



电讯技术
Telecommunication Engineering
ISSN 1001-893X, CN 51-1267/TN

《电讯技术》网络首发论文

题目：一种改进多尺度特征融合的交通标志识别算法
作者：余翔，靳闪闪，杨路
DOI：10.20079/j.issn.1001-893x.240120002
收稿日期：2024-01-20
网络首发日期：2024-07-03
引用格式：余翔，靳闪闪，杨路. 一种改进多尺度特征融合的交通标志识别算法[J/OL]. 电讯技术. <https://doi.org/10.20079/j.issn.1001-893x.240120002>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

DOI: 10.20079/j.issn.1001-893x.240120002.

一种改进多尺度特征融合的交通标志识别算法

余翔，靳闪闪，杨路

(重庆邮电大学 通信与信息工程学院, 重庆 400065)

摘要：在背景复杂且目标距离远的情况下，智能交通系统或自动驾驶中关键的交通标志识别算法的识别准确率有待进一步提升。为此，在 YOLOv5s 算法的基础上提出了一种改进的交通标志识别算法 MAFM-YOLO。首先，在颈部网络设计了基于空洞混合注意力机制的多尺度注意力特征融合模块，使网络在特征融合阶段能够高效保留图像中的细节信息，对小目标更加的敏感。其次，在回归阶段采用归一化 Wasserstein 距离改进原有的损失函数，提高了边界框的回归性能，从而进一步提高网络的识别性能。在 TT100K 数据集上的实验结果表明，MAFM-YOLO 较基准模型在精确率 P、召回率 R 和 mAP 上分别实现了 9.4%、3.3%、6.3% 的提升。

关键词：交通标志识别；YOLOv5s；特征融合；混合注意力机制；归一化 Wasserstein 距离

中图分类号：TN911.73; TP391.4 **文献标志码：**A

An Improved Traffic Sign Recognition Algorithm Based on Multi-Scale Feature Fusion

Yu Xiang, Jin Shanshan, Yang Lu

(School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: The recognition accuracy of the key traffic sign recognition algorithm in intelligent transportation systems or autonomous driving needs to be further improved in the case of complex backgrounds and long distances. Therefore, an improved traffic sign recognition algorithm MAFM-YOLO is proposed based on YOLOv5s algorithm. Firstly, a multi-scale attentional feature fusion module based on the atrous blend attention mechanism is designed in the neck network, so that the network can efficiently retain the detail information in the image in the feature fusion stage, and is more sensitive to the small targets. Second, the normalized Wasserstein distance is used to improve the original loss function in the regression stage, which improves the regression performance of the bounding box, thus further improving the recognition performance of the network. Experimental results on the TT100K dataset show that MAFM-YOLO achieves 9.4%, 3.3%, and 6.3% improvement in precision P, recall R, and mAP, respectively, over the baseline model.

Key words: traffic sign recognition; YOLOv5s; feature fusion; blend attention mechanism; Normalized Wasserstein Distance

收稿日期：2024-01-20; 修回日期：2024-04-14

基金项目：国家自然科学基金 (No. 62176035)

通信作者：靳闪闪 Email: s220131032@stu.cqupt.edu.cn

0 引言

交通标志是道路交通控制中重要的组成部分,在传递信息、引导方向、维护道路交通秩序和安全等方面有着重要作用^[1]。但是交通标志图像往往具有背景复杂、小目标较多等特点,这给交通标志识别带来了挑战。目前,研究者们采用基于深度学习的卷积神经网络(Convolutional Neural Network, CNN)目标识别算法和计算机视觉技术对交通标志识别进行了广泛研究^[2],较传统识别算法的精度有明显提升。

传统交通标志识别算法依赖其颜色和形状特征,并结合支持向量机(Support Vector Machine, SVM)等方法确定类别。然而,这些方法需要人工选择和设计特征,训练过程复杂且错误率高。随着深度学习的发展,基于卷积神经网络的识别方法逐渐成为主流。以 R-CNN(Region Convolutional Neural Network)为代表的两阶段识别算法和以 SSD(Single Shot MultiBox Detector)、YOLO(You Only Look Once)为代表的单阶段识别算法成为当前交通标志识别领域的主流技术^[3]。Zhang 等^[4]针对复杂环境干扰的挑战,提出了一种融入多尺度注意模块的级联 R-CNN 网络,有效解决不平衡样本的识别问题。Hu 等^[5]采用 Faster R-CNN 进行交通标志识别,并通过去雾算法对图像进行处理,提高检测效果。以上两阶段识别方法在某些方面表现出色,但通常模型复杂、速度较慢。相比之下,单阶段识别算法模型简单、速度快、精度高,逐渐成为实时目标识别的首选。Wan 等^[6]使用改进的 YOLOv3 进行模型剪枝和多尺度预测,在减小模型参数数量的同时提高识别速度,但是实时识别性能不高。李志刚等^[7]采用遗传算法和 K-means 聚类方法优化锚框设定,结合 Stem 模块和 ShufflenetV2 构建了轻量化网络,便于移动端部署,但该方案在提升交通标志识别精度上的效果相对有限。

