·面向智能驾驶的目标检测与道路通行区域感知技术专题·

融合边界增强与几何–深度先验的非结构化道路可通行区域检测方法*

于文泰[1] 高明达[2] 王靖岳[3]

（1.辽宁理工学院汽车工程学院，锦州 121007；2.辽宁工业大学汽车与交通工程学院，锦州 121001；3.沈阳理工大学汽车与交通学院，沈阳 110159）

【摘要】针对非结构化道路场景中地表纹理复杂、边界模糊和光照变化明显等问题，提出一种融合边界增强与几何–深度先验注意力的仅红绿蓝图像输入（RGB-only）可通行区域检测方法。该方法以SegFormer为基础框架，通过边界辅助监督提取显式边界信息，并利用边界特征融合模块将边界线索注入语义特征空间，增强模型对软边界和不规则轮廓的表达能力。同时，在注意力模块引入几何连续性先验和垂直方向伪深度先验，对特征进行结构化重加权，以提升模型对复杂地形的预测稳定性。试验结果表明：所提出方法在越野可通行区域检测（ORFD）数据集上的交并比（IoU）和F1分数分别达到90.8%、95.1%，且主要分割指标上优于单目RGB方法，具有较好的跨数据集泛化能力。

主题词：语义分割 可通行区域检测 非结构化道路 边界增强 几何–深度先验注意力

中图分类号：TP391.4 **文献标志码：**A **DOI:** 10.19620/j.cnki.1000-3703.20260254

An Off-Road Freespace Detection Method Based on Boundary Enhancement and Geometry-Depth Prior

Yu Wentai[1], Gao Mingda[2], Wang Jingyue[3]

（1. School of Automobile Engineering, Liaoning Institute of Science and Engineering, Jinzhou 121007; 2. School of Automobile and Traffic Engineering, Liaoning University of Technology, Jinzhou 121001; 3. School of Automobile and Transportation, Shenyang Ligong University, Shenyang 110159）

【Abstract】 To address the challenges of complex surface textures, blurred boundaries, and significant illumination variations in unstructured road scenarios, this paper proposes a freespace detection method using only red-green-blue (RGB) images and integrating boundary enhancement and geometry-depth prior attention. Based on the SegFormer framework, the proposed method extracts explicit boundary information through auxiliary boundary supervision and introduces boundary cues into the semantic feature space via a boundary feature fusion module, thereby enhancing the representation of soft boundaries and irregular contours. Meanwhile, a geometry-depth prior attention module is designed to incorporate geometric continuity priors and vertical pseudo-depth priors, enabling structured feature reweighting and improving prediction stability in complex terrain. Experimental results show that the proposed method achieves an IoU of 90.8% and an F1-score of 95.1% on the Off-Road Freespace Detection (ORFD) dataset, outperforming the compared monocular RGB methods on major segmentation metrics and demonstrating good cross-dataset generalization capability. These results indicate that the proposed method achieves a favorable balance among segmentation accuracy, boundary quality, and inference efficiency.

Key words: Semantic segmentation, Freespace detection, Unstructured road, Boundary enhancement, Geometry - depth prior attention

【引用格式】 于文泰，高明达，王靖岳．融合边界增强与几何–深度先验的非结构化道路可通行区域检测方法[J]. 汽车技术，2026(7): 11-19.

YU W T, GAO M D, WANG J Y. An Off-Road Freespace Detection Method Based on Boundary Enhancement and Geometry-Depth Prior[J]. Automobile Technology, 2026(7): 11-19.

*基金项目：辽宁理工学院开放合作科研课题（LNLGKFHZ03）；中国高校产学研创新基金项目（2024HT031）。

通信作者：王靖岳（1978—），教授，主要研究方向为车辆系统动力学与控制，abswell@126.com。

1 前言

可通行区域检测是自动驾驶与移动机器人系统中的关键感知任务，为路径规划与决策控制提供基础支撑[1-3]。现有研究集中于结构化道路场景，常依赖清晰的车道线、规则的道路边界以及相对稳定的路面纹理以实现高精度分割。然而，乡村土路、碎石路、林间道路及施工区域等大量场景均属于非结构化道路环境，此类场景存在道路边界模糊、路面形态不规则、地表纹理复杂和背景干扰明显等问题[4-6]。同时，遮挡频繁、光照变化剧烈等因素会进一步削弱边界可辨识性，使可通行区域与不可通行区域间呈现连续过渡的“软边界”特性[7]，从而引发边界漂移、轮廓断裂和误分割等现象。

在感知模态方面，基于单目红绿蓝（Red-Green-Blue，RGB）图像的方法利用全卷积网络（Fully Convolutional Network，FCN）、U-Net及DeepLabv3+语义分割网络等经典语义分割框架，在结构化道路场景中取得良好效果[8-10]，但在非结构化环境中易受到复杂纹理与光照退化的干扰，导致边界模糊、区域误判。基于激光雷达（Light Detection and Ranging，LiDAR）的方法利用点云的几何结构信息进行地面区域识别，对光照变化具有较强鲁棒性；使用圆柱形非对称三维卷积网络（Cylinder3D）建模点云的三维几何结构，为复杂环境下的LiDAR语义分割提供了有效思路[11]，但其受限于点云稀疏性与视场范围，在细节完整性与空间连续性方面仍存在不足。

随着RGB-LiDAR多模态融合方法逐渐成为研究热点，表面法向估计道路分割（Surface Normal Estimator Road Segmentation，SNE-RoadSeg）网络[12]引入表面法向信息，增强结构表达能力；位置嵌入变换第2版（Position Embedding Transformation v2，PETRv2）模型基于Transformer构建统一的空间表示，以提升复杂场景建模能力[13]。多模态特征融合网络（Multi-Modal Feature Fusion Network，M2F2-Net）和RGB-法向道路场景解析双路Transformer（Duplex Transformer for RGB-Normal Semantic Road Scene Parsing，RoadFormer）网络通过融合RGB外观信息与几何结构信息，提高了非结构化道路场景下的可通行区域检测效果[14-15]。尽管此类方法在多种数据集上取得了优异表现，但其对额外传感器和复杂预处理流程（如点云重投影与跨模态对齐）的依赖，大幅增加了系统成本与计算开销，在资源受限平台部署困难。

