尺度特征的重要性，在遥感与密集小目标场景中取得了明显增益。总体而言，现有方法在跨尺度信息交互、融合权重自适应以及表征增强上具有明显优势，因此更能应对尺度变化与复杂背景干扰。然而，多数方法依赖额外的分支路径、特征对齐或注意力计算，导致结构复杂度上升、推理开销增加。同时，过度引入跨尺度交互也可能放大背景噪声，造成特征冗余与优化难度提升。因此，如何在保证融合效果的同时控制计算与存储开销、抑制无效信息传递，仍是多尺度特征融合的关键问题。

2.3 模型轻量化

现有一些在小目标检测方面精度表现突出的算法，虽在交通标志检测任务中取得了较好的检测效果，但往往伴随较高的计算开销，难以部署在算力与功耗受限的边缘设备上，从而限制了其在实际场景中的落地应用。为弥补这一不足，研究者提出并广泛采用了多种模型轻量化方案。基于YOLO系列的轻量化改进是交通标志检测中较为常见的思路，在轻量级检测架构中结合切片感知、注意力机制等策略，以提升小目标的特征表示能力[24]。Ghost模块[25]、RepConv模块[26]、轻量级骨干网络[27]以及紧凑的特征金字塔结构[28]等

被广泛用于替换原有卷积块，在接近原始精度的前提下显著减少参数量与每秒浮点运算次数（FLOPs）。同时，部分工作进一步结合量化方法[6]，以获得更小的模型与更高的推理速度。

总体而言，尽管现有研究在小目标检测、多尺度融合及模型轻量化方面取得了显著进展，但在算力资源相对有限的边缘设备上，要兼顾小目标检测精度高与模型实时性好之间达到较优平衡仍比较困难。为此，本文提出的 YOLO-DyMiF 从轻量化骨干网络和增强型动态多尺度颈部网络两个层面协同设计，旨在保持模型紧凑高效的同时，进一步提升小目标交通标志检测性能并提高部署友好性。

3　实现方法

本文提出的 YOLO-DyMiF 是一种基于 YOLOv10n 网络改进的交通标志检测器，如图 1 所示，其核心为基于 AEConv 和 EDMS 的 C3k2_EDMS 模块以及基于 HMSE 模块的动态特征融合颈部网络。这些模块和结构使模型更适合检测小目标交通标志，并有效控制了模型参数量和计算量的增长。图 1 中的 CBS_3 模块表示 3×3 卷积层、批归一化与 SiLU 激活函数的组合，SCDown 模块表示基于可分离卷积的下采样模块，SPPF 模块表示快速空间金字塔池化模块，PSA 模块表示位置敏感注意力模块。上述模块为常规 YOLOv10n 网络中的组成部分，本文不再展开说明。

为避免各改进结构之间产生混淆，本文结合图 1 统一说明。AEConv 为基础算子，用于多方向分支的自适应加权。EDMS 以 AEConv 为核心单元并通过两级残差组织多路径计算，同时完成通道交互与冗余抑制。C3k2_EDMS 将 EDMS 嵌入 C3k2 框架，用于主干替换 C2f 并在颈部特征更新位置复用。HMSE 位于颈部，负责多尺度特征融合与增强，颈部输出 P3、P4、P5 三个尺度特征至检测头。

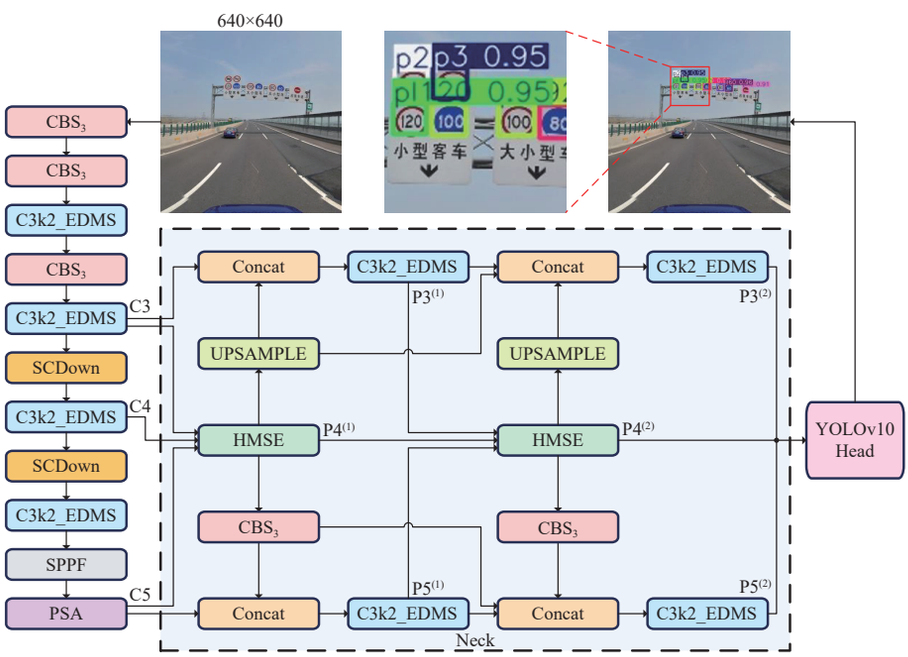

图 1　YOLO-DyMiF 网络结构

Fig. 1　Network architecture of YOLO-DyMiF

3.1　AEConv 模块

在交通标志检测任务中，目标尺度变化显著且小目标占比高，浅层特征细节丰富但语义不足，深层特征语义强但空间分辨率低。传统卷积采用静态且固定的卷积核，其感受野单一、跨尺度特征建模能力有限，并且缺乏对输入内容的自适应调整能力，导致在复杂背景与尺度变化场景下容易出现小目标漏检、定位偏差以及对干扰背景

的误响应。

为此，本文提出轻量高效的自适应混合卷积模块AEConv，通过多尺度深度卷积路径和动态权重融合实现对输入的自适应特征建模，以兼顾精度与速度。AEConv模块主要由多方向深度卷积分支、分支权重计算以及动态加权融合3部分组成，如图2所示。