基于上述文献的启发,本文选取综合性能较好的单阶段目标识别模型 YOLOv5s 作为主体框架,提出了一种改进的交通标志目标识别算法 MAFM-YOLO。为了应对交通标志复杂的背景信息以及小目标过多的问题,设计了一种全新的多尺度注意力特征融合模块,该模块引入主干网络提取的更浅层特征信息,加强不同特征层信息的融合,减少小目标特征丢失,在该模块中设计了基于空洞卷积的混合注意力模块,增大感受野,有效去除复杂背景信息,增强对目标的关注度。针对小目标定位不准的问题,采用归一化 Wasserstein 距离优化损失函数,提高目标边界框的定位精度,有效改善小目标识别效果,进一步提高识别精度。实验结果表明,改进后的算法相较于原有算法对交通标志有更好的识别能力。

1 MAFM-YOLO 改进算法

1.1 MAFM-YOLO 整体架构

本文在 YOLOv5s 算法^[8]的基础上,提出改进的交通标志识别算法 MAFM-YOLO,整体架构如图 1 所示。在颈部网络加入多尺度注意力特征融合模块,替换 YOLOv5s 模型中与检测头输出相连的三个 Concat 模块,通过引入主干网络提取的更浅层的信息,实现浅层图形特征和深层语义特征的融合,该模块嵌入的空洞混合注意力机制能够有效抑制背景,增强前景,有效解决尺度变化引起的误检、漏检等问题。引入归一化 Wasserstein 距离优化损失函数,进一步提高小目标的识别精度。