由于现有方法难以兼顾分割精度、系统成本与推理效率，仅RGB输入越野可通行区域检测（RGB-only fast and efficient Off-road freespace Detection，ROD）方法在不依赖LiDAR等额外传感器的情况下，可兼顾越野场景可通行区域检测的精度与推理效率[16]。然而，单目RGB图像缺乏显式几何与深度信息，对软边界和复杂地形结构的表征能力仍然不足。为此，本文基于SegFormer语义分割模型[17]，提出一种融合边界增强与几何-深度先验注意力、仅以红绿蓝图像（Red-Green-Blue-only，RGB-only）为输入的可通行区域检测方法。主要工作如下：

a. 设计边界增强分支，利用索贝尔（Sobel）算子构造伪边界监督，通过边界特征融合模块将显式边界线索注入语义特征，增强模型对软边界及不规则轮廓的表征能力。

b. 构建几何-深度先验注意力模块，联合利用边界几何连续性和垂直方向伪深度信息对融合特征进行结构化重加权，提升模型对复杂地形空间结构的建模能力。

c. 在越野可通行区域检测（Off-Road Freespace Detection，ORFD）数据集上开展对比试验与消融试验，并在RELLIS-3D数据集上进行跨数据集泛化测试，验证所提方法的有效性和泛化能力。

2 基本方法

2.1 问题定义

本文将可通行区域检测建模为3类像素级语义分割任务，已知输入RGB图像 $\boldsymbol{I}_{\mathrm{RGB}} \in \mathbf{R}^{H\times W\times 3}$，其中，$H$、$W$分别为图像的高度和宽度。对应的像素级标注为 $\boldsymbol{Y}^{\mathrm{gt}} \in \{1,2,3\}^{H\times W}$，其中，类别标签1、2和3分别表示可通行区域、不可通行区域和天空区域。使用分割网络对输入图像进行特征提取与像素级分类，输出预测概率张量 $\boldsymbol{P} \in [0,1]^{H\times W\times 3}$，其中，$\boldsymbol{P}_{i,j,c}$为像素$(i,j)$属于第$c$类的预测概率，且对任意像素$(i,j)$均满足 $\sum_{c=1}^{3}\boldsymbol{P}_{i,j,c}=1$，则最终预测类别可表示为：

$$\hat{\boldsymbol{y}}_{i,j}=\underset{c\in\{1,2,3\}}{\operatorname{argmax}}\left(\boldsymbol{P}_{i,j,c}\right) \tag{1}$$

在训练过程中，主分割分支采用多类交叉熵损失进行监督：

$$L_{\mathrm{seg}}=-\frac{1}{|\Omega|}\sum_{(i,j)\in\Omega}\sum_{c=1}^{3}\boldsymbol{y}_{i,j,c}^{\mathrm{gt}}\log\boldsymbol{P}_{i,j,c} \tag{2}$$

式中：Ω为参与训练的有效像素集合，$|\Omega|$为有效像素数量，$\boldsymbol{y}_{i,j,c}^{\mathrm{gt}}$为像素$(i,j)$在第$c$类上的one-hot形式真实标签。

由于在非结构化道路场景中，可通行区域边界常呈现软过渡和弱纹理特征，仅依赖区域分割损失难以获得精细的边界定位结果。为此，本文引入边界监督损失L_{edge}，并将其与主分割损失共同构成最终优化目标：

$$L = L_{seg} + \lambda L_{edge} \tag{3}$$

式中：λ为边界损失权重参数，用于调节区域分割损失与边界监督损失在总优化目标中的贡献比例。

当λ取值较小时，模型训练更偏向于区域语义分割，边界约束作用较弱，容易导致边界模糊；当λ取值过大时，模型会过度关注边界细节，从而影响主体区域的语义一致性。通过设置适当的λ权重，可以在区域完整性与边界精细度间实现有效平衡。

2.2 整体网络结构

本文提出的可通行区域检测网络整体结构如图1所示，该模型主要由混合Transformer-B4编码器（Mix Transformer-B4，MiT-B4）、统一感知解析头解码器（Unified Perceptual parsing Head，UPerHead）、边界增强分支、边界特征融合（Boundary Feature Fusion，BFF）模块和几何-深度先验注意力（Geometry-Depth Prior Attention，GDPA）模块组成。