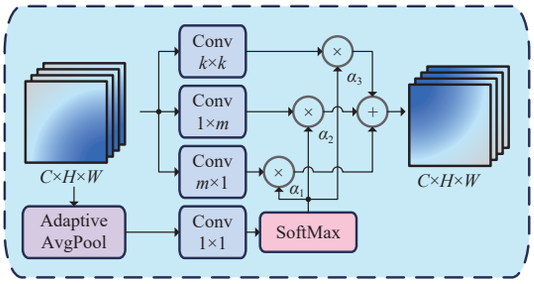

图2 AEConv结构示意图

Fig. 2 Schematic diagram of the AEConv architecture

设输入特征图为$\boldsymbol{X}_{\text{in}} \in \mathbb{R}^{C\times H\times W}$，首先并行通过3条卷积核分别为$k\times k$的、形如$1\times m$的以及形如$m\times 1$的深度可分离卷积分支，获取具有方向性的特征信息，以建立互补特征。随后利用含全局平均池化与逐点卷积的轻量化门控网络，在3条分支上分别计算软权重α_i。最后将3条分支输出的特征图分别与软权重相乘，再依次相加，经过批归一化与非线性激活后，AEConv的输出为$\boldsymbol{X}_{\text{out}} \in \mathbb{R}^{C\times H\times W}$。模块内的计算过程可以表示为：

$$\boldsymbol{X}_0 = \text{DWConv}_{k\times k}(\boldsymbol{X}_{\text{in}}), \tag{1}$$

$$\boldsymbol{X}_1 = \text{DWConv}_{1\times m}(\boldsymbol{X}_{\text{in}}), \tag{2}$$

$$\boldsymbol{X}_2 = \text{DWConv}_{m\times 1}(\boldsymbol{X}_{\text{in}}), \tag{3}$$

$$\alpha_i = \text{SoftMax}(\text{Conv}_{1\times 1}(\text{AAP}(\boldsymbol{X}_{\text{in}}))), \tag{4}$$

$$\boldsymbol{X}_{\text{out}} = \text{SiLU}\left(\text{BN}\left(\sum_{i=0}^{2}\boldsymbol{X}_i \odot \alpha_i\right)\right), \tag{5}$$

其中：$\boldsymbol{X}_i$表示深度可分离卷积分支输出的中间特征图；α_i表示计算获得的软权重，且应满足$\sum_{i=0}^{2}\alpha_i = 1$；$\text{DWConv}_{k\times m}(\cdot)$表示卷积核大小为$k\times m$的深度可分离卷积函数；$\text{SoftMax}(\cdot)$表示归一化指数函数；$\text{Conv}_{i\times i}(\cdot)$表示卷积核大小为$i\times i$的卷积函数；$\text{AAP}(\cdot)$表示全局平均池化操作；$\text{BN}(\cdot)$表示批归一化；$\text{SiLU}(\cdot)$表示SiLU激活函数；$\boldsymbol{X}_{\text{out}}$表示AEConv模块最终输出的特征图。

由于3条分支均采用深度卷积实现，其计算形式等价于按通道分组的并行分支提取，因此AEConv可视为一种轻量的分组动态分支结构。方形卷积分支提供各向同性的局部感受野，适合增强细碎纹理与小目标边缘细节；带状卷积分支对应各向异性带状感受野，能够在水平或垂直方向上更高效地聚合长程上下文信息，有利于刻画交通标志轮廓边缘、条带状目标以及运动模糊带来的方向性形变。由于带状卷积在计算开销上通常低于同尺度的方形大核卷积，结合深度卷积的分组计算特性，AEConv能够在较低额外开销下提升方向性建模能力。

为使3条多方向深度卷积分支能够依据输入内容实现更细粒度的协同，AEConv在分支融合阶段引入了通道级软门控机制。具体而言，利用由全局平均池化与逐点卷积构成的轻量门控网络，从输入特征的全局统计信息中为每个通道预测3条分支的融合权重，通过归一化约束使各分支权重在同一通道内可比较且和为1，随后将分支输出按对应权重进行加权叠加，从而在保持计算开销较低的前提下，实现对不同方向感受野的输入自适应选择与组合。当某一方向结构更显著时，该分支对应的权重提升，使输出在该通道上更侧重与对应方形或带状感受野的响应，实现通道级软选择。

这种通道级软门控的设计等价于在每条通道上进行可解释的、输入自适应的分支选择。当场景包含细碎纹理或小目标边缘时，方形卷积分支获得更高权重。当目标呈现细长状态时，对应方向的带状卷积分支获得更高占比，从而在无需额外引入自注意力的条件下实现各向异性感受野的动态调度。

3.2 C3k2_EDMS模块

交通标志检测中的特征往往同时呈现多尺度、多方向与结构复杂的特点。传统卷积依赖固定尺寸卷积核与静态参数共享，难以同时兼顾局部细节与全局上下文，对方向性结构刻画不足，并且缺乏对输入变化的自适应响应。更关键的是，在车载平台上，标准卷积与常规瓶颈结构通常带来较大的参数量与计算量开销，不利于实时部署。

为此，本文在C3k2骨架中引入EDMS模块，形成C3k2_EDMS，用以替换YOLOv10n中的C2f模块。该设计通过通道划分、多路径动态深

度卷积以及轻量融合压缩，在增强多方向性建模能力的同时，实现了对参数规模的有效压缩。

C3k2_EDMS 以 C3k2 为骨架，将其中的 Bottleneck 结构替换为 EDMS 单元。如图 3 所示，EDMS 模块由两级残差结构串联组成。设 EDMS 模块的输入特征图为 $\boldsymbol{E}_{\mathrm{in1}} \in \mathbb{R}^{C \times H \times W}$，在第一层残差结构中，输入特征沿通道维度均分为两组，分别得到 $E_0, E_1 \in \mathbb{R}^{C/2 \times H \times W}$。两组子特征分别采用不同的卷积核尺寸 k_1、k_2 的 AEConv 计算后，获得尺度不变的多尺度增强特征。随后将两路输出在通道维度拼接，恢复为 $\mathbb{R}^{C \times H \times W}$，并通过逐点卷积完成通道交互与压缩。经门控卷积和第二层残差耦合后，最终得到输出特征图 $\boldsymbol{E}_{\mathrm{out1}} \in \mathbb{R}^{C \times H \times W}$。EDMS 模块内的计算流程如下：

