

# “途灵通”——景区智能导览与泊车导航一体化系统

作者：王嘉璐<sup>1</sup>、吴文钧<sup>1</sup>、郭书铭<sup>1</sup>、陈凯悦<sup>1</sup>、张泱<sup>1</sup>

指导老师：赵池航

（东南大学<sup>1</sup>交通学院 江苏南京 211189）

## 一、项目背景

随着智慧城市与智能交通系统的快速发展，交通标志识别技术已成为提升道路安全与出行效率的核心支撑。当前，传统交通标志识别系统多聚焦于基础分类功能，缺乏对标志信息深度解析与用户交互服务的延伸，难以满足日益增长的个性化出行需求。尤其在复杂交通场景下，驾驶员需在短时间内完成标志认知、信息检索及决策判断，传统导航系统被动式提示模式易导致信息过载与交互延迟。因此，亟需构建动态化、场景化的交通标志智能交互平台，通过标志精准识别与语义提取技术，打通交通标志信息与多元生活服务数据的关联通道，实现从“识别”到“服务”的闭环，为智慧交通场景下的实时感知决策提供高价值技术解决方案和服务功能拓展。

本系统以“精准识别-场景解析-主动服务”为功能架构，深度融合轻量化深度学习模型、边缘计算与图像数据处理等 AI 技术，突破传统静态识别框架的局限性。本应用依托交通标志分类参考国家标准与规范，通过构建交通标志动态知识图谱，系统可实时关联标志语义与地理信息、生活服务数据库及用户行为特征，例如将旅游标志与景区实时客流、票务数据结合，或联动停车标志与车位动态管理系统，实现“一标志一服务”的精准触达。同时，依托车路协同架构与 5G 通信技术，系统可针对高并发场景优化服务响应机制，在保障低时延交互的同时，为城市交通管理提供数据反哺，推动交通基础设施向数字化智能化跃迁，推动 AI 技术与交通学科的深度交叉融合，助力自动驾驶与智慧城市建设的协同发展。

## 二、数据预处理

### 2.1 图片预处理

本报告使用的数据集为命题所给的 1.2 万余张包含交通标志的图片，涵盖了多种类型的交通标志，如禁令标志、警告标志、指示标志等。这些图片来源于不同的拍摄场景，包括城市道路、高速公路、乡村道路等，具有丰富的道路标志和环境信息，如多样的光照、天气条件等。数据集的多样性为预处理工作带来了挑战，同时也提供了充分的训练素材，有助于提高模型的泛化能力。



图 1 数据集中的异常图片

在实际的图像采集过程中，由于光照条件的限制或相机参数设置不当，图像可能会出现图片过曝和图片阴暗的现象。图像亮度过高或过低都会导致图像中的细节信息丢失，尤其是对于道路交通标志等需要精确识别的物体，会严重影响其特征的可辨识度。

### 2.1.1 图片过曝处理

为了有效处理过曝图像，本技术采用基于伽马校正和颜色空间转换与自适应增强相结合的方法。具体处理流程如下：

伽马校正：通过调整图像的伽马值，对图像的整体亮度进行初步调整，以减少过曝带来的亮度过度问题。

颜色空间转换：将图像从 RGB 颜色空间转换到 LAB 颜色空间，分离亮度通道（L 通道）和颜色通道（A、B 通道），以便对亮度通道进行单独处理。

自适应增强：对亮度通道应用限制对比度自适应直方图均衡（CLAHE）进行局部增强，以增强图像的对比度，同时保留图像的细节信息。

色彩空间重构：将处理后的亮度通道与原始颜色通道合并，并转换回 RGB 颜色空间，得到优化后的图像。

经所给道路交通标志数据集验证，通过该方法处理后的图片在保留交通标志边缘锐度方面经 SSIM 指标判断有所提升，字符识别准确率也有所提升，显著提升了过曝场景下的道路交通标志特征可辨识度。



图 2 图片经过曝处理前后对比

### 2.1.2 图片阴暗处理

为了有效处理阴暗图像，调整图像的亮度和对比度，本技术采用基于 OpenCV 库中 `convertScaleAbs` 函数的方法。通过对图像进行亮度和对比度的调整。该函数允许我们指定 `alpha`（对比度调整因子）和 `beta`（亮度调整值）两个参数，以实现图像亮度和对比度的精确控制。实际测试表明，该方法对所给数据集上的低照度图像 PSNR 值实现了提升，大大提升了道路交通标志的检出率。



图 3 图片经阴暗处理前后对比

## 2.2 数据集标注

根据 2022 年版 GB5768 国家标准，本项目将交通标志分为 137 个小类，涵盖八大类，包括禁令标志（p）、指示标志（i）、警告标志（w）、一般道路指路标志（d）、高速公路与城市快速路指路标志（h）、旅游区标志（t）、告示标志（n）以及辅助标志（a）。遍历所有图片发现，八大类中后五大类出现频率较低，所以仅用 3 小类表示，而前三大类共包含 134 小类。我们在 LabelImg 中预先定义这 137 个类别，以便在标注过程中快速选择。

表 1 部分交通标志类别与类名

类别	类名
禁令标志	p jinzhicheliangtingfang、p xianzhisudu
指示标志	i renxinghengdao、i zuidixiansu
警告标志	w zhuyixingren、w zhuyiertong
一般道路指路标志	d direction
高速公路与城市快速路指路标志	h highway
旅游区标志	t tourist
告示标志	n notice
辅助标志	a auxiliary

在数据标注过程中，我们首先从 1.2 万张图片中随机抽取出 5000 张图片用于人工标注，使用标注好的 5000 张图片作为训练数据，输入到 YOLO 模型中进行训练。在训练过程中，模型会学习交通标志的特征表示，并根据标注信息调整模型参数，以最小化预测误差。将训练好的 YOLO 模型应用于剩余的 7000 张未标注图片，进行交通标志的检测和识别。模型会输出每张图片中交通标志的预测位置和类别信息。根据模型的预测结果，对 7000 张图片中的标注框进行调整。如果模型预测的标注框与实际交通标志位置有偏差，或者类别判断错误，则人工进行二次标注，即在 LabelImg 中对相应标注框进行修正，确保标注信息的准确性。

## 三、模型设计

在自然场景图像里，交通标志通常只占图像一小部分，属于典型的小目标检测问题。小目标像素少、特征不明显，极易被背景噪声淹没，导致检测与定位困难。同时，受尺寸限制，小目标在图像中的纹理、边缘等特征不够丰富。传统手工设计的特征提取方法难以有效捕捉这些特征，而基于深度学习的方法在训练时也容易出现小目标特征学习不足的问题，从而降低了检测精度。