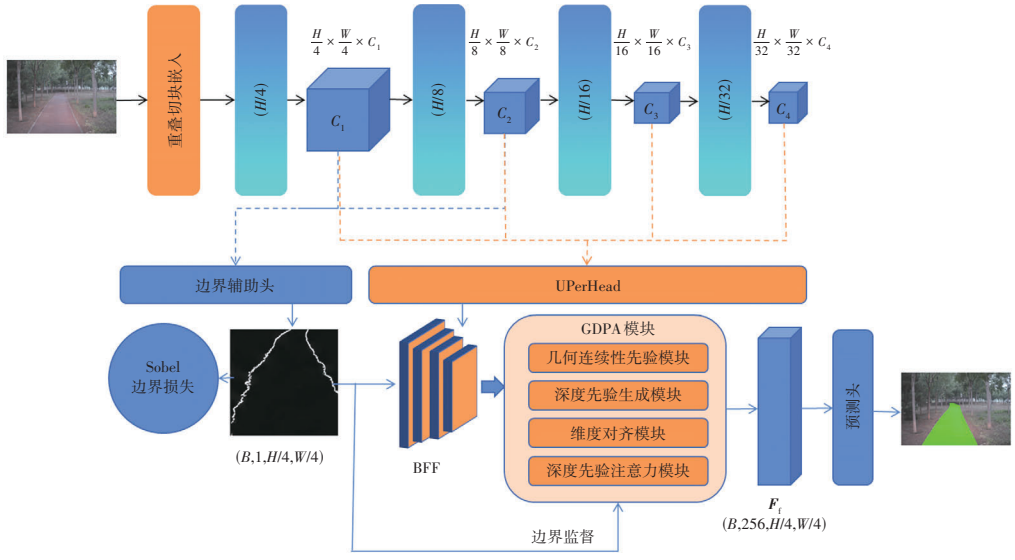


图1　本文网络架构

为了增强模型对软边界区域的感知能力，本文引入边界增强分支。该分支利用编码器浅层特征预测边界概率图$\boldsymbol{E}$，通过Sobel算子在分割真值中构造伪边界监督信号，并对边界预测结果进行约束，计算边界监督损失L_{edge}。同时，边界概率图可作为显式边界提示参与特征融合。首先，采用插值操作将$\boldsymbol{E}$调整至与语义特征图$\boldsymbol{F}_s$相同的空间分辨率；随后，将二者沿通道维度进行拼接，并通过1×1卷积进行特征映射，得到边界增强特征$\boldsymbol{F}_b$。该过程将显式边界线索注入语义特征空间，从而增强模型对可通行区域轮廓、软边界及不规则边界形状的表征能力。

最后，边界增强特征输入几何-深度先验注意力模块。该模块结合几何连续性先验、垂直方向伪深度先验以及空间-通道联合注意力机制，对特征进行结构化重加权。其中，几何连续性先验用于增强边界区域的空间结构一致性，伪深度先验用于提高模型对图像下方近处区域的关注程度，空间-通道联合注意力则进一步自适应突出有效区域和关键特征通道。经GDPA模块增强后的特征$\boldsymbol{F}_f$输入预测头，最终生成可通行区域、不可通行区域和天空区域的3类分割结果。

2.3 边界增强模块

为了有效提升模型在非结构化道路场景中对边界结构的刻画能力，基于SegFormer框架设计边界增强模块，主要包括边界辅助头、Sobel边界监督和边界特征融合。该模块通过显式引入边界约束，引导模型在学习区域语义信息的同时，关注可通行区域边缘，从而缓解非结构化道路中边界模糊、轮廓断裂和局部错分等问题。

由于浅层特征保留了较多纹理、边缘和局部结构信息，故更适用于边界预测。将编码器浅层特征C_1作为边界辅助分支的输入，经过1×1卷积降维和卷积映射后生成边界特征图，并进一步预测边界概率图$\hat{\boldsymbol{E}}$。采用Sobel算子构造伪边界标注进行边界预测监督。鉴于语义类别编号仅用于区分类别，直接对3类语义标签图进行梯度计算，类别编号差值对边界相应产生影响。因此，本文仅提取可通行区域边界，首先，将3类语义标签图映射为可通行区域二值掩码：

$$\boldsymbol{M}_{i,j}^{\mathrm{gt}}=\begin{cases}1, & \boldsymbol{Y}_{i,j}^{\mathrm{gt}}=1\\ 0, & \boldsymbol{Y}_{i,j}^{\mathrm{gt}}\in\{2,3\}\end{cases} \tag{4}$$

随后，分别采用水平和垂直方向的Sobel卷积核计算二值掩码的梯度，并根据各方向的梯度响应计算梯度幅值。对梯度幅值进行阈值二值化，得到伪边界标签为：

$$\begin{cases}\boldsymbol{G}_x=\boldsymbol{K}_x*\boldsymbol{M}^{\mathrm{gt}}\\ \boldsymbol{G}_y=\boldsymbol{K}_y*\boldsymbol{M}^{\mathrm{gt}}\\ \boldsymbol{E}_{i,j}^{\mathrm{gt}}=\mathbb{I}\left(\sqrt{\boldsymbol{G}_{x,i,j}^2+\boldsymbol{G}_{y,i,j}^2}>\tau\right)\end{cases} \tag{5}$$

式中：$\boldsymbol{K}_x$、$\boldsymbol{K}_y$分别为水平和垂直方向的Sobel卷积核，*为卷积运算，$\boldsymbol{G}_x$和$\boldsymbol{G}_y$分别为水平和垂直方向的梯度响应，τ为二值化阈值，$\mathbb{I}()$为指示函数。

边界预测分支采用二值交叉熵损失进行监督：

$$L_{\mathrm{edge}}=-\frac{1}{|\Omega_e|}\sum_{(i,j)\in\Omega_e}\left[\begin{array}{l}E_{i,j}^{\mathrm{gt}}\log\left(\hat{E}_{i,j}+\varepsilon\right)\\ (1-\boldsymbol{E}_{i,j}^{\mathrm{gt}})\log\left(1-\hat{E}_{i,j}+\varepsilon\right)\end{array}\right] \tag{6}$$

式中：Ω_e为参与边界损失计算的有效像素集合，$|\Omega_e|$为有效像素数量，$E_{i,j}^{\mathrm{gt}}$和$E_{i,j}$分别为像素(i,j)处的伪边界标签和预测边界概率，ε为防止数值计算不稳定的极小常数。

L_{edge}能够引导模型关注边界变化显著的区域，从而提升模型对软边界和复杂轮廓的感知能力。对预测边界概率图$\boldsymbol{E}$进行插值，得到与语义特征$\boldsymbol{F}_{\mathrm{s}}$相同的空间分辨率的边界概率图$\hat{\boldsymbol{E}}$。随后，将$\hat{\boldsymbol{E}}$与$\boldsymbol{F}_{\mathrm{s}}$沿通道维度进行拼接，通过$1\times1$卷积进行特征压缩与映射，得到边界增强特征：

$$\boldsymbol{F}_b=Conv_{1\times1}\left(Concat\left(\boldsymbol{F}_{\mathrm{s}},\hat{\boldsymbol{E}}\right)\right) \tag{7}$$

式中：$Concat()$为通道维度拼接操作。

通过上述融合过程，边界结构线索被显式注入语义特征空间，使模型能够更好地刻画可通行区域的边缘形状与过渡区域，从而提升模型在非结构化道路场景下的分割稳定性。

2.4 几何-深度先验注意力模块

为进一步提升模型在非结构化道路场景中的空间结构建模能力与预测稳定性，本文设计几何-深度先验注意力模块，以边界特征融合后的增强特征作为输入，通过引入几何连续性先验、垂直方向伪深度先验以及空间-通道联合注意力机制，对语义特征进行结构化重加权，从而增强模型在复杂地形、弱纹理区域以及软边界条件下的可通行区域识别能力。