$$\boldsymbol{E}_0, \boldsymbol{E}_1 = \mathrm{Split}(\boldsymbol{E}_{\mathrm{in1}}), \tag{6}$$

$$\begin{cases} \boldsymbol{E}_0' = \mathrm{AEConv}(\boldsymbol{E}_0) \\ \boldsymbol{E}_1' = \mathrm{AEConv}(\boldsymbol{E}_1) \end{cases}, \tag{7}$$

$$\boldsymbol{U}_0 = \boldsymbol{E}_{\mathrm{in1}} + \mathrm{CBS}_1(\mathrm{Concat}(\boldsymbol{E}_0' + \boldsymbol{E}_1')), \tag{8}$$

$$\boldsymbol{U}_1 = \mathrm{Conv}_{1\times1}(\boldsymbol{U}_0) \odot \mathrm{DWConv}_{3\times3}(\mathrm{Conv}_{1\times1}(\boldsymbol{U}_0)), \tag{9}$$

$$\boldsymbol{E}_{\mathrm{out1}} = \boldsymbol{U}_0 + \mathrm{Conv}_{1\times1}(\boldsymbol{U}_1), \tag{10}$$

其中，$\boldsymbol{E}_{\mathrm{in1}}$ 表示输入 EDMS 模块的特征图，$\boldsymbol{E}_0$、$\boldsymbol{E}_1$、$\boldsymbol{E}_0'$、$\boldsymbol{E}_1'$、$\boldsymbol{U}_0$、$\boldsymbol{U}_1$ 表示计算过程中产生的中间特征图，Spilt(·)表示通道划分操作，Concat(·)表示通道拼接操作，$\boldsymbol{E}_{\mathrm{out1}}$ 表示 EDMS 结构的输出特征图。

与 YOLOv10n 模型中使用的 C2f 结构相比，C3k2_EDMS 中的 EDMS 模块降低了结构整体的参数量。记 C3k2_EDMS 的参数量为：

$$N_{\mathrm{C3k2_EDMS}} = N_{\mathrm{cv1}} + n \cdot N_{\mathrm{EDMS}} + N_{\mathrm{cv2}}, \tag{11}$$

其中：N_{cv1}、N_{cv2} 表示 C3k2 结构输入和输出部分卷积模块的参数量，n 表示 EDMS 模块的重复次数，N_{EDMS} 表示 EDMS 模块的参数量。

在输出特征图和相关计算参数一致的前提下，C3k2_EDMS 模块和 C2f 模块的参数量差异主要来源于两者内部结构不同，具体体现在 EDMS 结构和 Bottleneck 结构的卷积形式上。设模块的输入通道为 C，Bottleneck 结构大量采用常规卷积计算，其参数量随 C^2 增长。相比之下，EDMS 结构以深度可分离卷积为主，其参数量随 C 线性增长。因此在重复次数 n 和卷积核尺寸相同的情况下，使用 C3k2_EDMS 模块可以降低模型的参数量。

当特征中包含密集细节或小目标边缘时，小核分支获得更高权重以保留精细纹理；当目标呈现较强的条带结构或方向性形态时，对应方向分支占比上升以强化上下文建模。随后，两级残差结构将通道交互与冗余抑制结合起来。在第一层残差中，两路分支特征先进行拼接，拼接结果通过逐点卷积实现跨分支通道混合，从而完成通道交互。同时，逐点卷积的压缩映射会削弱高度相关通道的重复响应，借助残差连接保留原始信息，避免过度压缩带来的判别信息损失。在第二层残差中，门控卷积对中间特征进行自适应重标定，并以加性残差形式注入对当前输入有贡献的增量信息，使重复强化的背景纹理和噪声响应被门控抑制，从而在增强多方向性建模能力的同时，避免多分支叠加引起的冗余累积。消融实验中引入 C3k2_EDMS 后精确率提升，说明该结构在提升判别性的同时可以减少误检。

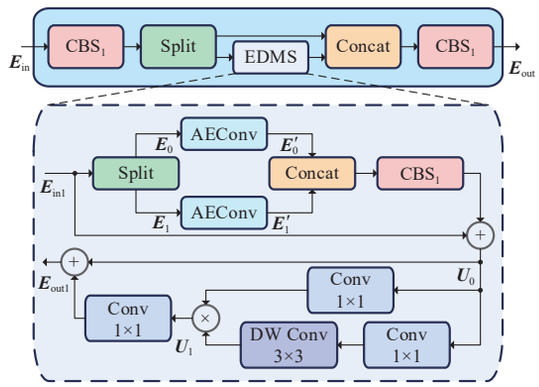

图 3　C3k2_EDMS 模块结构示意图

Fig. 3　Schematic diagram of the C3k2_EDMS architecture

3.3　HMSE 模块

在多次下采样过程中，小目标的边缘和纹理等细粒度信息容易被削弱，导致不同特征之间的语义与空间细节难以统一建模，从而引发小目标漏检与定位偏差。为了解决这一问题，本文提出了 HMSE 模块，将高、中、低层特征对齐到同一空间分辨率并在通道维度拼接，利用多尺度深度卷积，在保持轻量化的前提下显著扩大有效感受野，从而增强小目标的上下文建模能力与定位稳定性。

HMSE 的工作流程可以分为跨层对齐融合与多尺度空间增强两步。如图 4 所示，设输入模块的高层、中层、低层特征分别为 $\boldsymbol{F}_{\mathrm{h}} \in \mathbb{R}^{C_1 \times H_1 \times W_1}$、$\boldsymbol{F}_{\mathrm{m}} \in \mathbb{R}^{C_2 \times H_2 \times W_2}$、$\boldsymbol{F}_{\mathrm{l}} \in \mathbb{R}^{C_3 \times H_3 \times W_3}$。首先，将 3 路输入特征对齐到相同的空间尺寸，并在通道维度进