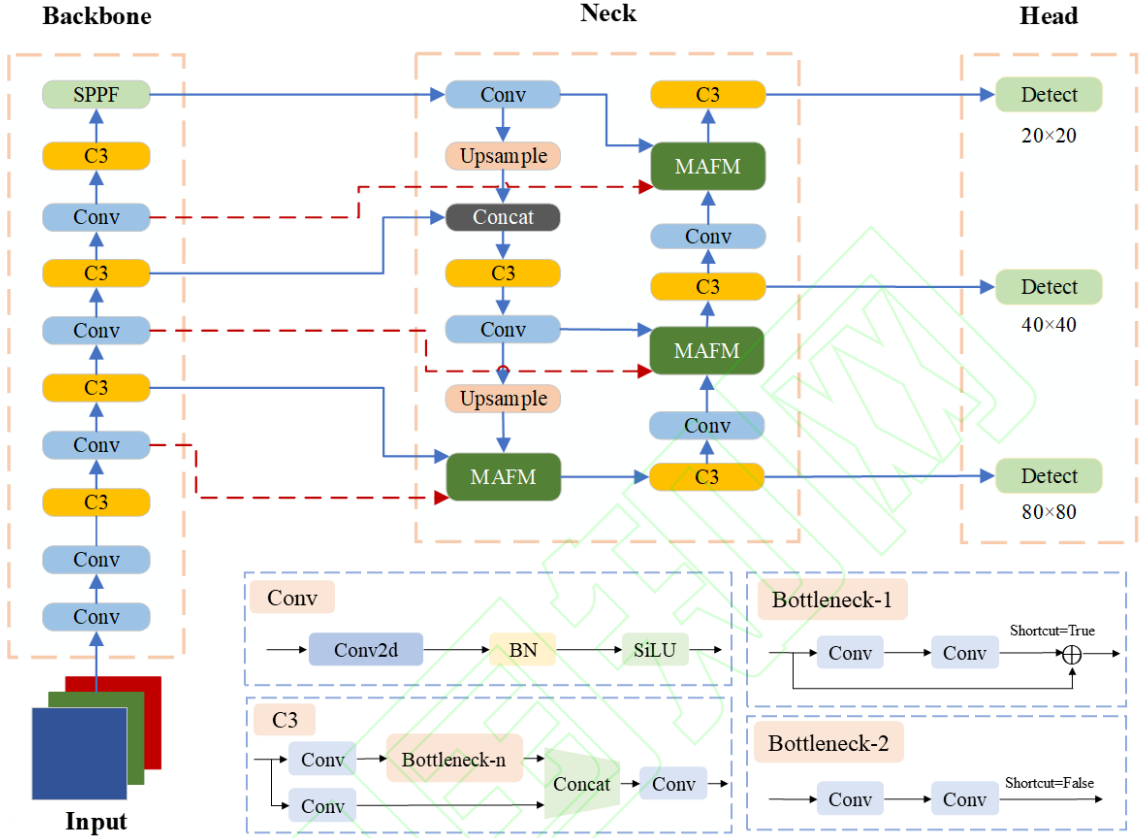


图 1 MAFM-YOLO 网络架构

Fig.1 MAFM-YOLO network architecture

1.2 多尺度注意力特征融合模块

交通标志图像具有背景信息复杂、小目标多等特点。YOLOv5s 模型通过结合 FPN(Feature Pyramid Networks)和 PAN(Path Aggregation Network)融合不同尺度特征信息，然而，在 FPN-PAN 中，浅层特征需经多层卷积至顶层，下采样导致小目标特征丢失，背景冗余信息可能引发误检和漏检。此外，交通标志图像受拍摄角度、天气等因素影响，轮廓纹理模糊，增加了识别难度。注意力机制借鉴了人类对信息的选择性关注机制，使神经网络在图像处理时能够更智能地适应和感知。在图像识别任务中，注意力机制常常与卷积神经网络结合使用，通过给模型的不同部分赋予不同权重，使模型能够聚焦于图像中的关键信息，从而提高识别精度和优化模型性能。

为此，本文设计了多尺度注意力特征融合模块（Multi-scale Attention Feature Fusion Module, MAFM），结构如图 2 所示。MAFM 是一个特征融合模块，它由普通卷积和基于空洞卷积的混合注意力模块（Atrous Blend Attention Mechanism, ABAM）组成，能够融合不同尺度的特征信息。该模块借鉴 ResNet^[9]的思想，在原有特征融合的基础上，添加跨层连接引入更浅层信息，融合主干网络提取的浅层、中层与深层特征图，同时在该模块设计了空洞混合注意力机制，能够有效增大感受野，降低背景信息的影响。其中，浅层特征图 F_1 和中层特征图 F_2 之间相差一个 C3 模块，而深层特征图 F_3 则是经过 SPPF 模块池化后，再通过 Neck 网络处理得到的。为了实现特征融合，首先通过卷积调整浅层特征图的尺寸，使其与中层特征图尺寸一致。接着，将调整后的浅层特征图与中层特征图融合，并对每个通道的特征进行相加，以增强特征表达而不增加通道数。随后，与深层特征进行融合之后通过 ABAM 注意力模块，该模块通过引入注意力机制，增强了目标信息的特征响应，同时降低了背

景信息的权重，提升了模型在复杂背景下的抗干扰能力。最后，通过 F 将三个层次的特征信息进行拼接，合并通道数，以充分利用浅层特征信息。

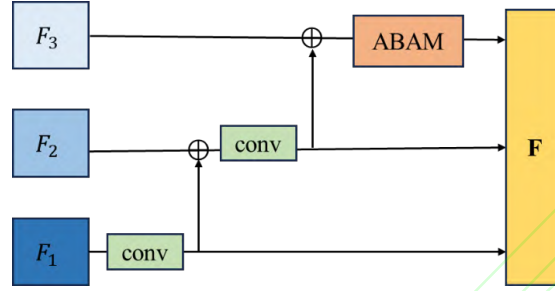


图 2 MAFM 结构
Fig.2 MAFM structure

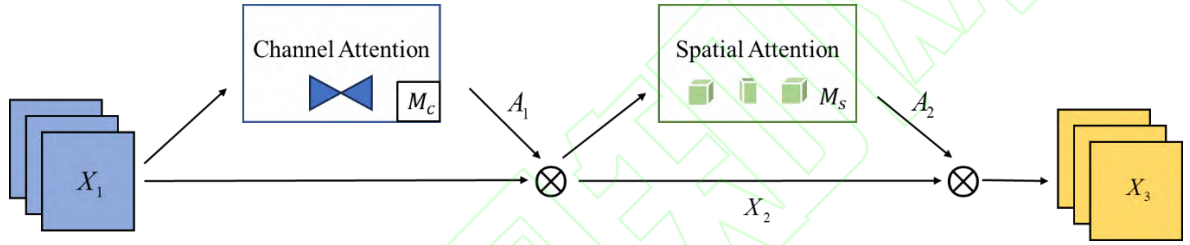


图 3 ABAM 结构
Fig.3 ABAM structure

ABAM 模块如图 3 所示，设输入特征图为 X_1 ，依次通过通道注意力与空间注意力模块，同时关注特征的通道信息与空间信息。首先，输入尺寸为 $H \times W \times C$ 的特征图进入通道注意力模块，如图 4 所示， C 代表图像的通道数， H 代表高度， W 代表宽度。通道注意力子模块使用三维排列来在三个维度上保留信息，采用两层的 MLP 放大跨维通道-空间依赖性，MLP 中先用一个线性层，将通道 C 压缩，压缩比率为 r ，通道数变为 C/r ，再使用非线性 $relu$ 激活函数，最后通过一个线性层将通道还原为 C ，得到的输出通道注意力权重表达式为：

$$A_1 = \text{sigmoid}(p'(linear(relu(linear(p(X_1))))) \quad (1)$$

式中， p 与 p' 表示三维重排， $linear$ 表示线性操作， $relu$ 表示非线性激活函数， sigmoid 表示激活归一化函数， A_1 为输出通道注意力权重。