需要说明的是，本文所述深度先验并不依赖真实深度图、LiDAR点云或额外深度估计网络，而是根据道路图像中近处区域通常位于图像下方、远处区域通常位于图像上方的成像规律，构造一种沿垂直方向变化的伪深度权重分布。因此，该模块仍然保持RGB-only输入形式。

假设GDPA模块的输入为BFF模块输出的边界增强特征$\boldsymbol{F}\in\mathbf{R}^{B\times C\times H'\times W'}$，其中，$B$为批量大小，$C$为通道数，$H'$和$W'$分别为特征图的高度和宽度。同时，将预测边界概率图通过插值方式对齐至与F相同的空间尺度，即$\hat{\boldsymbol{E}}\in\mathbf{R}^{B\times1\times H'\times W'}$。

首先，为增强模型对道路边界连续结构的感知能力，基于边界概率图构建几何连续性先验。采用水平方向和垂直方向的非对称卷积核，分别提取边界区域在两个方向上的长程结构响应，并通过Sigmoid函数归一化，得到几何先验掩码：

$$\boldsymbol{M}_{\mathrm{g}}=\sigma\left(\boldsymbol{W}_{\mathrm{h}}*\hat{\boldsymbol{E}}+\boldsymbol{W}_{\mathrm{v}}*\hat{\boldsymbol{E}}\right) \tag{8}$$

式中：$\sigma()$为Sigmoid函数，$\boldsymbol{W}_{\mathrm{h}}$和$\boldsymbol{W}_{\mathrm{v}}$分别为捕捉不同方向边界连续性的卷积核。

为突出近处可通行区域在导航决策中的重要性，构造垂直方向伪深度先验。定义归一化坐标为：

$$y_i=\frac{i}{H'-1},\quad i=0,1,\cdots,H'-1 \tag{9}$$

式中：i为特征图的行索引。

$y_i\in[0,1]$在图像中该值从上至下逐渐增大，基于该归一化坐标，构造伪深度先验描述：

$$M_{\mathrm{d}}(i)=\beta+(1-\beta)y_i^{\alpha},\quad \alpha>0,\beta\in(0,1) \tag{10}$$

式(10)中，α用于控制权重增长曲线的形状，β用于控制最小权重。通过该设计，图像下方区域能够获得更大的先验权重，使模型更加关注近处可通行区域及其边界变化。在实际计算中，M_{d}通过广播机制扩展至与输入特征相同的空间维度。

进一步提升模型对关键区域和有效通道的自适应选择能力，GDPA模块引入空间-通道联合注意力机制。使用1×1卷积生成空间注意力权重$\boldsymbol{A}_{\mathrm{s}}$，利用压缩与激励(Squeeze-and-Excitation，SE)模块生成通道注意力权重$\boldsymbol{A}_{\mathrm{c}}$，通过逐元素乘法得到联合注意力权重$\boldsymbol{A}$：

$$\begin{cases}\boldsymbol{A}_{\mathrm{s}}=\sigma(Conv_{1\times1}(\boldsymbol{F})), & \boldsymbol{A}_{\mathrm{s}}\in\mathbf{R}^{B\times1\times H'\times W'}\\ \boldsymbol{A}_{\mathrm{c}}=f_{\mathrm{se}}(\boldsymbol{F}), & \boldsymbol{A}_{\mathrm{c}}\in\mathbf{R}^{B\times C\times1\times1}\\ \boldsymbol{A}=\boldsymbol{A}_{\mathrm{s}}\odot\boldsymbol{A}_{\mathrm{c}}\end{cases} \tag{11}$$

式中：$\odot$为逐元素乘法，通过广播机制完成维度对齐。

随后，将几何连续性先验、伪深度先验与空间-通道联合注意力进行融合，通过广播方式扩展至与$\boldsymbol{A}$相同的维度得到混合增强掩码：

$$\boldsymbol{M}_{\mathrm{h}} = \boldsymbol{A} \odot \boldsymbol{M}_{\mathrm{g}} \odot \boldsymbol{M}_{\mathrm{d}} \tag{12}$$

最后，采用可学习系数的残差增强方式输出GDPA模块的增强特征：

$$\boldsymbol{F}_{\mathrm{f}} = \boldsymbol{F} + \gamma(\boldsymbol{F} \odot \boldsymbol{M}_{\mathrm{h}}) \tag{13}$$

式中：γ为可学习系数，用于控制先验增强分支对原始特征的影响强度。

通过残差连接，模型能够在保留原始语义特征的基础上，自适应融合边界结构、空间位置和通道响应信息，避免先验约束对特征学习产生过强干扰。GDPA模块将边界几何连续性、垂直方向伪深度先验和数据驱动注意力机制进行统一建模。几何连续性先验用于增强软边界区域的结构表达，伪深度先验用于提高模型对近处可通行区域的关注程度，空间通道联合注意力则进一步提升特征选择能力。通过上述设计，模型能够在复杂地形、弱纹理和不规则边界条件下获得更加稳定的分割结果。

3 试验设置

本文在ORFD数据集[18]上完成模型训练、验证与测试，并进一步在RELLIS-3D非结构化道路多模态数据集[19]上开展跨数据集泛化试验，以验证所提方法在不同非结构化道路场景下的鲁棒性与泛化能力。其中，ORFD数据集标注类别包括可通行区域、不可通行区域和天空3类，适用于可通行区域检测任务。选取RELLIS-3D数据集[19]中不同地形与环境条件下的983张图像作为泛化测试集，用于跨场景性能评估。

在训练设置方面，本文采用带解耦权重衰减的自适应矩估计（Adaptive Moment Estimation with Decoupled Weight Decay，AdamW）优化器，初始学习率设为4×10^{-5}，批大小设置为12，最大训练迭代次数为160 000次。学习率调整采用poly多项式衰减策略，衰减指数设为0.9。