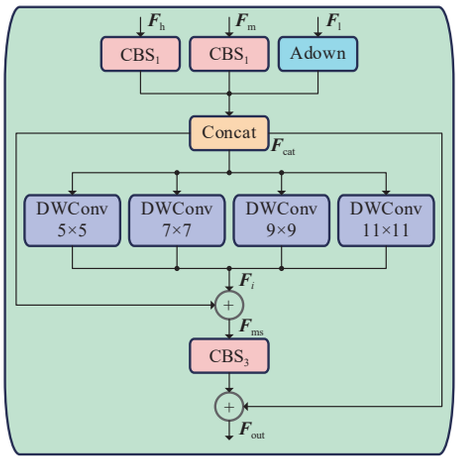

图4 HMSE模块结构示意图

Fig. 4 Schematic diagram of the HMSE architecture

行拼接，得到融合特征。随后，对融合特征施加多组卷积核尺寸不同的深度可分离卷积，以获取多尺度空间响应，并通过残差连接将多尺度响应与原始融合特征进行融合。最后，使用逐点卷积完成通道交互与调整，输出增强后的特征图。HMSE模块内部的计算流程如下：

$$\begin{cases} F_h' = \mathrm{CBS}_{1\times1}(\mathrm{Up}(F_h)) \\ F_m' = \begin{cases} \mathrm{CBS}_{1\times1}(F_m), e \neq 1 \\ F_m, e = 1 \end{cases} \\ F_l' = \mathrm{ADown}(F_l) \end{cases}, \tag{12}$$

$$F_{cat} = \mathrm{Concat}(F_h', F_m', F_l') \in \mathbb{R}^{C\times H\times W}, \tag{13}$$

$$F_i = \mathrm{DWConv}_{k_i\times k_i}(F_{cat}), \tag{14}$$

$$F_{ms} = F_{cat} + \sum_{i=0}^{3} F_i, \tag{15}$$

$$F_{out} = F_{cat} + \mathrm{CBS}_3(F_{ms}), \tag{16}$$

其中Up(·)代表两倍上采样；ADown(·)代表YOLOv9模型[29]中的可学习下采样；e为通道缩放因子，用于调整特征通道宽度；F_h'、F_m'、F_l'为对齐后的特征图，其空间尺寸一致，均为$c\times H\times W$，其中c表示对齐后的特征图通道数；F_{cat}表示拼接后的特征图，其通道数为$C=3c$；F_i表示多尺度空间响应特征，由卷积核尺寸为$k_i\in\{5,7,9,11\}$的深度可分离卷积计算获得；F_{ms}表示融合后的特征图；F_{out}表示HMSE模块的最终输出。

具体而言，在特征对齐时，以中层特征F_m的分辨率为基准，将3路输入特征对齐到$c\times H\times W$；随后将输入特征拼接为$C\times H\times W$，其中，$C=3c$；然后使用4组不同卷积核尺寸的深度可分离卷积获取不同感受野；最后采用特征融合残差以及逐点卷积残差的两次加和，有助于稳定训练并抑制多尺度卷积引入的冗余响应。

3.4 动态特征融合颈部网络

在交通标志检测任务中，小目标占比高且背景复杂。传统PAFPN(Path Aggregation Feature Pyramid Network)在跨层信息交互环节通常依赖有限的自顶向下或自底向上传递，易出现浅层特征语义支撑不足、深层特征空间分辨率受限的问题，从而导致小目标漏检、边界模糊与定位偏差。为此，本文在颈部网络部分设计一种以HMSE为核心的动态特征融合结构，通过在中间尺度上进行更充分的跨层融合，并将增强后的上下文信息有效传播到各检测尺度，以提高小目标检测的稳定性与定位精度。

该颈部网络以骨干网络输出的三尺度特征{C3,C4,C5}为输入，在网络内部生成并更新{P3,P4,P5}3个检测尺度特征，最终输入检测头。网络整体拓扑采用两次HMSE融合以及两轮尺度传播更新的机制。首先，将{C3,C4,C5}输入HMSE模块，在P4层级上完成跨层融合，得到$P4^{(1)}$。随后，由$P4^{(1)}$分别进行上采样和下采样，再与C3、C5拼接，经C3k2_EDMS模块更新后得到$P3^{(1)}$与$P5^{(1)}$。接着，将$\{P3^{(1)},P4^{(1)},P5^{(1)}\}$再次输入HMSE模块，进一步融合重建，获得$P4^{(2)}$。最后，采用与第一次尺度传播相似的方法得到$P3^{(2)}$和$P5^{(2)}$，颈部网络输出$\{P3^{(2)},P4^{(2)},P5^{(2)}\}$，以达到支撑多尺度检测头实现对大、中、小目标协同检测的目的。

下文中，记↑(·)为2倍上采样，↓(·)为2倍下采样，由CBS(Convolution-BatchNorm-SiLU)模块实现；Cat(·)为通道拼接；ϕ(·)表示C3k2_EDMS模块。有：

$$\begin{cases} P3^{(1)} = \phi(\mathrm{Cat}(\uparrow(P4^{(1)}), C3)) \\ P4^{(1)} = \mathrm{HMSE}(C3, C4, C5) \\ P5^{(1)} = \phi(\mathrm{Cat}(\downarrow(P4^{(1)}), C5)) \end{cases}. \tag{17}$$

保留第一次传播时产生的中间尺度转换特征$\tilde{P}3^{(1)}=\uparrow(P4^{(1)})$、$\tilde{P}5^{(1)}=\downarrow(P5^{(1)})$，进行第二次融合与尺度传播：

$$\begin{cases} P3^{(2)} = \phi(\mathrm{Cat}(\tilde{P}3^{(1)}, \uparrow(P4^{(2)}), P3^{(1)})) \\ P4^{(2)} = \mathrm{HMSE}(P3^{(1)}, P4^{(1)}, P5^{(1)}) \\ P5^{(2)} = \phi(\mathrm{Cat}(\tilde{P}5^{(1)}, \uparrow(P4^{(2)}), P5^{(1)})) \end{cases}. \tag{18}$$