将通道注意力权重通过乘积与输入特征图相结合，为：

$$X_2 = A_1 \otimes X_1 \quad (2)$$

式中， X_2 为通道注意力输出特征图。

特征图进入空间注意力模块，如图 5 所示，在该模块，输入特征图分别经过空洞率为 1、2、3 的空洞卷积^[10]，为了保持特征图的尺寸一致，padding 分别设置为：1，2，3，步长为 3，等价于 3×3 、 5×5 和 7×7 的普通卷积。其中小的卷积核可聚焦像素自身周围的特征信息，大的卷积核能提取深层语义信息。另外，采用空洞卷积可以在不增加过多参数和复杂度的情况下提高网络的感受野，充分利用多层次信息。经过 3 个卷积操作后的特征图通过 $concat$ 拼接起来，然后经过 1×1 卷积恢复原有维度，最后经过 sigmoid 函数得到归一化权重，与特征图 X_2 相结合，使得目标区域得以增强。输出空间注意力权重表达式为：

$$A_2 = \text{sigmoid}(\text{Conv}_4(\text{concat}(\text{Conv}_1(X_2), \text{Conv}_2(X_2), \text{Conv}_3(X_2)))) \quad (3)$$

式中， Conv_1 表示空洞率为 1，padding 为 1 的 3×3 空洞卷积， Conv_2 表示空洞率为 2，padding 为 2 的 3×3 空洞卷积， Conv_3 表示空洞率为 3，padding 为 3 的 3×3 空洞卷积， Conv_4 表示 1×1 普通卷积， $concat$ 表示拼接操作， A_2 表示输出通道注意力权重。

将空间注意力权重通过乘积与输入特征图相结合，通过 ABAM 的输出特征图为：

$$X_3 = A_2 \otimes X_2 \quad (4)$$

式中， X_3 为 ABAM 输出特征图。

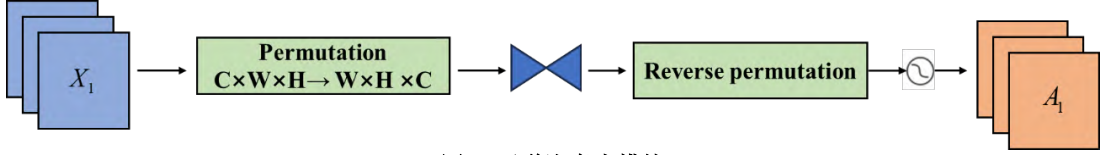


图4 通道注意力模块
Fig.4 Channel attention module

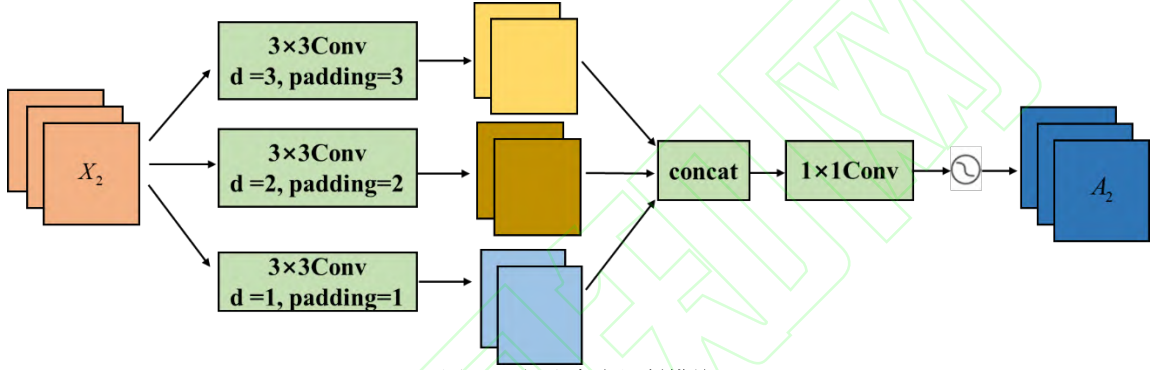


图5 空间注意力机制模块
Fig.5 Spatial attention module

1.3 NWD 损失函数

原始的 YOLOv5s 损失函数由分类损失、定位损失和目标损失加权而来，通过最小化损失函数，模型可以逐渐优化目标的类别预测、位置信息和目标存在概率，提升目标识别的准确性。原始 YOLOv5s 模型采用 IoU (Intersection over Union) 定位损失函数，当预测框与实际框之间无重叠或没有相互包含时，IoU 损失函数可能会导致梯度消失，从而引发误判。而 CIoU (Complete Intersection over Union) 和 DIoU (Distance Intersection over Union) 等基于 IoU 的损失函数同样对位置偏移较为敏感，因此需要考虑新的边界框度量方式。