为提升模型的鲁棒性，训练过程中采用随机裁剪、随机水平翻转、亮度扰动和随机旋转等数据增强操作。其中，随机裁剪尺寸设置为1 024×512。模型训练阶段仍采用式(3)定义的联合损失函数，多组试验对比后边界损失权重参数设置为$\lambda = 1$，GDPA模块中的伪深度先验参数设置为$\alpha = 2$、$\beta = 0.2$。

模型性能评估采用准确率（Accuracy）A、精确率（Precision）P、召回率（Recall）R、F1分数（F1-score）、交并比（Intersection over Union，IoU）和帧速率等指标。其中，准确率衡量像素分类的整体正确程度，精确率和召回率分别反映可通行区域预测的准确性与完整性，F1分数综合评价精确率与召回率，交并比衡量预测可通行区域与真实可通行区域之间的重叠程度，帧速率衡量模型的推理速度。

以可通行区域为正类，将不可通行区域和天空区域视为负类，各项分割指标的计算公式为：

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

$$P = \frac{TP}{TP + FP} \tag{15}$$

$$R = \frac{TP}{TP + FN} \tag{16}$$

$$A_{\mathrm{F1-score}} = \frac{2TP}{2TP + FP + FN} \tag{17}$$

$$I_{\mathrm{IoU}} = \frac{TP}{TP + FP + FN} \tag{18}$$

式中：TP为被正确预测为可通行区域的像素数量，TN为被正确预测为负类的像素数量，FP为被错误预测为可通行区域的负类像素数量，FN为被错误预测为负类的可通行区域像素数量。

4 试验结果与分析

4.1 ORFD数据集上的对比试验

使用ORFD数据集验证本文方法在非结构化道路可通行区域检测任务的有效性，并与多种主流方法进行对比试验。基于单目RGB输入的对比方法包括U-Net、采用ResNet-101骨干网络的DeepLabv3+语义分割模型DLV3+-R101、融合RGB与LiDAR信息的SNE-RoadSeg、OFF-Net、M2F2-Net和RoadFormer，涵盖了不同输入模态和网络结构，能够较全面地反映模型在分割精度、推理效率和工程部署方面的综合性能，结果如表1所示。

由表1可知，本文方法在单目RGB方法中取得了最优的综合表现。相较于U-Net模型，本文方法的IoU从41.1%提升至90.8%，Precision从63.7%提升至95.6%，Recall从53.7%提升至95.2%，F1-score从58.3%提升至95.1%，在可通行区域识别的准确性与完整性方面均有明显提升。

相较于同为RGB输入的DLV3+-R101方法，本文方法的IoU、Precision和F1-score分别提高了20.8百分点、17.5百分点和12.7百分点，表明边界增强与几何-深度先验注意力机制能够有效弥补单目模型在复杂边界和弱纹理区域中的识别不足。本文方法的准确率略低于DLV3+-R101，主要源于模型准确率易受到背景像素占比影响，难以充分反映可通行区域的分割质量，因此，IoU、Precision、Recall和F1-score更能体现模型在目标区域上的综合性能。

表1　ORFD数据集上的对比试验结果

方法	输入模态	*A*/%	*P*/%	*R*/%	F1-score/%	IoU/%	帧速率/帧·s^{-1}
U-Net[9]	RGB	95.9	63.7	53.7	58.3	41.1	4.44
DLV3+-R101[10]	RGB	98.0	78.1	87.1	82.4	70.0	38.73
SNE-RoadSeg[12]	RGB+LiDAR	93.8	86.7	92.7	89.6	81.2	1.05
OFF-Net[18]	RGB+LiDAR	94.5	86.6	94.3	93.0	82.3	1.58
M2F2-Net[14]	RGB+LiDAR	98.1	97.3	95.5	96.4	93.1	1.73
RoadFormer[15]	RGB+Normal	97.9	95.1	97.2	96.1	92.5	
本文	RGB	96.4	95.6	95.2	95.1	90.8	14.4

与RGB+LiDAR多模态方法相比，本文方法的IoU达到90.8%，高于SNE-RoadSeg的81.2%和OFF-Net的82.3%，说明在不依赖LiDAR点云信息的情况下，该方法仍可获得较强的可通行区域检测能力。

相比M2F2-Net和RoadFormer，本文的IoU分别降低2.3百分点和1.7百分点，F1-score分别降低1.3百分点和1.0百分点。虽然差距相对较小，但本文方法仅使用RGB图像作为输入，避免了点云投影、跨模态对齐和多传感器同步等额外处理流程，在系统成本和部署复杂度方面更具优势。

在推理效率方面，本文方法达到14.4帧/s，明显优于SNE-RoadSeg、OFF-Net和M2F2-Net等融合方法。虽然DLV3+-R101的推理速度更快，但其IoU和F1-score明显低于本文方法，难以在复杂非结构化场景中取得同等分割效果。

因此，本文方法在保持RGB-only输入形式的前提下，实现了接近多模态先进方法的分割精度，兼具较高的推理效率，在检测精度、计算效率和工程部署成本间取得了较好的平衡。

将ORFD数据集上的对比试验结果进行可视化分析，如图2所示。从可视化结果可知，OFF-Net在部分复杂纹理区域容易出现误分割现象，RoadFormer虽然整体检测效果较好，但在局部边界细节仍存在边缘不够平滑、区域过渡不够自然的问题。本文方法通过引入显式边界监督与边界特征融合机制，使模型对软边界区域具有更强的感知能力，能够较好地保持可通行区域的完整性，在道路边界处的预测结果更加清晰、连续，所以在复杂背景干扰和边缘过渡区域中表现更加稳定。可见，本文方法在单目RGB输入条件下可有效提升非结构化道路可通行区域检测的精度与边界质量。