这样的设计使得每个尺度的语义都经历了多轮跨层交互和HMSE的多尺度重建，对于小目

标密集、尺寸跨度极大的交通标志场景，会比传统PAFPN颈部更稳定、更细致，且不引入计算量较大的注意力机制，在增强跨尺度交互的同时控制了额外参数与推理开销。

4　实验结果及分析

4.1　数据集

为了验证所提方法的有效性，本文在TT100K（Tsinghua-Tencent 100K）数据集[30]上进行实验验证与对比。TT100K是面向自动驾驶场景的大规模交通标志数据集，包含约10万张街景图像和数万个标注实例，覆盖限速、禁令、警告、指示等200余类中国道路标志。图像分辨率约为2 000像素×2 000像素。交通标志多为远距离小目标，且伴随复杂背景、光照变化和遮挡，非常适合评估小目标交通标志检测与识别算法的性能。由于本文主要聚焦小目标交通标志检测问题，因此将数据集中图片的分辨率统一调整为640像素×640像素，且只挑选TT100K数据集中的45种样本数量大于100的交通标志进行实验，以避免样本分布长尾现象的影响。在实验中，选取了TT100K数据集的9 738张图片作为数据集，按照7∶1∶2的比例划分训练集、验证集和测试集。

4.2　实验环境及参数

本文的实验在Intel Xeon Platinum 8481C CPU、NVIDIA GeForce RTX 4090D GPU、80 GB内存的PC机上进行。在输入图像尺寸为640像素×640像素的情况下，批大小设为32，在0.01的学习率下使用SGD优化器对模型训练。为了公平起见，本文的所有算法均训练200轮。

4.3　评价指标

与之前的交通标志检测算法类似，本文采用参数量（Parameters）、浮点运算量（GFLOPs）、平均精度均值（mAP50）、精确率（P）、召回率（R）以及F1分数对模型性能以及资源消耗量进行评价。这些指标用以下公式计算：

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}, \tag{19}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{20}$$

$$\mathrm{F1}=\frac{2\times P\times R}{P+R}, \tag{21}$$

$$\mathrm{mAP50}=\frac{1}{N}\sum_{i=1}^{N}\mathrm{AP}_{50}^{i}, \tag{22}$$

其中：TP表示正样本被正确检测的数量；FP是负样本被检测为正样本的数量，即误报数量；FN是正样本未被检出的数量，即漏报数量；N表示类别总数；AP_{50}^{i}表示每个类别在50% IoU时的平均精度；F1分数是精确率和召回率的综合指标，是二者的调和平均数，用来在P和R之间取折中评价。

4.4　实验结果

4.4.1　消融实验

为了分析不同模块在YOLO-DyMiF中起到的作用，本文针对C3k2_EDMS模块、HMSE模块以及动态多尺度融合颈部进行了消融实验。由于HMSE模块是动态多尺度融合颈部的关键组成部分，因此下文将使用HMSE_FPN指代HMSE模块及动态多尺度融合颈部。如表1所示，本文以YOLOv10n作为基线模型，逐步引入C3k2_EDMS模块以及HMSE_FPN结构，对各模块的贡献进行定量分析。

在仅使用基线结构时，模型参数量为2.27M、计算量为6.6 GFLOPs，在TT100K数据集上获得63.4%的mAP50和61.9%的F1。引入C3k2_EDMS后，参数量和计算量分别下降至1.92M和5.5 GFLOPs，较基线减少约15.4%和16.7%，同时mAP50提升至65.5%、F1提升至63.9%，精确率从65.1%提升到70.3%，仅召回率略微下降0.5个百分点。由于C3k2_EDMS的核

表1　不同模块与结构在YOLO-DyMiF中的消融实验结果

Tab. 1　Ablation study results of different modules and structures in YOLO-DyMiF

Base	C3k2_EDMS	HMSE_FPN	Parameters/M	GFLOPs	mAP50/%	P/%	R/%	F1/%
√	×	×	2.27	6.6	63.4	65.1	59.0	61.9
√	√	×	**1.92**	**5.5**	65.5	70.3	58.5	63.9
√	×	√	2.84	8.9	66.8	**73.2**	59.0	65.3
√	√	√	2.50	7.9	**68.6**	**73.2**	**61.5**	**66.9**

注：√表示应用了该模块或结构，×表示没有应用该模块或结构，Base表示YOLOv10n基线模型，加粗数字表示最优结果。

心改动在于将EDMS嵌入C3k2结构，因此表1中Base与Base+C3k2_EDMS的对比可视为对EDMS有效性的直接验证，也是对其输入输出特征建模能力的定量说明。这表明C3k2_EDMS在显著压缩模型复杂度的同时，有效增强了浅层与中层特征的判别性，使检测结果更稳健，虽然召回率略有保守，但综合性能得到提升。

由表1可知，Base与Base+HMSE_FPN的对比可视为动态特征融合颈部引入前后的整体性能比较。当在基线网络上仅引入HMSE_FPN模块时，参数量和GFLOPs分别增加到2.84M和8.9 GFLOPs，但mAP50进一步提升至66.8%、F1提升至65.3%，精确率达到73.2%，召回率保持在59.0%。相较于基线，这说明HMSE_FPN通过分层多尺度融合与残差细化，显著提升了特征表达质量和定位精度，但计算代价有所增加。

在同时引入C3k2_EDMS与HMSE_FPN构成完整的YOLO-DyMiF时，参数量为2.50M、计算量为7.9 GFLOPs，介于两种单独改进方案之间，此时mAP50达到68.6%，精确率为73.2%，召回率提升到61.5%，F1进一步提高到66.9%，在4种配置中取得最佳综合性能。与仅使用HMSE_FPN相比，完整模型在保持相同精确率的同时，将参数量和GFLOPs分别降低了0.34M和1.0 GFLOPs，并在mAP50和F1上分别提升1.8个百分点和1.6个百分点，召回率也提升了2.5个百分点。上述结果表明，C3k2_EDMS与HMSE_FPN在结构上具有良好的协同效应：前者提供更高效的特征提取与轻量化能力，后者增强多尺度融合与上下文建模，二者结合既保证了小目标交通标志检测精度，又兼顾了模型复杂度与计算开销。