本文采用归一化 Wasserstein 距离 (Normalized Wasserstein Distance, NWD)^[11] 来测量边界框的相似性。NWD 在小目标识别方面具有尺度不变性，同时能够平滑位置偏差，能够测量不重叠或相互包含的边界框之间的相似性^[12]，很好的解决了传统基于 IoU 对小目标造成的定位不准确问题。具体来说，首先将边界框建模为二维高斯分布，然后使用归一化 Wasserstein 距离来测量导出的高斯分布的相似性，从而衡量两个边界框之间相似性。具体定义如下：

水平边界框 $R = (c_x, c_y, \omega, h)$ ，其中 (c_x, c_y) , ω, h 分别表示中心坐标、宽度和高度，边界框的中心像素权重最高，像素的重要性从中心向边界递减。将边界框建模为二维高斯分布 $N(\mu, \Sigma)$ ， μ 和 Σ 分别表示高斯

分布的均值向量与协方差矩阵，定义为： $\mu = \begin{bmatrix} c_x \\ c_y \end{bmatrix}$, $\Sigma = \begin{bmatrix} \frac{\omega^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix}$ 。

边界框 A 和 B 之间的相似度可以转换为两个高斯分布之间的分布距离，二阶高斯分布 $\mu_1 = N(m_1, \Sigma_1)$, $\mu_2 = N(m_2, \Sigma_2)$ 之间的 Wasserstein 距离定义为：

$$W_2^2(\mu_1, \mu_2) = \|\mathbf{m}_1 - \mathbf{m}_2\|_2^2 + \text{Tr}(\Sigma_1 + \Sigma_2 - 2(\Sigma_1^{1/2} \Sigma_1 \Sigma_2^{1/2})^{1/2}) \quad (5)$$

归一化为：

$$NWD(\mu_1, \mu_2) = \exp\left(-\frac{\sqrt{W_2^2(\mu_1, \mu_2)}}{C}\right) \quad (6)$$

式中, C 为与数据集相关的常数。

将 NWD 用作损失函数:

$$L_{NWD} = 1 - NWD(\mu_1, \mu_2) \quad (7)$$

2 实验与结果分析

2.1 实验环境及参数配置

实验在 64 位 windows11 的操作系统上进行, GPU 为 NVIDIA GeForce RTX 3070 8G 显存, CPU 为 Intel(R) Core(TM) i5-12600K, 深度学习框架为 PyTorch2.0, 采用 python3.9 编程语言。后续的消融实验和对比实验中的训练参数设置与此一致。实验的 epochs 为 150, batchsize 为 16, 采用小批量随机梯度下降。输入图像的尺寸是 512×512 像素, 使用默认数据增强的方法, 无额外预训练, 其余实验参数均保持 YOLOv5s 默认设置。

2.2 数据集预处理

TT100K^[13]是由清华和腾讯联合创建的数据集, 该数据集由中国多个城市的 6 台高分辨率广角单反相机拍摄而成, 共收集了 1 万张包含交通标志的图像。主要包括三类交通标志: 提示标志、警告标志和禁止标志, 如图 6 所示。对其中的每个交通标志进行类别标签、边界框和像素掩码的标注。与其他交通标志数据集相比, TT100K 的特点在于分类更多, 目标更小, 识别难度更大。由于数据集中各类别标志的分布不平衡, 会影响模型的收敛速度, 导致模型在学习某些类别的特征时存在不足, 从而导致模型的泛化能力存在偏差。对该数据集进行处理, 步骤如下: (1) 对数据集进行筛选, 得到交通标志实例个数在 100 个以上的类别, 共有 42 类, 9448 张图片; (2) 对得到的图片进行划分, 其中, 6589 张作为训练集, 1889 张作为验证集, 970 张作为测试集进行实验。