4.2　RELLIS-3D数据集上的泛化试验

考虑到实际应用中非结构化道路场景具有较强的

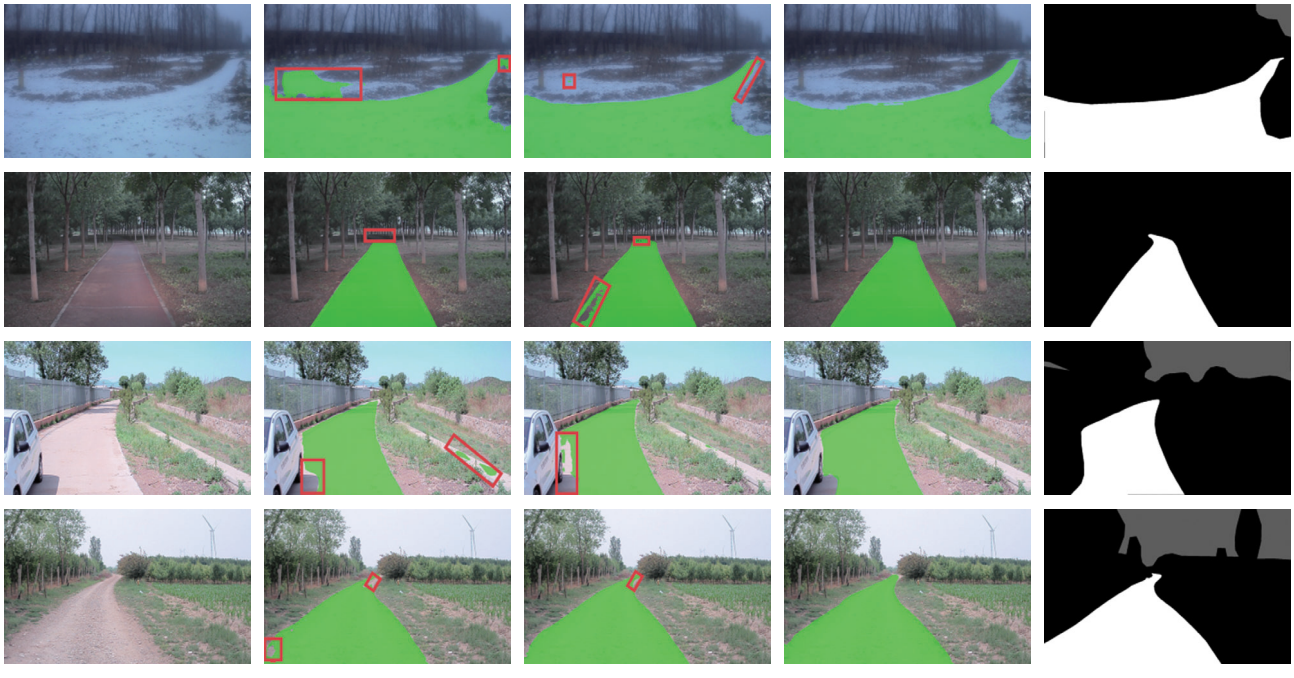


图2　ORFD数据集上不同方法的可视化结果

不确定性，使用同一数据集测试难以充分反映模型在陌生环境中的适应能力，本文使用RELLIS-3D数据集进行泛化试验。由于ORFD与RELLIS-3D在地形起伏、地表材质、道路边界形态、障碍物分布以及光照条件等方面存在明显差异，该试验能够更真实地检验模型在不同非结构化道路场景下的鲁棒性。

模型仅在ORFD数据集上完成训练，在不进行任何微调的情况下，迁移至RELLIS-3D验证集进行测试。选取RELLIS-3D验证集的983张图像作为测试样本，评价指标与ORFD数据集上的设置保持一致。不同数据集中模型性能对比结果如表2所示。

表2　本文方法在不同数据集上的泛化试验结果

数据集	*A*/%	*P*/%	*R*/%	F1-score/%	IoU/%	帧速率/帧·s^{-1}
ORFD	96.4	95.6	95.2	95.1	90.8	14.4
RELLIS-3D	91.2	90.1	89.3	89.7	84.7	14.4
Δ(Drop)	5.2	5.5	5.9	5.4	6.1	0

由表2可知，当模型由ORFD迁移至RELLIS-3D后，各项评价指标均出现一定程度下降，表中Δ(Drop)表示模型跨数据集测试后各项指标的下降值，其计算方式为ORFD数据集的指标值减去RELLIS-3D数据集上的指标值。准确率、精确率、召回率、F1分数和交并比分别下降5.2百分点、5.5百分点、5.9百分点、5.4百分点和6.1百分点，而帧速率保持不变。各项分割指标的下降幅度总体保持在5~6百分点，说明不同数据集间的场景分布差异会对模型性能产生影响。尤其在RELLIS-3D中，道路区域形态更加不规则，地表纹理和背景干扰更加复杂，软边界与障碍物邻接区域更容易造成局部误分割，从而使模型的区域判别难度增加。尽管如此，本文方法在RELLIS-3D数据集上的IoU和F1-score分别达到84.7%、89.7%，说明模型在未经过目标数据集微调的情况下，依然能够保持较好的区域分割能力。模型在2个数据集上的帧速率均为14.4帧/s，表明跨数据集测试并未增加模型推理开销，所提方法在保持较好泛化表现的同时，仍具有稳定的实时推理能力。

为直观比较所提方法的跨数据集泛化效果，选取未引入边界增强模块、边界特征融合模块和几何-深度先验注意力模块的SegFormer(MiT-B4+UPerHead)作为基线模型。对泛化试验结果进行可视化分析，如图3所示。RELLIS-3D数据集中存在较多地表纹理复杂、道路边界不清晰以及可通行区域形态变化较大的场景。本文方法在可通行区域轮廓恢复方面更加连续，能够较好地保持道路主体区域的完整性，在道路边缘、远近区域过渡以及障碍物邻接区域表现出更稳定的分割效果。尤其在弱边界和复杂背景干扰条件下，本文方法能够减少部分区域误分割现象，使预测结果与GT标注在整体形态上更加接近。

因此，泛化试验结果表明：本文方法虽然在跨数据集测试时存在一定性能下降，但下降幅度相对可控，且在RELLIS-3D数据集上仍保持了较好的分割精度与实时性，说明边界辅助监督与边界特征融合机制能够增强模型对软边界区域的刻画能力，几何-深度先验注意力模块则有助于提升模型对复杂地形结构和空间层次信息的利用能力，从而提高模型在陌生非结构化道路场景