为进一步验证动态特征融合颈部网络在生成增强输出特征图过程中的关键操作，本文在保持训练策略和评估设置不变的条件下，对3项关键设计进行细粒度消融分析。具体而言，E1对应颈部网络中由C3k2_EDMS实现的自适应加权融合，E2对应HMSE中通过残差连接将多尺度响应与原始融合特征进行融合，E3对应HMSE中通过逐点卷积完成通道交互与调整。由表2可见，当3项机制同时启用时模型性能最优，mAP50为68.6%，F1为66.9%。关闭E1后，mAP50降至67.2%，精确率与召回率同步下降，说明通道级软门控实现的自适应加权融合能够根据输入动态选择有效分支响应，从而提升融合特征质量。关闭E2后，mAP50降至66.7%，召回率由61.5%降至57.2%，表明残差连接将多尺度响应与原始融合特征进行叠加有助于缓解信息丢失并减少漏检，从而提升小目标检测的稳定性。关闭E3虽使参数量与计算量略有下降，但精确率由73.2%降至69.0%，mAP50降至67.0%，说明逐点卷积在完成通道交互与调整的同时能够抑制冗余响应并降低误检。上述结果表明，E1的自适应加权融合、E2的残差融合以及E3的逐点卷积通道交互共同作用，能够有效提升颈部输出特征的判别性与多尺度一致性。

表2 动态特征融合颈部网络中关键机制的消融实验结果

Tab. 2 Ablation results of key mechanisms in the dynamic feature fusion neck network

E1	E2	E3	Parameters/M	GFLOPs	mAP50/%	*P*/%	*R*/%	F1/%
×	√	√	2.50	7.9	67.2	71.1	60.7	65.5
√	×	√	2.50	7.9	66.7	**75.4**	57.2	65.1
√	√	×	**2.38**	**7.6**	67.0	69.0	60.5	64.5
√	√	√	2.50	7.9	**68.6**	73.2	**61.5**	**66.9**

注：√表示应用了该设计，×表示未应用该设计，E1表示颈部网络中由C3k2_EDMS实现的自适应加权融合，E2表示HMSE中的残差融合，E3表示HMSE中的逐点卷积通道交互与调整，加粗数字表示最优结果。

4.4.2 对比实验

为了验证本文所提YOLO-DyMiF方法的性能，在TT100K数据集上完成与YOLO-Master-n、Hyper-YOLOt和Mamba-YOLOt等11种针对小目标检测的最新代表性方法的对比实验，结果如表3所示。

通过表3可见，YOLO-DyMiF在参数量仅为2.50M、计算量为7.9 GFLOPs的前提下，取得了

表 3　YOLO-DyMiF 与其他方法的对比实验结果

Tab. 3　Comparative experimental results of YOLO-DyMiF and other methods

Models	Input size	Parameters/M	GFLOPs	mAP50/%	*P*/%	*R*/%	F1/%
YOLOv5n	640×640	2.51	7.1	61.7	70.1	53.8	60.9
YOLOv8n	640×640	3.01	8.1	60.7	63.7	56.4	59.8
YOLOv9t	640×640	**2.01**	7.9	60.7	66.4	54.4	59.8
YOLOv10n	640×640	2.27	6.6	63.4	65.1	59.0	61.9
YOLOv11n	640×640	2.58	6.3	63.4	65.1	59.0	61.9
YOLOv12n	640×640	2.58	6.5	63.7	65.2	59.7	62.3
RT-DETR-L	640×640	32.0	103.0	58.3	64.5	50.2	59.7
YOLO26n	640×640	2.52	**5.9**	64.2	68.8	56.5	62.0
YOLO-Master-n	640×640	2.67	8.7	64.4	64.9	58.8	61.7
Hyper-YOLOt	640×640	3.01	9.0	66.2	66.6	60.5	63.4
Mamba-YOLOt	640×640	5.99	13.7	67.6	72.7	60.0	65.7
Ours	640×640	2.50	7.9	**68.6**	**73.2**	**61.5**	**66.9**

注：加粗数字表示最优结果

68.6%的mAP50、73.2%的精确率、61.5%的召回率以及66.9%的F1分数，在所有对比方法中取得最佳检测性能。另外，与同为轻量级的YOLO系列算法相比，YOLO-DyMiF在参数量略有增加、计算量保持同一量级的情况下，mAP50从63.4%提升到68.6%，F1从61.9%提升到66.9%，说明引入EDMS骨干、HMSE模块以及动态多尺度融合颈部后，模型对远距离小尺寸交通标志的表征与判断能力得到显著增强。与当前针对小目标检测的最新代表性方法YOLO-Master-n、Hyper-YOLOt、Mamba-YOLOt相比，YOLO-DyMiF在mAP50上分别提高了4.2%、2.4%和1%，F1从61.7%、63.4%和65.7%提高到66.9%，同时参数量从2.67M、3.01M和5.99M降至2.50M，运算量从8.7 GFLOPs、9.0 GFLOPs和13.7 GFLOPs降至7.9 GFLOPs，实现了在更低模型复杂度下获得更高检测精度的效果，表明本文提出的网络结构在精度与效率的权衡上优于现有轻量化交通标志检测器。

为了直观展示YOLO-DyMiF模型的检测效果，图5给出YOLO-DyMiF模型与最新的YOLO26n和Mamba-YOLOt算法的部分对比实验结果。图5(a～e)分别表示倾斜角度、逆光环境、阴影干扰、目标遮挡以及常规环境。3种方法在相同测试集与一致阈值设置下进行评估，以确保对比公平。由图5可见，YOLO-DyMiF在复杂背景与远距离小目标场景下检测质量更稳定，说明多尺度信息交互带来的小目标可见性增强能够有效提高检测性能，并减少漏检。