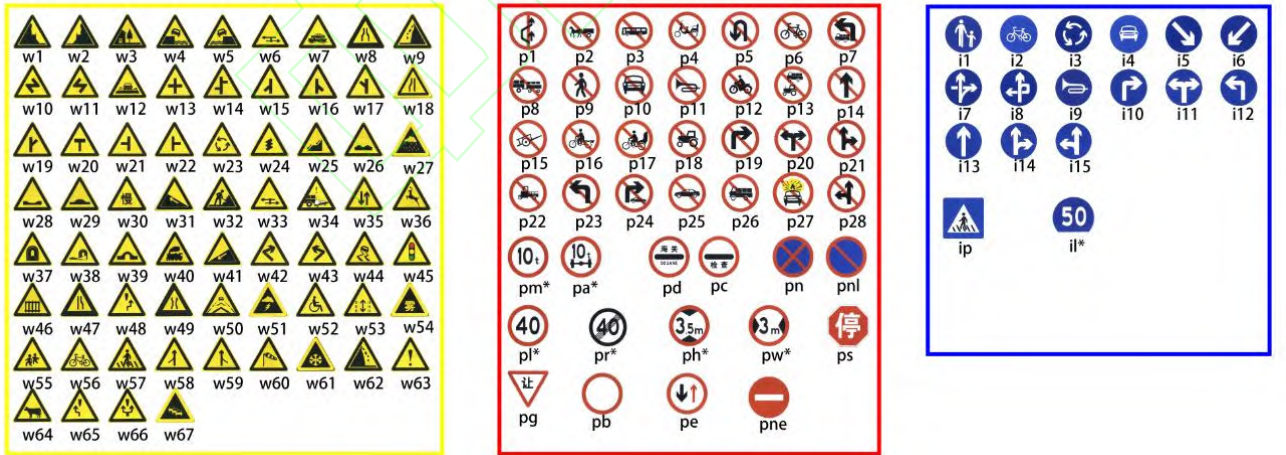


图 6 TT100K 交通标志分类

Fig.6 TT100K traffic sign category

2.3 评价指标及训练过程分析

本次实验的评价指标主要有精确率 Precision、召回率 Recall, 平均精度均值 mAP。定义如下:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{FN + TP} \quad (9)$$

$$AP = \int_0^1 P(R) dR \quad (10)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i) \quad (11)$$

式中, TP 表示将正样本正确预测为正样本的数量, FP 表示将负样本错误预测为正样本的数量, FN 表示将正样本错误预测为负样本的数量, mAP 值是不同类别精度的平均值, n 表示类别总数, i 代表第 n 个类别, AP 值是某一类的平均精度, 定义为不同召回率下的平均精度。

如图 7 所示, 为 MAFM-YOLO 模型训练曲线, box_loss 、 obj_loss 和 cls_loss 都分别呈现从下降到趋于平缓的趋势, mAP 稳步上升并逐步趋于稳定, 可以看出在训练 150 轮之后模型达到收敛。

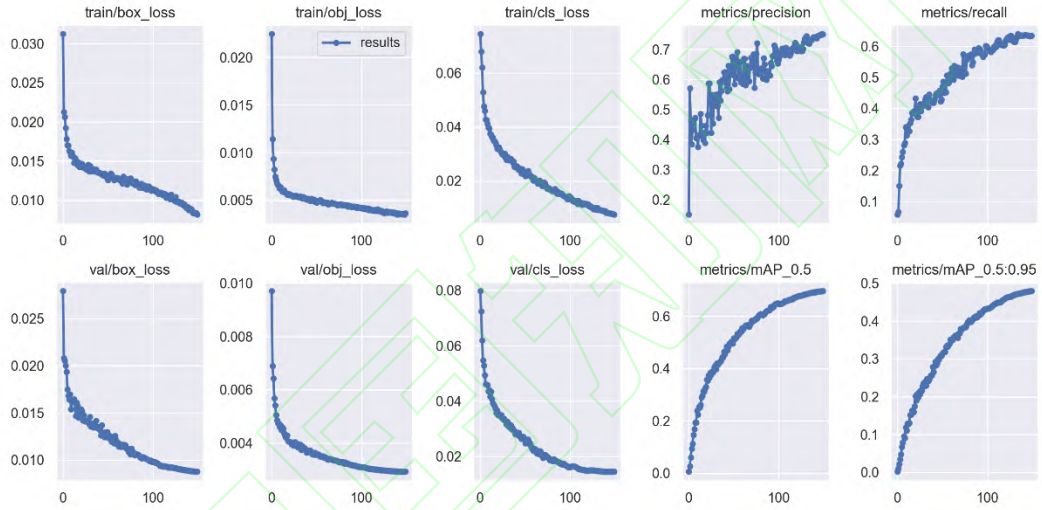


图 7 MAFM-YOLO 模型训练曲线

Fig.7 MAFM-YOLO model training curves

2.4 实验结果及分析

2.4.1 消融实验

为了验证 MAFM-YOLO 算法中提出的 MAFM 模块和 NWD 损失函数的有效性, 以 YOLOv5s 为基准模型进行下列的消融实验。(1) 为验证 MAFM 模块的有效性, 将 MAFM 加入到 YOLOv5s 的特征融合模块中。(2) 为验证 NWD 损失函数的有效性, 将 NWD 损失函数加入到 YOLOv5s 原有的损失函数中。(3) 将 MAFM 模块和 NWD 损失函数同时加入到 YOLOv5s 中的相应位置, 实现 MAFM-YOLO 算法。在相同的条件下以 Precision、Recall 和 mAP 为指标进行实验, 结果如表 1 所示。

表 1 消融实验

Tab.1 Ablation experiment

MAFM	NWD	Precision %	Recall %	mAP %
×	×	66.9	64.6	68.5
√	×	69.6	66.5	71.9
×	√	73.5	67.4	73.1
√	√	76.3	67.9	74.8

由表 1 实验结果可知, 对比基准模型 YOLOv5s, 实验 (1) Precision、Recall 和 mAP 分别提高了 2.7%、1.9%、3.4%; 这说明设计的 MAFM 模块能够有效提高模型的多层次信息融合能力, 特别是充分利用了浅层的