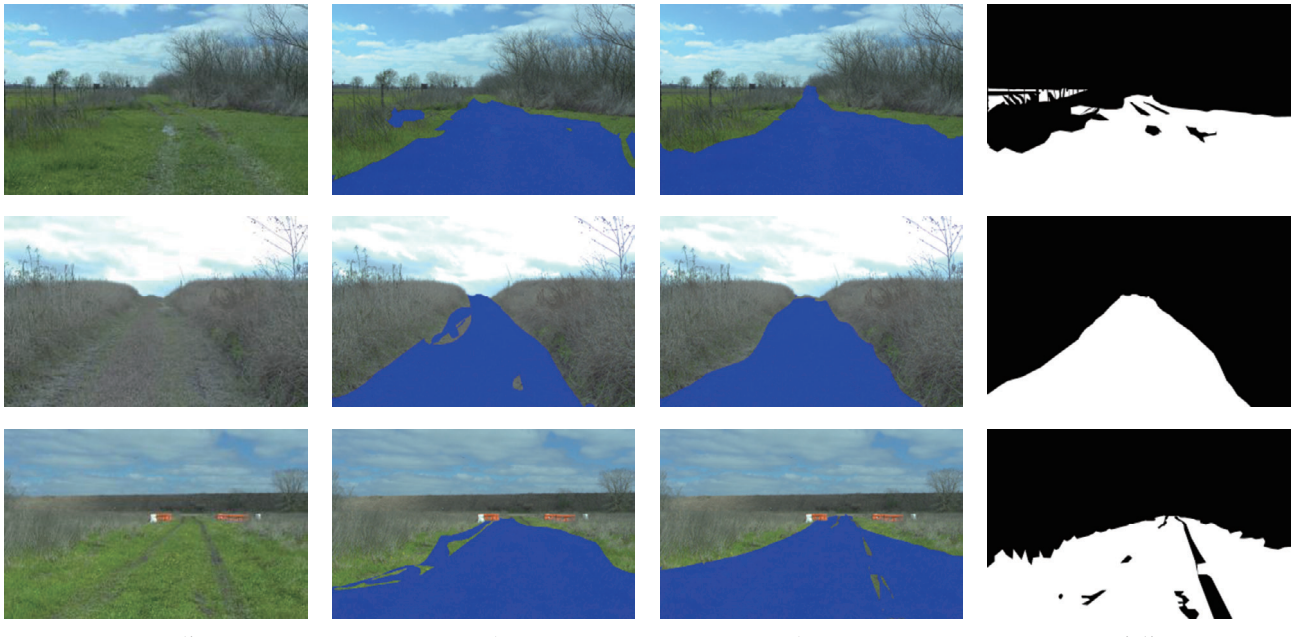


图3　RELLIS-3D数据集上的跨场景泛化可视化结果

下的泛化表现。

4.3 模块消融试验

在ORFD数据集上开展消融试验，分析各模块对可通行区域检测性能的影响。试验以SegFormer（MiT-B4+UPerHead）作为基线模型，逐步引入边界增强模块（Boundary）、边界特征融合模块（BFF）以及几何－深度先验注意力模块（GDPA），构建不同模块组合进行对比分析，结果如表3所示。

表3 ORFD数据集上本文模型消融试验结果

Boundary	BFF	GDPA	*A*/%	*P*/%	*R*/%	F1-score/%	IoU/%	帧速率/帧·s^{-1}
			92.1	92.1	91.7	91.4	84.2	18.7
√			93.7	93.7	93.0	92.9	86.7	17.2
	√		93.4	92.8	92.5	92.6	86.2	17.5
		√	93.0	92.4	92.2	92.3	85.7	17.0
√	√		95.2	95.2	94.7	94.3	89.3	15.1
√		√	94.6	94.1	93.7	93.8	88.3	15.6
	√	√	94.2	93.8	93.4	94.0	88.7	15.8
√	√	√	96.4	95.6	95.2	95.1	90.8	14.4

注：√表示使用该模块。

由表3可知，基线模型的IoU和F1-score分别为84.2%和91.4%。在单模块设置下，引入Boundary模块后，模型IoU提升至86.7%，F1-score提升至92.9%，分别较基线模型提高了2.5百分点和1.5百分点，说明显式边界监督能够增强模型对非结构化道路软边界区域的感知能力，缓解边界模糊、轮廓断裂以及局部误分割等问题。

单独引入BFF或GDPA时，IoU分别提升至86.2%、85.7%，虽然均优于基线模型，但增幅相对较小。这表明仅依靠特征融合或注意力重加权难以充分解决非结构化道路中的边界不确定性问题，显式边界约束仍是提升分割性能的重要因素。在双模块组合情况下，各模块间的互补作用进一步体现。当Boundary与BFF共同引入时，模型IoU提升至89.3%，F1-score提升至94.3%，性能提升较为明显。这说明边界监督获得的结构信息能够通过BFF进一步融入语义特征空间，从而增强模型对道路轮廓、细长区域以及过渡边界的表达能力。Boundary与GDPA组合、BFF与GDPA组合时，IoU分别达到88.3%、88.7%，但整体略低于Boundary与BFF组合，说明GDPA主要发挥空间结构增强与特征重加权作用，其效果在与清晰边界提示和语义融合共同作用时更易充分体现。

当Boundary、BFF和GDPA模块共同作用时，模型取得最优性能，IoU达到90.8%，F1-score达到95.1%，相比基线模型分别提高了6.6百分点和3.7百分点。结果表明：边界增强、边界特征融合以及几何-深度先验注意力之间能够形成有效协同；Boundary模块提供显式边界约束，BFF模块将边界信息注入语义特征，GDPA模块进一步利用空间结构先验对融合特征进行重加权，使模型在复杂纹理、弱边界和不规则地形条件下能够同时保持较好的区域完整性与边界连续性。