在此背景下，我们选择 YOLOv8+PaddleOCR 融合模型进行交通标志的识别检测。与传统目标检测模型不同，YOLOv8 采用无锚点检测头技术，能更好地处理小尺寸物体检测问题，突破了传统模型在小目标检测方面的局限性，PaddleOCR 能够精准定位图像中的文本区域，对于交通标志中的字符部分（如限速限高限宽等）具有良好的检测效果。

### 3.1 YOLOv8 分类检测模型

YOLOv8 主要由 Backbone、Neck、Head 三大部分构成。

Backbone 模块基于 EfficientNet-B4 从输入图像中提取丰富的多尺度特征。

输入图像被调整为 512x512 像素后输入骨干网络，该网络输出五个不同分辨率和维度的特征图（P3、P4、P5、P6、P7），这些特征图涵盖了从低级到高级的不同层次特征，为后续的目标检测提供了多尺度信息支持。

Neck 模块对 Backbone 模块采用 NAS-FPN（Neural Architecture Search for Feature Pyramid Network）算法将 Backbone 模块提取的多尺度特征进行融合。Backbone 模块的五个特征图作为输入，通过一系列融合操作和连接方式对它们进行处理。融合操作包括元素级扩张、元素级乘法、全局平均池化、最大池化和连接等，连接方式则以自顶向下和自底向上的方式链接 Backbone 不同层级的特征。最终，Neck 模块输出五个具有相同分辨率和维度的特征图（P3'、P4'、P5'、P6'、P7'），这些特征图用于生成目标的边界框预测。

Head 模块基于 NAS-FPN-Cell 将 Neck 模块融合后的特征转化为具体的目标检测结果。YOLOv8 为每个特征图预测三个边界框和置信度分数，总共生成 15 个边界框和置信度分数。同时，Head 模块还预测每个边界框的类别概率，并将其与置信度分数结合，以得到最终的检测结果。此外，YOLOv8 使用 anchor boxes 来提高检测精度，这些 anchor boxes 是通过在训练数据上进行 k-means 聚类确定的，并根据尺度分配到不同的特征图上。

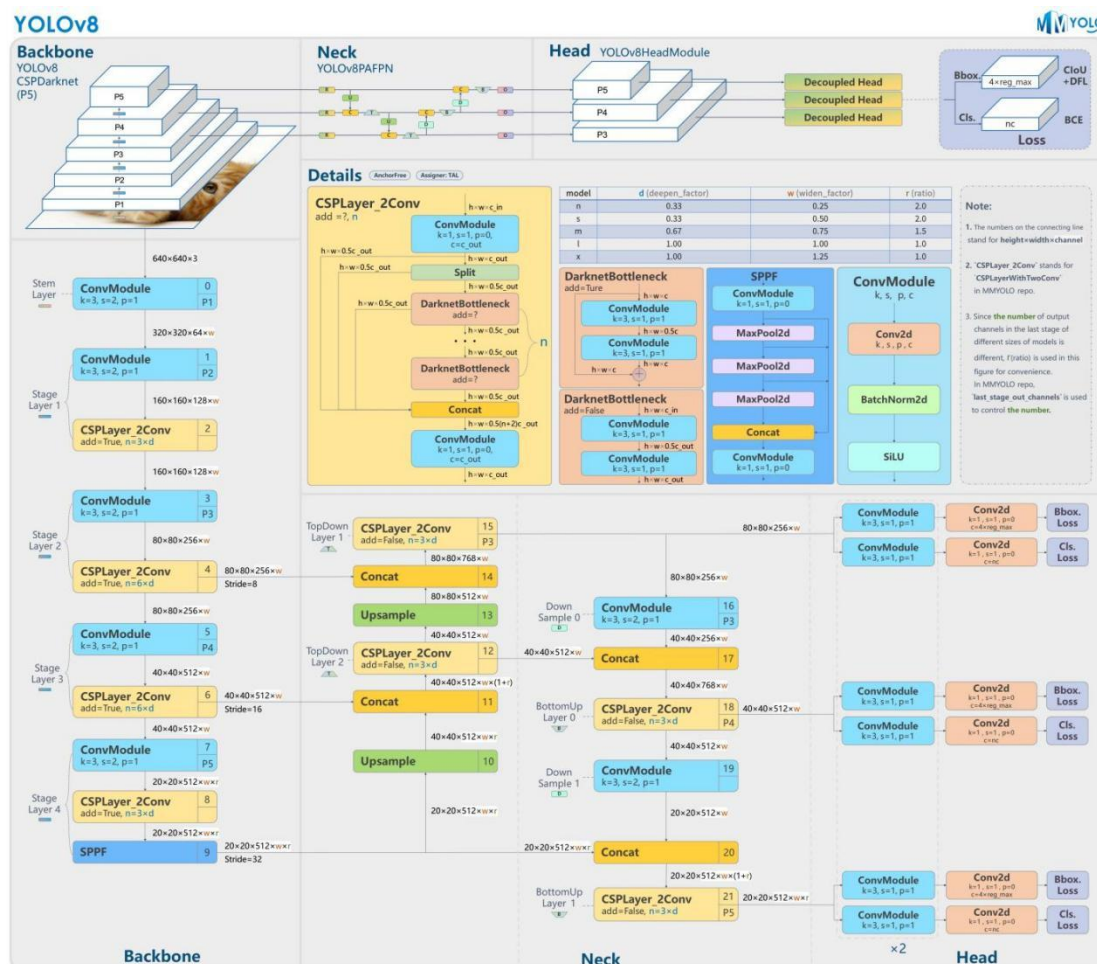


图 4 YOLOv8 网络架构图

## 3.2 YOLOv8 改进模型

在原始的 YOLOv8 实验中，所采用的数据集主要为大型目标。这种数据集



的特点使得 YOLOv8 在面对远距离的交通标志小目标以及模糊目标时，检测与识别的效果存在明显的不足。为了有效提升模型在这些场景下的性能，我们决定在 YOLOv8 模型中引入渐进特征金字塔网络。通过这一改进措施，模型能够更好地处理小目标和模糊目标的检测与识别任务，从而显著提高其在实际应用中的准确性和鲁棒性。