对比分析表明，C3k2_EDMS以轻量动态分支提升细粒度与方向性特征建模，有利于保留小目标纹理与边缘。动态特征融合颈部网络通过跨层融合与多轮重建提升上下文一致性，为低分辨率输入下的小目标提供了更充分的语义支撑。二者协同作用，在复杂度增幅较小的前提下提升了交通标志检测精度。

4.5　硬件部署

为了验证本文所提方法的工程实现可行性，在本节中还对YOLO-DyMiF模型与上述对照实验计算量最小的YOLO26n模型的部署效果进行详细对比。硬件部署实验在边缘计算平台NVIDIA Jetson Xavier NX上进行，如图6所示。两者均使用TensorRT FP16进行量化加速。TensorRT可以将PyTorch深度学习模型转换为优化的TensorRT引擎，通过层融合、精度校准和内存优化等技术增强部署在NVIDIA边缘设备上的模型的推理性能和效率。

在Jetson Xavier NX边缘平台上的测试实

图 5 TT100K 数据集上的交通标志检测结果。(a)倾斜视角;(b)逆光环境;(c)阴影干扰;(d)目标遮挡;(e)常规环境。

Fig. 5 Traffic sign detection results on the TT100K dataset. (a) Oblique views; (b) Backlit scenes; (c) Shadow disturbances; (d) Object occlusions; (e) Normal scenes.

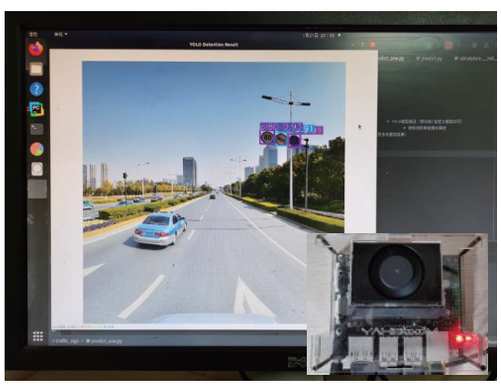

图 6 硬件部署平台及部署效果

Fig. 6 Hardware deployment platform and deployment performance

验使用 TT100K 数据集的测试集进行,主要将 YOLO-DyMiF 与计算量最低的 YOLO26n 模型进行对比,结果如表 4 所示。

通过表 4 可以发现,YOLO-DyMiF 在端侧推理场景下仍保持与服务器端一致的精度水平,同时精确率和召回率与服务器端实验结果相比变化不大。这表明 YOLO-DyMiF 在嵌入式平台上对背景噪声与相似干扰的抑制更强,能够显著减少误检,从而提升整体检测可信度。在速度方

表 4 YOLO-DyMiF 与 YOLO26n 在边缘平台的部署结果

Tab. 4 Deployment results of YOLO-DyMiF and YOLO26n on edge platforms

Models	mAP50/%	P/%	R/%	FPS
YOLO26n	64.3	62.7	**60.2**	**34.3**
Ours	**68.5**	**75.7**	59.9	32.1

注:加粗数字表示最优结果

面，YOLO-DyMiF 的推理帧率虽然稍微落后于 YOLO26n，但仍保持在 30 FPS 以上的实时视频流，可以满足工程的需要。考虑到 YOLO-DyMiF 在端侧平台更显著的精度增益，说明所提出结构在真实部署条件下具有较优的精度-效率权衡，能够以较小的速度代价换取更稳定、更可靠的小目标交通标志检测性能。

5　结　　论

本文提出了一种面向算力受限车载平台的轻量化交通标志检测网络 YOLO-DyMiF，旨在在较低输入分辨率下提升小目标识别能力，同时满足实时性需求。该方法以 YOLOv10n 为基础，围绕高效特征提取与跨尺度信息交互对结构进行优化。一方面，设计基于 AEConv 的高效动态混合器，并构建了即插即用的 C3k2_EDMS 模块，通过大量使用深度可分离卷积和动态门控系统，在不降低特征提取效果的前提下显著降低了模块的计算开销，使远距离、小尺寸交通标志在 640×640 输入条件下仍能保留足够的判别信息；另一方面，构建以 HMSE 为核心的增强型动态多尺度融合颈部网络，通过跨层特征交互机制与自适应重加权方法提升不同尺度特征的互补性与一致性，有效缓解多次下采样与复杂背景干扰造成的小目标特征弱化和尺度不匹配问题，从而提高整体召回率与检测稳定性。实验结果表明，在 TT100K 数据集上，YOLO-DyMiF 在仅有 2.50M 参数量和 7.9 GFLOPs 计算量的约束下取得 68.6% 的 mAP50 分数，并在 mAP50、F1 等指标上整体优于常规轻量级 YOLO 系列模型，表现出更具竞争力的小目标检测性能。实验结果验证了 C3k2_EDMS 在计算量受限条件下的特征提取能力，以及动态多尺度融合颈部网络的跨层信息融合能力，表明所提方法在现实道路环境中具有较好的应用前景和工程价值。

参　考　文　献：

[1] XU J, DU Y J, YI Y W, *et al*. An improved lightweight algorithm for traffic sign detection [J]. *Scientific Reports*, 2025, 15(1): 33554.

[2] ZHAO Z C, DU J X, LI C L, *et al*. Dense Tiny object detection: a scene context guided approach and a unified benchmark [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5606913.

[3] LIU Y Y, PENG J Y, XUE J H, *et al*. TSingNet: scale-aware and context-rich feature learning for traffic sign detection and recognition in the wild [J]. *Neurocomputing*, 2021, 447: 10-22.

[4] WANG P Z, MOHAMED R, MUSTAPHA N, *et al*. YOLO-AEF: traffic sign detection on challenging traffic scenes via adaptive enhancement and fusion [J]. *Neurocomputing*, 2025, 655: 131430.

[5] LIANG T L, GLOSSNER J, WANG L, *et al*. Pruning and quantization for deep neural network acceleration: a survey [J]. *Neurocomputing*, 2021, 461: 370-403.

[6] DING Y F, FENG W L, CHEN C Y, *et al*. Reg-PTQ: regression-specialized post-training quantization for fully quantized object detector [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). Seattle: IEEE, 2024: 16174-16184.