从推理效率来看，随着模型模块逐步增加，帧速率由基线模型的18.7帧/s下降至完整模型的14.4帧/s，推理速度存在一定下降。主要因为边界分支、特征融合和注意力计算引入了额外计算量。但总体来看，完整模型仍保持了较为稳定的推理效率，在分割精度明显提升的同时，计算开销增加相对可控。因此，本文方法能够在精度提升与推理效率间取得平衡。

5 结束语

针对非结构化道路场景中可通行区域边界模糊、地表纹理复杂以及局部误分割等问题，本文提出一种融合边界增强与几何-深度先验注意力的RGB-only可通行区域检测方法。通过试验分析，可以得到以下结论：

a. 本文构建的显式边界增强机制能够有效提升模型对软边界与复杂轮廓区域的表达能力。通过边界辅助监督和边界特征融合，模型能够在学习区域语义信息的同时进一步关注道路边缘结构，从而在一定程度上缓解非结构化道路场景中的边界模糊、轮廓断裂和误分割问题。

b. 本文设计的几何-深度先验注意力模块能够增强模型对复杂地形结构的感知能力。该模块将几何连续性先验与垂直方向伪深度先验引入特征重加权过程，并结合空间-通道注意力机制，使模型在弱纹理、地形变化明显以及远近区域过渡复杂的场景下具有更好的鲁棒性和结构一致性。

c. 本文方法在ORFD数据集上取得了较好的分割性能，并通过RELLIS-3D数据集上的跨数据集泛化试验进一步验证了模型的适应能力。

本文方法在分割精度、边界质量和推理效率间取得了较好的平衡，为非结构化道路可通行区域检测提供了一种可行的RGB-only解决方案。

参 考 文 献

[1] YURTSEVER E, LAMBERT J, CARBALLO A, et al. A Survey of Autonomous Driving: Common Practices and Emerging Technologies[J]. IEEE Access, 2020, 8: 58443-58469.

[2] AWAN M, WHANGBO T K, SHIN J. Deep Learning Methods for Autonomous Driving Scene Understanding Tasks: A Review[J]. Expert Systems with Applications, 2025, 287.

[3] FENG D, HAASE-SCHÜTZ C, ROSENBAUM L, et al. Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(3): 1341-1360.

[4] 宋亮, 谷玉海, 石文天. 基于改进BiSeNet的非结构化道路分割算法研究[J]. 应用光学, 2023, 44(3): 556-564.
SONG L, GU Y H, SHI T W. Unstructured Road Segmentation Algorithm Based on Improved BiSeNet[J]. Journal of Applied Optics, 2023, 44(3): 556-564.

[5] 周越洋, 徐江, 钟珊, 等. 点云空间与反射强度融合的非结构化道路可行驶区域检测[J]. 控制理论与应用, 2024, 41(5): 847-854.
ZHOU Y Y, XU J, ZHONG S, et al. Detection of Drivable Areas on Unstructured Roads Fused with Point Cloud Space and Reflection Intensity[J]. Control Theory & Applications, 2024, 41(5): 847-854.

[6] 时利, 赵盼, 林玲龙, 等. 基于道路拓扑结构约束的未知非结构化场景可通行区域检测方法[J]. 电子测量技术, 2024, 47(16): 147-154.
SHI L, ZHAO P, LIN L L, et al. Robust Drivable Area Detection Method for Unknown Off-Road Scenes under Road Topological Constraints[J]. Electronic Measurement Technology, 2024, 47(16): 147-154.

[7] CHEN X, DONG Y J, LI X Y, et al. Drivable Area Recognition on Unstructured Roads for Autonomous Vehicles Using an Optimized Bilateral Neural Network[J]. Scientific Reports, 2025, 15(1).

[8] SHELHAMER E, LONG J, DARRELL T. Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(4): 640-651.

[9] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer, 2015: 234-241.

[10] CHEN L, ZHU Y K, PAPANDREOU G, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation[C]// Proceedings of the European Conference on Computer Cision (ECCV). Munich, Germany: Springer, 2018: 801-818.

[11] ZHU X G, ZHOU H, WANG T, et al. Cylindrical and Asymmetrical 3D Convolution Networks for LiDAR Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(10): 6807-6822.

[12] FAN R, WANG H L, CAI P D, et al. SNE-RoadSeg: Incorporating Surface Normal Information into Semantic Segmentation for Accurate Freespace Detection[C]// European Conference on Computer Vision. Glasgow, United Kingdom: Springer, 2020: 340-356.

[13] LIU Y F, YAN J J, JIA F, et al. PETRv2: A Unified Framework for 3D Perception from Multi-Camera Images[C]// 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE, 2023: 3262-3272.

[14] YE H L, MEI J L, HU Y. M2F2-Net: Multi-Modal Feature Fusion for Unstructured Off-Road Freespace Detection[C]// 2023 IEEE Intelligent Vehicles Symposium (IV). Anchorage, AK, USA: IEEE, 2023: 1-7.

[15] LI J H, ZHANG Y K, YUN P, et al. RoadFormer: Duplex Transformer for RGB-Normal Semantic Road Scene Parsing [J]. IEEE Transactions on Intelligent Vehicles, 2024, 9(7): 5163-5172.

[16] SUN T, YE H L, MEI J L, et al. ROD: RGB-only Fast and Efficient Off-Road Freespace Detection[C]// 2025 IEEE International Conference on Robotics and Automation (ICRA). Atlanta, GA, USA: IEEE, 2025: 9787-9793.

[17] XIE E Z, WANG W H, YU Z D, et al. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers[C/OL]// Advances in Neural Information Processing Systems. NeurIPS, 2021, 34: 12077-12090.

[18] MIN C, JIANG W Z, ZHAO D W, et al. ORFD: A Dataset and Benchmark for Off-Road Freespace Detection[C]// 2022 International Conference on Robotics and Automation (ICRA). Philadelphia (PA), USA: IEEE, 2022: 2532-2538.

[19] JIANG P, OSTEEN P, WIGNESS M, et al. RELLIS-3D Dataset: Data, Benchmarks and Analysis[C]// 2021 IEEE International Conference on Robotics and Automation (ICRA). Xi'an, China: IEEE, 2021: 1110-1116.

（责任编辑　瑞　秋）

修改稿收到日期为2026年5月19日。