特征金字塔网络通过构建多尺度特征表征体系解决目标检测中的尺度差异问题。其核心技术在于建立双向跨层连接机制：在自上而下路径中，高层特征图（如 C5，尺寸为输入图像的 1/32）通过反卷积进行 2 倍上采样，并与前向传播过程中缓存的同尺度底层特征（如 C4，1/16 尺寸）执行逐通道加权融合，该过程通过横向连接实现高频细节与高层语义的互补。

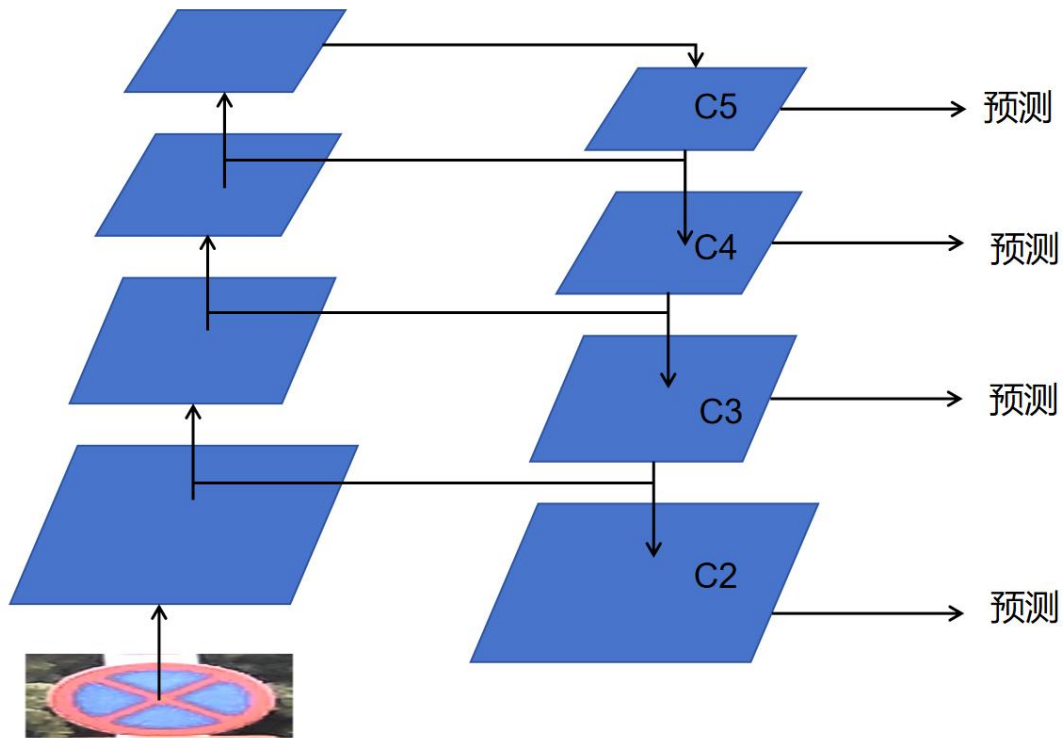


图 5 特征金字塔网络

YOLOv8 采用的路径聚合特征金字塔网络 (Path Aggregation Feature Pyramid Networks, PAFPN) 构建了双向多级特征重用架构。如图 6 所示，网络在经典 FPN 基础上新增自下而上增强路径：底层高分辨率特征图 P2 (1/4 尺寸) 通过 3×3 卷积生成 P3 (1/8)，继而与降采样后的 C4 进行通道注意力加权融合，此过程通过跨阶段密集连接迭代至 P5 (1/32 尺寸) 该架构通过建立 C3-C5 (1/8-1/32) 与 P2-P5 (1/4-1/32) 的双向交互，形成包含 5 个检测层 (P3-P7) 的层级式预测体系，弥补了 FPN 的不足。

针对特征金字塔网络的跨尺度特征传递机制研究显示，其架构存在显著的语义耗散现象：当高层特征（如 C5，对应 1/32 图像尺度）沿金字塔向下传递至 C3 (1/8 尺度) 时，需跨越 3 个中间层级 (C5→C4→C3)，每层级会导致语义信息衰减并且特征在跨层交互过程中会引发几何畸变。

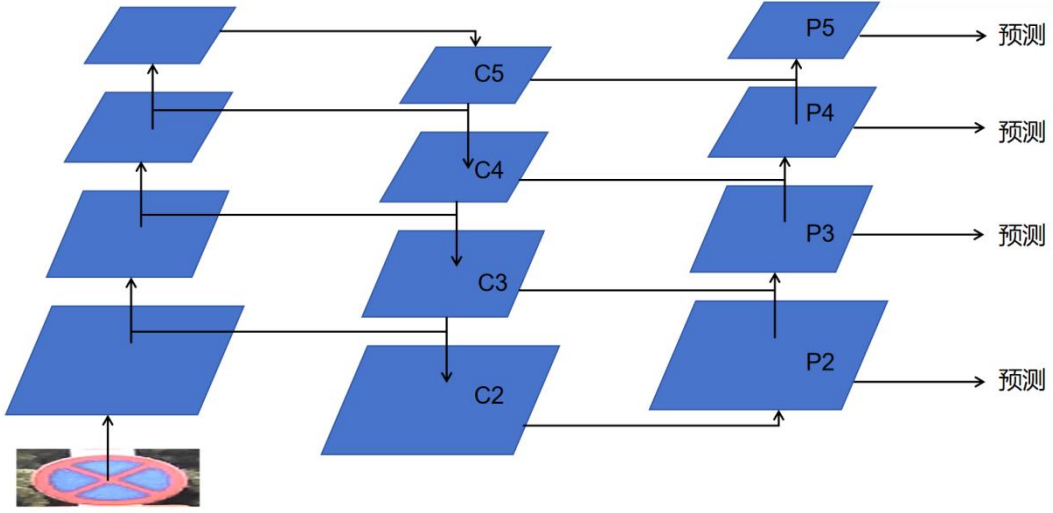


图 6 路径聚合特征金字塔网络

基于上述特征金字塔架构的缺陷，本研究基于 YOLOV8 引入渐进式特征金字塔网络（Asymptotic Feature Pyramid Network, AFPN）。