特征信息，降低复杂背景信息的干扰。实验（2）三项指标分别提高了 6.6%、2.8%、4.6%，这是因为采用 NWD 优化损失函数能够克服原有小目标识别存在的缺陷，进一步提高了网络对于小目标的定位效果，提高对小目标的识别能力。实验（3）将 MAFM 模块与 NWD 损失函数同时加入 YOLOv5s 中，进一步的提升网络的识别性能，三项指标分别提高了 9.4%、3.3%、6.3%。MAFM-YOLO 在所有指标上都优于原始 YOLOv5s 模型，证明了本文改进算法对交通标志识别的有效性。

2.4.2 注意力模块对比实验

为了验证空洞混合注意力机制 ABAM 的有效性，将 MAFM 中的 ABAM 模块分别用经典注意力机制（SE^[14]、CBAM^[15]、ECA^[16]、CA^[17]、NAM^[18]）进行替换，实验参数与本文模型参数保持一致，在 TT100K 数据集上进行测试，以 Precision、Recall 和 mAP 为指标，结果如表 2 所示。

表 2 对比实验			
Tab.2 Contrast experiment			
模块	Precision %	Recall %	mAP %
SE	74.5	66.3	72.8
CBAM	75.7	65.8	72.9
ECA	73.1	66.7	72.3
CA	73.3	67.3	72.4
NAM	75.6	65.8	73.5
MAFM	76.3	67.9	74.8

由表 2 可知，在网络中加入设计的 ABAM 注意力机制，相较于经典注意力机制 SE、CBAM、ECA、CA、NAM 注意力机制，在 Precision 上分别提高了 1.8%、0.6%、3.2%、3.0%和 0.7%，在 Recall 上别提高了 1.6%、2.1%、1.2%、0.6%和 2.1%，在 mAP 上分别提高了 2.0%、1.9%、2.5%、2.4%和 1.3%。本文设计的 ABAM 注意力机制相较于其他注意力机制取得了最好的效果，这说明 ABAM 能够有效去除复杂背景干扰，更好地映射特征之间的联系，提高交通标志目标识别的效果。

2.4.3 对比试验

为了验证改进算法 MAFM-YOLO 的先进性，与现有的主流算法 Faster R-CNN、YOLOv3-tiny、YOLOv4、YOLOX-s、YOLOv5、YOLOv7-tiny、YOLOv8n 和现有改进算法进行对比实验，实验在 TT100K 数据集上进行，实验参数与本文模型参数保持一致，差异参数则保持默认设置。以 Precision、Recall 和 mAP 以及每秒处理图像帧数（Frames Per Second, FPS）作为评价指标，实验结果如表 3 所示。

表 3 对比实验				
Tab.3 Contrast experiment				
模型	Precision/%	Recall/%	mAP/%	FPS
Faster R-CNN	66.4	60.3	63.4	83
YOLOv3-tiny	64.3	58.8	62.2	104
YOLOv4	66.7	61.1	64.3	114
YOLOX-s	68.2	60.5	66.4	131
YOLOv5s	66.9	64.6	68.5	161
YOLOv7-tiny	46.8	42.4	37.0	120
YOLOv8n	62.2	62.2	65.4	263
Ghost-YOLOv8	71.7	66.7	71.9	227
MAFM-YOLO	76.3	67.9	74.8	169

由表 3 可知，与先进的两阶段目标识别算法 Faster R-CNN 进行对比，MAFM-YOLO 在精确率、召回率、

mAP 上分别提升了 9.9%、7.6%、11.4%。与单阶段目标识别 YOLO 系列算法 YOLOv3-tiny、YOLOv4、YOLOX-s、YOLOv7-tiny、YOLOv8n 和现有基于单阶段改进算法 Ghost-YOLOv8 进行对比，MAFM-YOLO 在 Precision、Recall 和 mAP 上都取得了最好的性能。改进算法 MAFM-YOLO 在 FPS 上相较于原始 YOLOv5s 模型有所提升，尽管其 FPS 略低于 YOLOv8n 和 Ghost-YOLOv8 模型，但仍然保持在一个相对较高的水平，充分展现了其优化后的实时性能优势。这是由于在特征融合阶段设计的 MAFM 模块能够充分利用多尺度特征，去除背景信息。另外，NWD 损失函数的引用进一步提高了对小目标的识别性能，使得 MAFM-YOLO 在目标识别任务中能够在保持较高精度的同时，满足实时性要求。

2.4.4 可视化实验

为直观展示 MAFM-YOLO 算法性能，本文选取多种交通标志图像进行实验，包括远距离小目标、单一目标、多目标和遮挡目标，与原始 YOLOv5s 算法对比识别效果。如图 8 所示，第一行为原始 YOLOv5s 算法识别结果，第二行为 MAFM-YOLO 算法识别效果。原始 YOLOv5s 存在漏检（限速 30 p130 类）和误检（注意行人 w57 类误判为注意儿童 w55 类）问题。而 MAFM-YOLO 能准确识别 p130、w57 等交通标志，并保持高置信度，对多目标和遮挡目标也有良好识别能力。经过验证，本文提出的 MAFM-YOLO 算法有效解决了交通标志的漏检与错检问题，并显著提升了交通标志的识别精度。