[7] RAZA S M, ABIDI S M H, MASUDUZZAMAN M, *et al*. Lightweight deep learning for visual perception: a survey of models, compression strategies, and edge deployment challenges [J]. *Neurocomputing*, 2025, 656: 131357.

[8] HAO C Y, ZHANG H, SONG W R, *et al*. SliNet: slicing-aided learning for small object detection [J]. *IEEE Signal Processing Letters*, 2024, 31: 790-794.

[9] TANG J X, CHENG J D, XIANG D L, *et al*. Large-difference-scale target detection using a revised Bhattacharyya distance in SAR images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 4506205.

[10] LI Y S, LIU F X, QIN Y S, *et al*. Small object detection network based on progressive enhanced multi-level feature fusion [J]. *Neurocomputing*, 2025, 649(13): 130755.

[11] ZHU Z Q, ZHENG R Z, QI G Q, *et al*. Small object detection method based on global multi-level perception and dynamic region aggregation [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, 34(10):

10011-10022.

[12] GAO Y X, WANG Y C, ZHANG Y X, *et al*. Feature super-resolution fusion with cross-scale distillation for small-object detection in optical remote sensing images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 6008105.

[13] CHEN J Z, HUANG H Q, ZHANG R H, *et al*. YOLO-TS: real-time traffic sign detection with enhanced accuracy using optimized receptive fields and anchor-free fusion [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 26(11): 19995-20011.

[14] XU Y L, XIAO X X. PST-YOLO: a lightweight, robust approach for traffic sign detection via polynomial-based and progressive scale fusions [J]. *Digital Signal Processing*, 2026, 169: 105745.

[15] 李鹏飞,熊召新,王桂宝. YOLOv8-NTS:一种针对交通标志检测的目标识别方法[J]. 液晶与显示,2025,40(12):1868-1880.
LI P F, XIONG Z X, WANG G B. YOLOv8-NTS: a target detection approach for traffic sign recognition [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(12): 1868-1880. (in Chinese)

[16] QU S M, YANG X Y, ZHOU H F, *et al*. Improved YOLOv5-based for small traffic sign detection under complex weather [J]. *Scientific Reports*, 2023, 13(1): 16219.

[17] 宋绍剑,夏海姐,李刚. YOLOv5的改进算法及其在自动驾驶多目标检测的应用研究[J]. 计算机工程与应用,2023,59(15):68-75.
SONG S J, XIA H J, LI G. Research on improved YOLOv5 algorithm and its application in multi-object detection for automatic driving [J]. *Computer Engineering and Applications*, 2023, 59(15): 68-75. (in Chinese)

[18] SUN Y, LI X, ZHAO D J, *et al*. Evolving traffic sign detection via multi-scale feature enhancement, reconstruction and fusion [J]. *Digital Signal Processing*, 2025, 160: 105028.

[19] YIN C F, WANG Z H, LI Y C, *et al*. SBEW-YOLOV8: a small object detection algorithm for autonomous driving based on multi-scale feature fusion [J]. *The Journal of Supercomputing*, 2025, 81(10): 1125.

[20] 王伯霄,宋延嵩,董小娜. 难点注意力感知红外小目标检测网络[J]. 中国光学(中英文),2024,17(3):538-547.
WANG B X, SONG Y S, DONG X N. Indistinguishable points attention-aware network for infrared small object detection [J]. *Chinese Optics*, 2024, 17(3): 538-547. (in Chinese)

[21] YANG L H, GU Y, FENG H. Multi-scale feature fusion and feature calibration with edge information enhancement for remote sensing object detection [J]. *Scientific Reports*, 2025, 15(1): 15371.

[22] TONG K, WU Y Q. Small object detection using deep feature learning and feature fusion network [J]. *Engineering Applications of Artificial Intelligence*, 2024, 132: 107931.

[23] AHMAD I, LU W, CHEN S B, *et al*. Lightweight oriented object detection with dynamic smooth feature fusion network [J]. *Neurocomputing*, 2025, 628: 129725.

[24] XU S, LI Y J, LIN M B, *et al*. Q-DETR: an efficient low-bit quantized detection transformer [C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE, 2023: 3842-3851.

[25] ZHANG S, CHE S B, LIU Z, *et al*. A real-time and lightweight traffic sign detection method based on ghost-YOLO [J]. *Multimedia Tools and Applications*, 2023, 82(17): 26063-26087.

[26] 王龙春,方巍,张丽娟,等. 基于YOLOv8s改进的自动驾驶目标检测[J]. 液晶与显示,2025,40(5):773-784.
WANG L C, FANG W, ZHANG L J, *et al*. Improved autonomous driving object detection based on YOLOv8s [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(5): 773-784. (in Chinese)

[27] QIU J, ZHANG W B, XU S Y, *et al*. DP-YOLO: a lightweight traffic sign detection model for small object detection [J]. *Digital Signal Processing*, 2025, 165: 105311.

[28] DOHERTY J, GARDINER B, KERR E, *et al*. BiFPN-YOLO: One-stage object detection integrating Bi-Directional Feature Pyramid Networks [J]. *Pattern Recognition*, 2025, 160: 111209.

[29] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information [C]//*Proceedings of 18th European Conference on Computer Vision-ECCV* 2024. Milan, Italy: Springer, 2024: 1-21.

[30] ZHU Z, LIANG D, ZHANG S H, *et al*. Traffic-sign detection and classification in the wild [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016: 2110-2118.

作者简介：

宋绍剑，男，硕士，教授，2001年于广西大学获得硕士学位，研究方向为机器学习、人工智能、新能源变换与控制。E-mail：sjsong03@163.com

李　刚，男，博士，教授，2009年于上海理工大学获得博士学位，主要研究方向为计算机视觉、模式识别及智能系统、自动驾驶技术。E-mail：liganglyd@126.com

李　昊，男，硕士研究生，2022年于广西大学获得学士学位，主要从事交通场景目标检测方面的研究。E-mail：lihao.0805@qq.com

李国进，男，博士，教授，2006年于华南理工大学获得博士学位，主要研究方向为模式识别与智能系统、机器人。E-mail：LGJGX@163.com