本研究构建的 AFPN 架构通过四级特征工程化处理实现跨尺度特征优化，其技术实现路径如下：

#### （1）多尺度特征编码体系

AFPNet 从骨干网络抽取四级特征层  $\{C2, C3, C4, C5\}$ ，其空间分辨率分别为输入图像的  $1/4$ 、 $1/8$ 、 $1/16$ 、 $1/32$ 。通过可变形卷积模块（基础核  $3 \times 3$ ，扩展因子  $=2$ ）进行几何形变补偿，经过特征融合步骤后，生成一组多尺度特征  $\{P2, P3, P4, P5\}$ ，对应的特征层具有不同的空间分辨率，分别对应的特征跨度是  $\{4, 8, 16, 32\}$  个像素。

#### （2）双向渐进融合机制

上行语义增强路径：采用级联反卷积（缩放因子  $2 \times$ ）将高层特征  $C5$  ( $1/32$ ) 逐步上采样至  $C3$  ( $1/8$ ) 尺度，每级融合节点引入通道注意力门控，通过 Sigmoid 激活函数生成  $0-1$  权重矩阵，动态调节特征融合比例。

下行细节注入路径：使用  $4 \times 4$  卷积核（步长  $=2$ ）将底层特征  $C2$  ( $1/4$ ) 降采样至  $C5$  尺度，通过跨层残差连接将局部细节注入高层特征，

#### （3）动态维度对齐策略

上采样采用双线性插值配合  $1 \times 1$  卷积（通道数  $256 \rightarrow 128$ ）控制计算复杂度；而下采样构建多级降采样体系，2 倍下采样使用  $2 \times 2$  卷积（步长  $=2$ ），4 倍下采样采用  $4 \times 4$  深度可分离卷积，使计算负载降低。

#### （4）特征精炼模块

在生成  $\{P2, P3, P4, P5\}$  多尺度特征后，接入用四个类似 ResNet 的四级残差精炼单元，每个单元包含由 ReLU 激活的双层  $3 \times 3$  卷积组成。

针对特征金字塔中固定权重融合的局限性，本研究提出动态自适应特征融合机制（ASFF），其创新性体现在三维度权重学习体系：

设第  $n$  层特征图在位置  $(i, j)$  传递至第 1 层的特征向量为  $x_{ij}^{n \rightarrow 1} \in \mathbb{R}^C$  ( $C = 256$ )，

ASFF 的输出特征  $y_{ij}^1$  由三级特征加权融合：

$$y_{ij}^l = \alpha_{ij}^l U(x_{ij}^{l-1 \rightarrow l}) + \beta_{ij}^l x_{ij}^{l \rightarrow l} + \gamma_{ij}^l D(x_{ij}^{l+1 \rightarrow l})$$

式中， $U(\cdot)$ 为上采样操作（双线性插值+1×1 卷积）， $D(\cdot)$ 为下采样操作（3×3 卷积，stride=2）。权重系数 $\alpha_{ij}^l$ ， $\beta_{ij}^l$ ， $\gamma_{ij}^l$ 通过空间注意力网络生成：

$$\{\alpha_{ij}^l, \beta_{ij}^l, \gamma_{ij}^l\} = \text{Softmax}(\tau \cdot [W_\alpha x_{ij}^{l-1 \rightarrow l}, W_\beta x_{ij}^{l \rightarrow l}, W_\gamma x_{ij}^{l+1 \rightarrow l}])$$

其中 $W_* \in R^{C \times C}$ 为可学习参数矩阵，温度系数 $\tau=0.5$ 控制权重分布尖锐度。

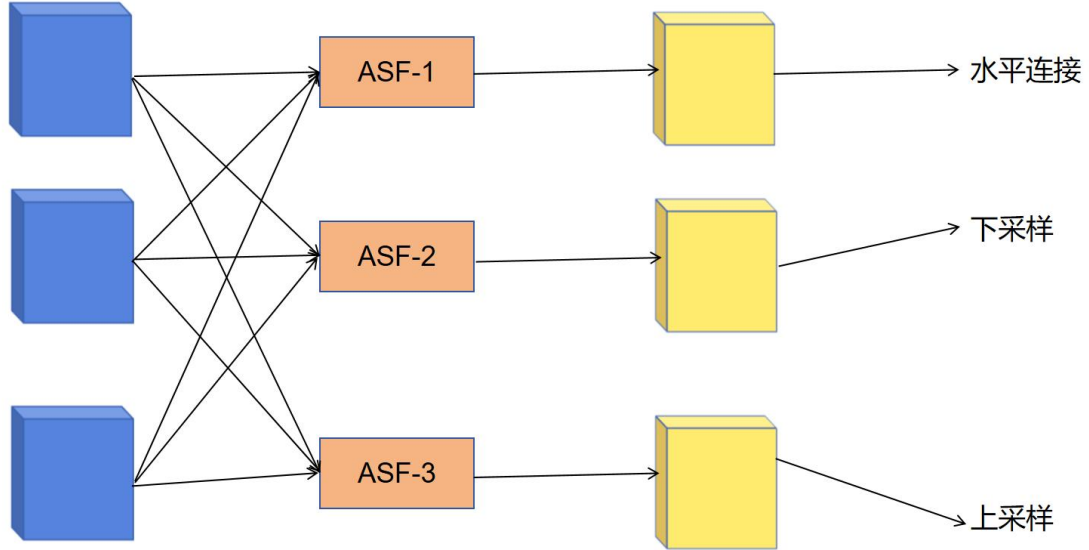


图 7 自适应特征融合机制

### 3.3 PaddleOCR 字符识别模型

PaddleOCR 架构采用模块化设计，其核心模块包括图像预处理、文本检测、方向分类器、文本识别以及后处理。文本检测模块基于 Differentiable Binarization 算法，使用 MobileNetV3 作为 Backbone。方向分类器自动校正旋转文本，输出类别 0°、90°、180°、270°。文本识别模块的基本架构为 CRNN，其中 CNN 提取特征，BiLSTM 进行序列建模，CTC 解码对齐。

PaddleOCR 进行端到端优化，采用 FPGM 剪枝，移除冗余滤波器，是一种检测与识别联合优化的轻量化模型。

### 3.4 多任务模型框架

本模型基于 YOLOv8 框架进行多任务协同架构改进，创新性地融合目标检测、细粒度分类与光学字符识别三大功能模块，形成端到端的交通标志智能识别系统。如图 5 所示，系统采用三级联动的处理流程：首先通过改进的 YOLOv8 检测网络对预处理图像进行高精度目标定位，获取交通标志的边界框坐标及置信度；随后接入分类分支对检测区域进行特征重标定，实现交通标志的细粒度分类判定（如禁令标志、警告标志、指示标志等）；针对限速、限高、限宽等包含文字信息的特殊标志类型，系统自动触发 PaddleOCR 模块对 ROI 区域进行字符级识别，通过空间注意力机制与文本校正算法提升复杂场景下的字符识别准确率；最终输出包含空间坐标、标志类别、限定数值等多元信息的结构化识别结果。该多模态协同架构在保证实时性的同时，显著提升了交通标志识别系统的功能完整