图 8 识别效果对比
Fig.8 Comparison of recognition effect

3 结束语

本文提出了一种基于多尺度特征融合的交通标志识别算法 MAFM-YOLO，设计了基于空洞混合注意力机制的多尺度特征融合模块，并采用 NWD 损失函数对原始的 IoU 损失函数进行优化，能够有效融合多尺度特征信息，去除复杂背景干扰，进一步提高对小目标的识别能力。在 TT100K 数据集上进行测试，与原始的 YOLOv5s 模型相比，MAFM-YOLO 算法在精确率 P、召回率 R 和平均精度均值 mAP 上分别实现了 9.4%、3.3%和 6.3%的显著提升。与其他先进算法相比，MAFM-YOLO 算法也展现出了优越的性能，各项评价指标均有所上升，充分验证了其有效性和先进性。本文算法在交通标志识别领域展现出高精度的优势，未来将进一步探索收敛速度快且缓存优化友好的网络模型，以持续提升算法性能并优化资源利用。

参考文献:

[1] Lin Y H, Wang Y S. Modular Learning: Agile Development of Robust Traffic Sign Recognition[J]. IEEE Transactions on Intelligent Vehicles, 2023, 9(1): 764-774.
[2] 陈科圻, 朱志亮, 邓小明等. 多尺度目标检测的深度学习研究综述[J]. 软件学报, 2021, 32(4): 1201-1227.

- [3] 陈昌川, 王海宁, 赵悦等. 一种基于深度学习的交通标志识别新算法[J]. 电讯技术, 2021, 61(1): 76-82.
- [4] ZHANG J, XIE Z, SUN J, et al. A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection[J]. IEEE access, 2020, 8(1): 29742-29754.
- [5] Hu R, Li H, Huang D, et al. Traffic sign detection based on driving sight distance in haze environment[J]. IEEE Access, 2022, 10(1): 101124-101136.
- [6] WAN J X, DING W, ZHU H L, et al. An efficient small traffic sign detection method based on YOLOv3[J]. Journal of Signal Processing Systems, 2021, 93 (8): 899-911
- [7] 李志刚, 张娜. 一种轻量型 YOLOv5 交通标志识别方法[J]. 电讯技术, 2022, 62(9): 1201-1206.
- [8] Yar H, Khan Z A, Ullah F U M, et al. A modified YOLOv5 architecture for efficient fire detection in smart cities[J]. Expert Systems with Applications, 2023, 231(1): 120465-120476.
- [9] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 770-778.
- [10] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. (2016-04-30) [2024-03-18]. <https://arxiv.org/abs/1511.07122.pdf>.
- [11] Yang B, Zhang X, Zhang J, et al. EFLNet: Enhancing Feature Learning Network for Infrared Small Target Detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62(1): 1-11.
- [12] 董刚, 谢维成, 黄小龙等. 深度学习小目标检测算法综述[J]. 计算机工程与应用, 2023, 59(11): 16-27.
- [13] ZHU Z, LIANG D, ZHANG S, et al. Traffic-sign detection and classification in the wild[C]//Proceedings of 2016 IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 2110-2118.
- [14] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of 2018 IEEE conference on computer vision and pattern recognition. Salt Lake City: IEEE, 2018: 7132-7141.
- [15] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of 2018 European conference on computer vision. Munich: Springer Science+Business Media, 2018: 3-19.
- [16] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of 2020 IEEE conference on computer vision and pattern recognition. Virtual: IEEE, 2020: 11534-11542.
- [17] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of 2021 IEEE conference on computer vision and pattern recognition. Virtual: IEEE, 2021: 13713-13722.
- [18] Chen Z, Shi W, Zhang X, et al. Temporal Self-Attentional and Adaptive Graph Convolutional Mixed Model for Sleep Staging[J]. IEEE Sensors Journal, 2024, 24(8): 12840-12852.
- [19] 熊恩杰, 张荣芬, 刘宇红等. 面向交通标志的 Ghost-YOLOv8 检测算法[J]. 计算机工程与应用, 2023, 59(20): 200-207.

作者简介:

第一作者: 余翔, 男, 1969 年生于重庆, 2004 年获硕士学位, 教授, 主要研究方向为网络协议与安全、无线通信技术、移动边缘计算、计算机视觉。

第二作者: 靳闪闪, 女, 1999 年生于河南洛阳, 硕士研究生在读, 主要研究方向为计算机视觉。

第三作者: 杨路, 女, 1969 年生于重庆, 2011 年获硕士学位, 高级工程师, 主要研究方向为: 移动边缘计算、深度学习。