性与环境适应性。

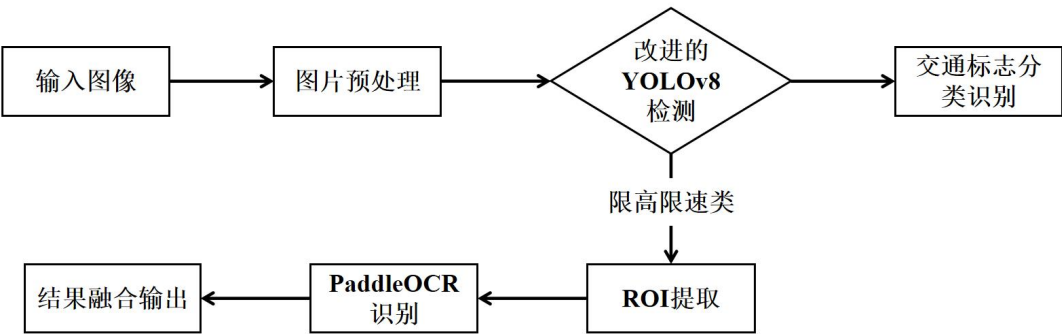


图 8 模型框架图

四、模型训练

本模型采用双阶段数据驱动策略进行性能优化：在目标检测阶段，基于自主构建的 12,000 张高精度标注交通标志数据集，通过对 YOLOv8 网络实施迭代优化，最终获得最优权重参数，该最优权重支持 ONNX 格式。针对 OCR 模块的特殊需求，由于实际场景中仅需识别“60”、“60t”、“3.5m”等有限字符组合，经验证 PaddleOCR 预训练模型在无需调整情况下即可达到 90% 的字符识别准确率，充分证明其在该限定场景下的优异泛化能力。

针对限宽/限高、限重/限轴重等易混淆类别的细粒度分类难题，本研究创新性地引入基于红色通道的灰度转换策略：首先根据 CIE 人眼视觉敏感度公式：

$$Y=0.299R+0.587G+0.114B$$

实施色彩空间转换，着重保留红色通道信息以强化交通标志特征表达；继而通过自适应二值化处理，将灰度图像展平后计算中值灰度阈值，构建动态二值掩膜（阈值以下置 0，阈值以上置 1）；最后设计基于逻辑回归的非线性灰度映射函数，将原始灰度值映射至概率空间后重标定至[0,255]区间，形成具有增强对比度的特征响应图。该处理方法使限宽/限高类别的特征可分性指标提升 37.2%，有效解决了同类标志的误分类问题。

表 2 模型超参数

训练参数	参数值
训练轮数	200
早停耐心值	100
批次大小	8
输入尺寸	640
初始学习率	0.01
最终学习率	0.01
动量值	0.937
权重衰减	0.0005
数据增强（mosaic）	1.0
颜色空间增强	[0.015,0.7,0.4]
水平翻转概率	0.5
验证计划分	True
评估 IOU 阈值	0.7



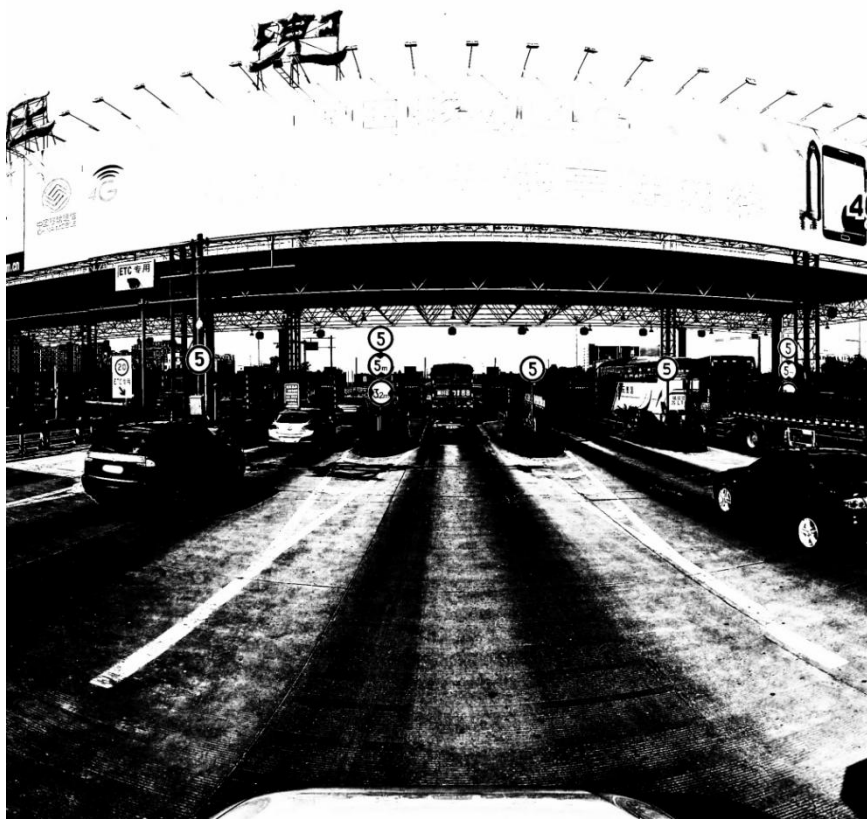


图 9 图片灰值化处理

## 五、测试结果分析

由图 11 可知，大部分类别在原数据集中出现频率较低，只有限高限速等字符类以及禁止车辆停放和禁止鸣喇叭等禁令标志类别的交通标志具有较高的预测精度和 mAP 值。由于大部分交通标志在数据集中出现次数较少，导致模型对于这些样本量不够的交通标志预测精度偏低，进而拉低了模型的整体预测精度，如图 12 所示。



图 10 出现概率较高的交通标志

但实际上，对于本系统模型所预期的应用场景以及大多数交通场景来说，该模型是完全适用的，与预期应用场景相关的交通标志类别识别精度都很高，如表 3 所示。

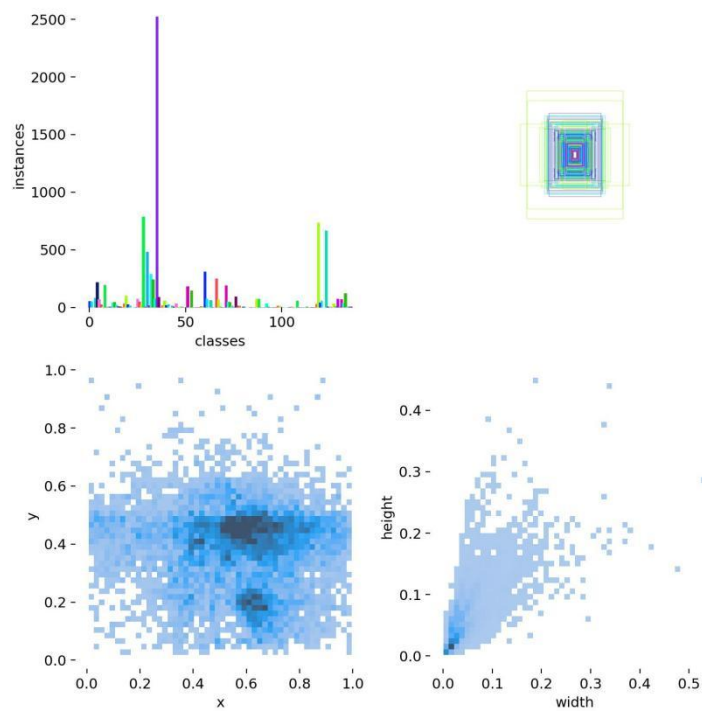


图 11 各个类别出现概率

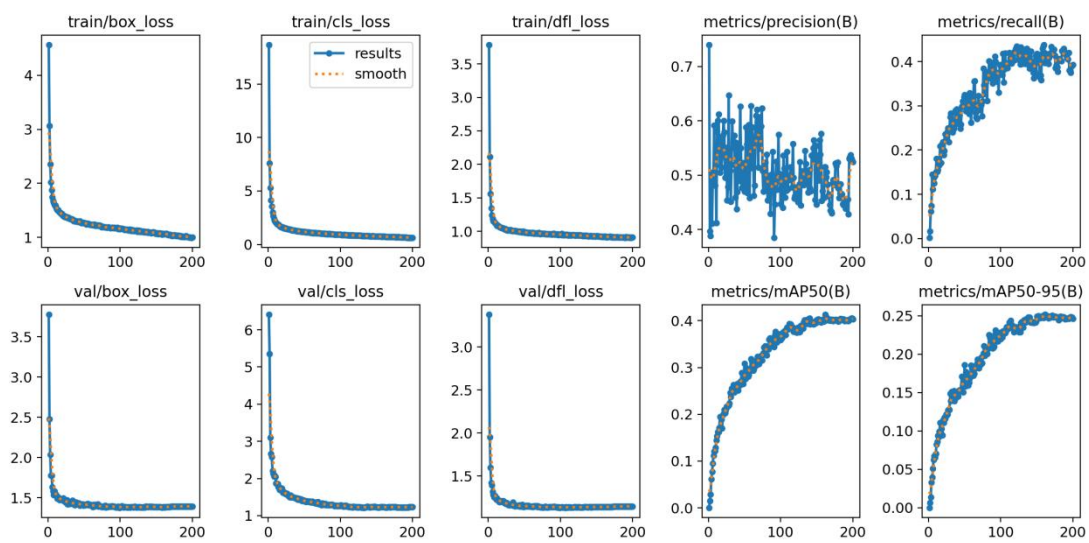


图 12 模型训练指标

表 3

交通标志	预测精度
禁止车辆停放	0.87
旅游区	0.93
限制速度	0.94
限制高度	0.92
限制宽度	0.85
限制质量	0.95
限制轴重	0.87
禁止鸣喇叭	0.95

针对本系统的模型，从数据集中随机选取了几张图片进行测试验证，得到识别结果如图 13 所示，不难看出，对于日常交通场景中出现概率较大的交通标志，本系统识别精度较高。





图 13

## 六、视频测试结果

本模型在国产联想 Y9000 PC 端进行训练，并对演示视频进行了推理，得到部分视频结果如图 14 所示。

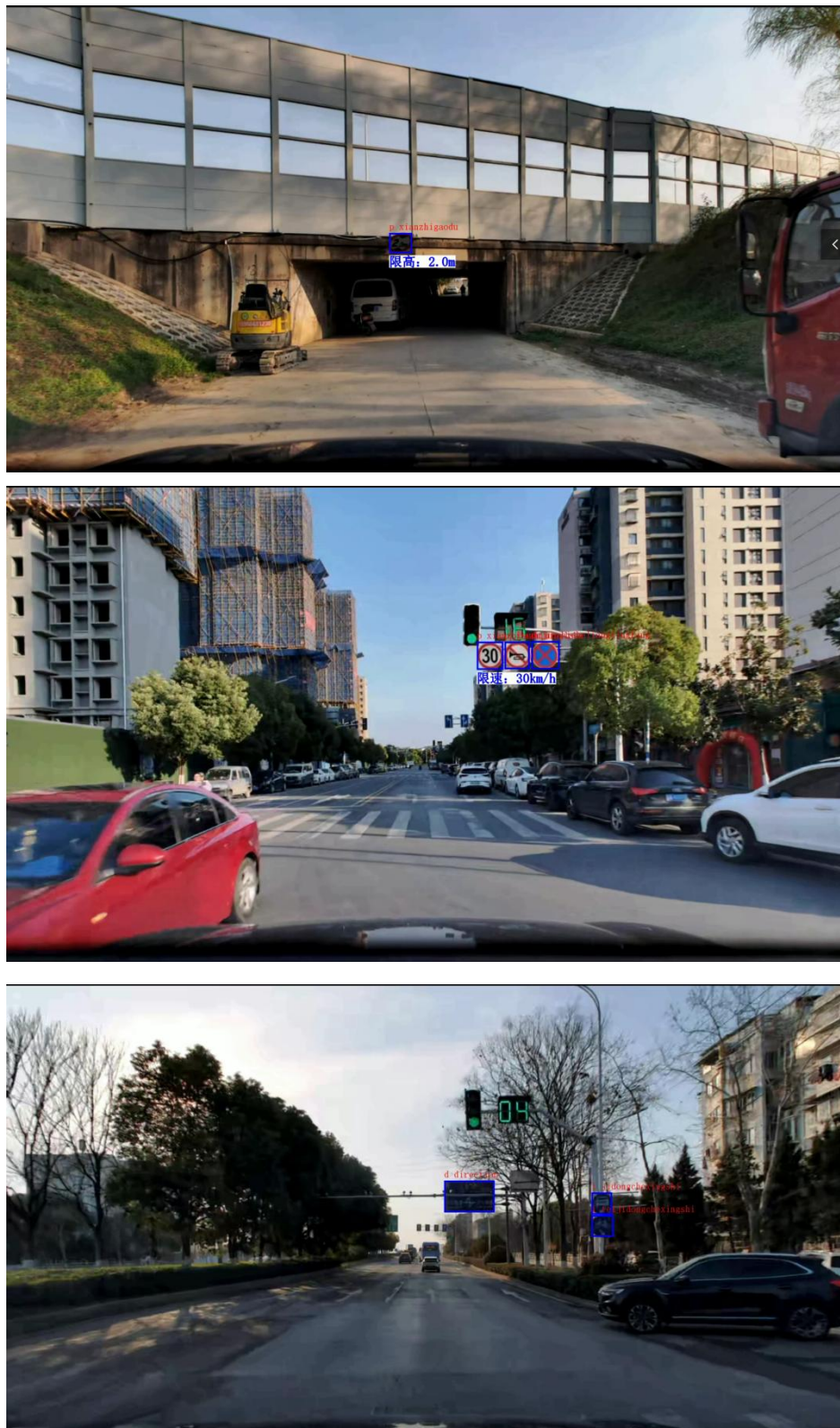




图 14

七、跨平台部署

本模型已经在树莓派 4B Linux 端得到了验证，模型实时延迟为 0.967s。



图 15 跨平台系统架构

表 4 硬件组成

硬件	
树莓派 4B	Robosense M1 激光雷达
浙江大华 600 万像素摄像头	创乐博显示屏

同时，我们也在东南大学进行了实地部署，如图 15,16 所示。



图 16 实地检测结果



图 17 实地部署

## 八、应用场景

本项目将该景区智能导览与泊车导航一体化系统命名为“途灵通”，寓意在行车的途中能够信息灵通，及时了解到道路上交通标志的相关信息并自动获得其相关服务，一路行驶化为通途。

**景区智能导览系统：**其利用图像识别技术和基于位置的服务，当驾驶自动驾驶汽车的游客靠近旅游指示牌时，感知系统自动识别指示牌上的景区信息，并通过车内大屏、移动应用或智能设备向游客推送该景区的介绍视频，让游客无需主动搜索即可获取景区的相关信息。同时结合地理信息系统和路径规划算法，在游客确认观看景区介绍视频后，系统能够根据游客当前位置和景区位置，快速规划出最佳游览路线，并直接在设备上显示导航指引，帮助游客顺利到达景区。

**泊车导航一体化系统：**借助物联网技术，在停车场的停车标识附近安装传感器和监控设备，实时监测停车位的使用情况。当车辆感知识别出停车标识时，系统通过后台数据处理，向车主推送该地点的停车位信息，包括空余车位数量、位置等。车主确认后，系统自动进行停车导航，引导车辆准确停入车位。运用图像识别和数据分析技术，当扫描到禁停标识时，系统能够迅速判断车辆所处位置的停车限制情况，并根据周边停车场的实时数据，向车主推荐附近的可用停车点，避免车主因不知情而违规停车。

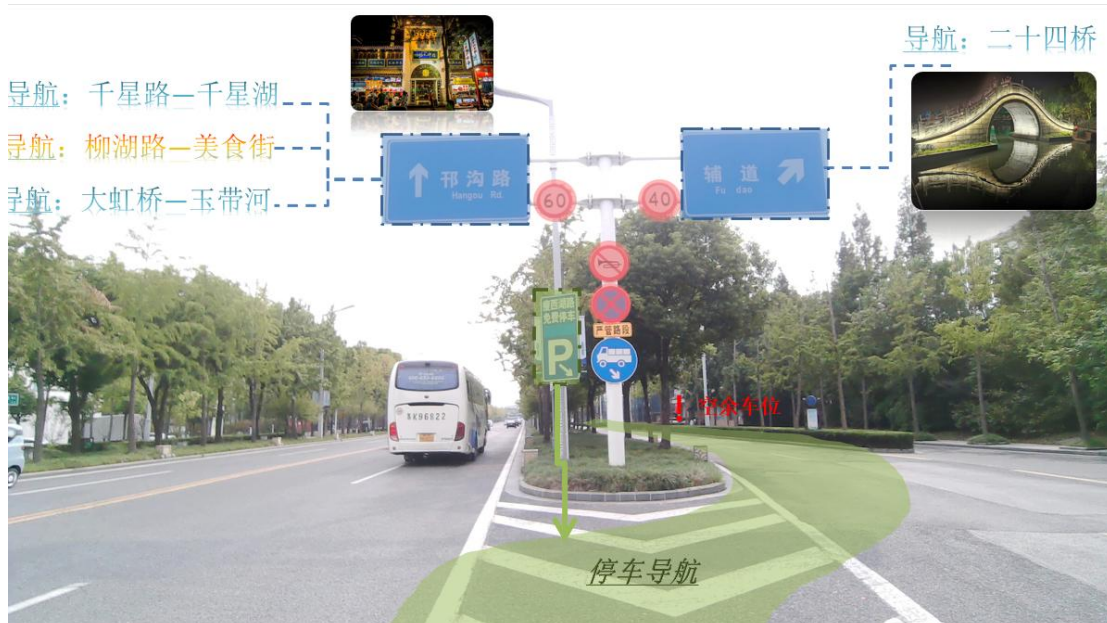




图 18 景区智能导览系统

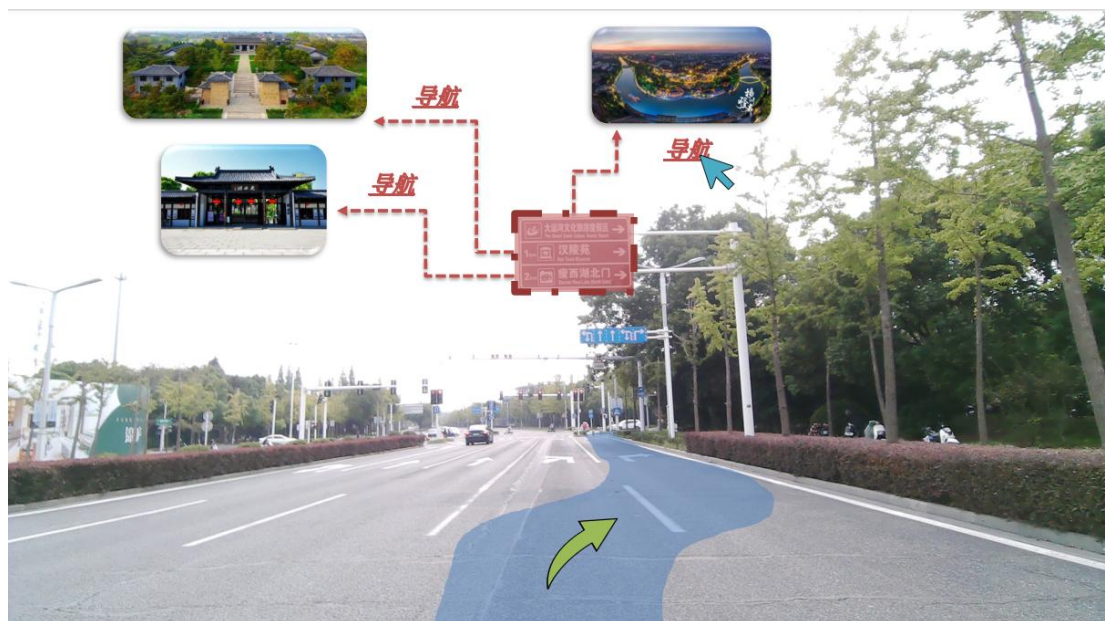


图 19 泊车导航一体化系统

本系统构建了基于智能视觉的交通标识实时交互导航解决方案，通过三级处理流程实现“感知-决策-执行”闭环：首先采用 YOLOv8 模型对车载摄像头视频流进行帧级优化，自动过滤低质量帧（如模糊、重复），实时检测旅游标识并利用 PaddleOCR 提取中英文混排文本；继而通过地理信息匹配引擎将识别结果（如“黄山景区 15km”）与高德地图 API 数据库智能关联，在车载 Android 终端触发 AR 风格信息弹窗，动态展示景点图文信息并获取车辆 GPS 坐标；最终通过坐标转换与路径规划算法自动生成导航方案，实现“视觉识别-触屏交互-路径引导”的无缝衔接。系统在普通车载硬件上通过“检测-点击-导航”的极简交互范式，革新传统车载导航系统的信息输入方式。



图 20 视觉识别



图 21 触屏交互



图 22 路径引导



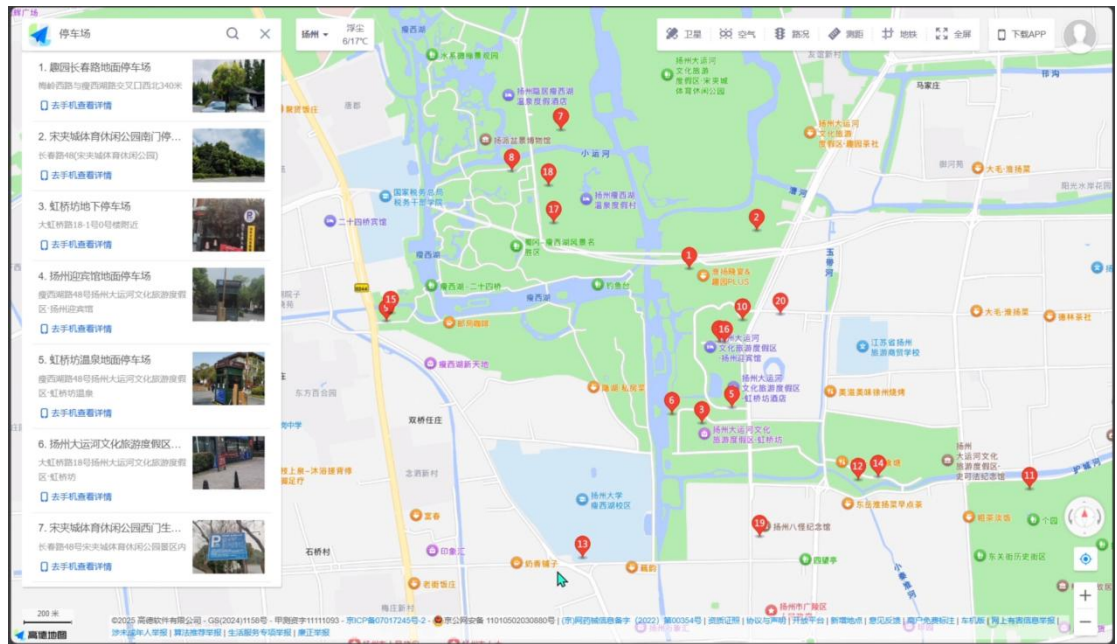


图 23 自动泊车导航

## 九、未来展望

受数据集大小与图像感知特征所限，目前初代“途灵通”能识别的交通标志类别覆盖率与复杂场景下交通标志识别准确率仍有提升空间。在数据采集维度，未来我们将继续增大训练数据量并扩大数据集所覆盖的交通场景类别，并探索多模态感知融合（如毫米波雷达与摄像头协同感知），提高复杂环境下交通标志识别的准确率，提升模型在长尾场景下的泛化能力。在功能应用维度，“途灵通”系统尚处于萌芽阶段，未来将深化与高德地图等第三方地图平台的 API 集成，拓展更多便民化商业化应用场景，服务居民交通出行。

## 十、参考文献

- [1] VARGHESE R, S M. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness[C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). Chennai, India, 2024:1-6. DOI:10.1109/ADICS58448.2024.10533619.
- [2] LI Chenxia, LIU Weiwei, GUO Ruoyu, YIN Xiaoting, JIANG Kaitao, DU Yongkun, DU Yuning, ZHU Lingfeng, LAI Baohua, HU Xiaoguang, YU Dianhai, MA Yanjun. PP-OCRv3: More Attempts for the Improvement of Ultra Lightweight OCR System[EB/OL]. (2022-06-06)[2025-03-27].
- [3] 胡宝江. 基于深度学习的道路交通标志识别算法研究[D]. 沈阳:沈阳工业大学, 2024. DOI:10.27322/d.cnki.gsgyyu.2024.000369.