

基于多传感器BEV特征融合的 轨道分割算法

石海丰, 苗壮壮, 闫正祥, 韩改堂, 杨连报, 汪小亮

(北京和利时系统工程有限公司, 北京 100176)

摘要: 轨道分割技术是列车主动检测障碍物的核心支撑。目前, 主流轨道分割法通常依赖于相机传感器, 无法捕获轨道位置深度信息, 影响分割精度。为此提出1种基于多传感器BEV特征融合的轨道分割算法, 通过结合摄像头和激光雷达的输入, 增强对轨道数据的感知能力。首先, 从摄像头捕获的图像数据和雷达生成的点云数据中提取特征; 然后, 设计1个基于注意力机制的融合模块整合这2种数据特征; 最后, 通过线性层对融合后的特征进行处理, 实现轨道精确分割。为验证算法有效性, 在OSDaR23数据集上进行训练与测试。结果表明, 该方法具备出色的分割性能。

关键词: 轨道分割; 雷达; 摄像头; 鸟瞰视图; 多传感器融合; 注意力机制

中图分类号: U284

文献标识码: A

文章编号: 1001-683X (2025) 04-0087-08

DOI: 10.19549/j.issn.1001-683x.2024.07.09.003

0 引言

轨道交通网络快速扩张带来巨大运输需求, 对列车运行安全提出更高要求。在轨道交通系统中, 列车行驶区域存在障碍物可能会对列车正常运行造成严重威胁, 甚至导致重大安全事故。传统轨道交通列车运行控制系统主要依靠地面信号系统及轨旁设备获取列

车运行前方障碍物和占用情况等, 属于被动检测^[1]。近年来, 通过在列车前方安装雷达和摄像头等传感设备, 开发和部署有效的主动障碍物检测系统, 可预防和减少风险, 实现列车自主感知运行前方轨行区域线路状况^[2-3]。轨道检测和分割技术是主动障碍物检测系统的基础, 能够识别列车行驶的轨道区域, 明确轨道所处的位置, 为障碍物检测提供空间信息^[4]。

已有研究通过提取图像传统特征, 如边缘特征、颜色直方图、频域特征等, 有效获得图像中的局部信息和纹理信息进而完成分割^[5]。然而, 传统特征提取方法无法提取图像的深层特征, 在鲁棒性、可扩展性、

基金项目: 中国国家铁路集团有限公司科技研究开发计划项目 (L2022X003, N2022G010, N2023G058)

第一作者: 石海丰 (1978—), 男, 高级工程师。

E-mail: shf13231467574@126.com

数据依赖性等方面存在劣势。随着深度学习技术的发展,语义分割领域取得显著进展^[6-7],主要使用基于区域的卷积神经网络(CNN)对图像进行分割^[8]。其中,Shelhamer等^[9]提出全卷积网络,能够直接生成像素级别的语义分割结果。为提高准确性,Chen等^[10]引入空洞卷积,扩大感受野,并提高语义分割的性能。He等^[11]提出Mask R-CNN方法,将目标检测和语义分割结合起来,实现像素级别的目标分割。然而,基于单摄像头传感器的分割模型仅能提供二维视觉信息,缺少三维空间的位置深度信息,限制了模型的分割性能,且单图像模型无法直接提供立体信息,在处理复杂场景时,难以区分前后关系和空间位置。

针对单传感器信息局限问题,提出1种基于多传感器鸟瞰视图(BEV)特征融合的轨道分割算法(简称融合算法)。利用相机和雷达2种传感器实现轨道分割。整个模型分为相机通道、雷达通道和多传感器融合3部

分。相机通道采用基于Transformer的特征提取网络,对图像进行初步的特征提取,然后设计多尺度特征提取模块,融合多尺度图像信息,最后利用BEV特征提取模块,将图像特征投影到BEV空间中。雷达通道采用主流点云特征提取网络,将点云也投影到BEV空间中。多传感器融合部分用于融合多传感器深层特征,实现轨道分割。

1 融合算法设计

融合算法架构(见图1)分为3部分:相机通道、雷达通道和多传感器融合。其中,相机通道包含Swin Transformer特征提取主干网络、多尺度特征融合模块、BEV特征提取模块,输出图像BEV特征。雷达通道包含Point Pillars点云特征提取主干网络输出点云BEV特征。多传感器融合包含多模态表示融合模块以及分割回归头,回归轨道的分割结果。

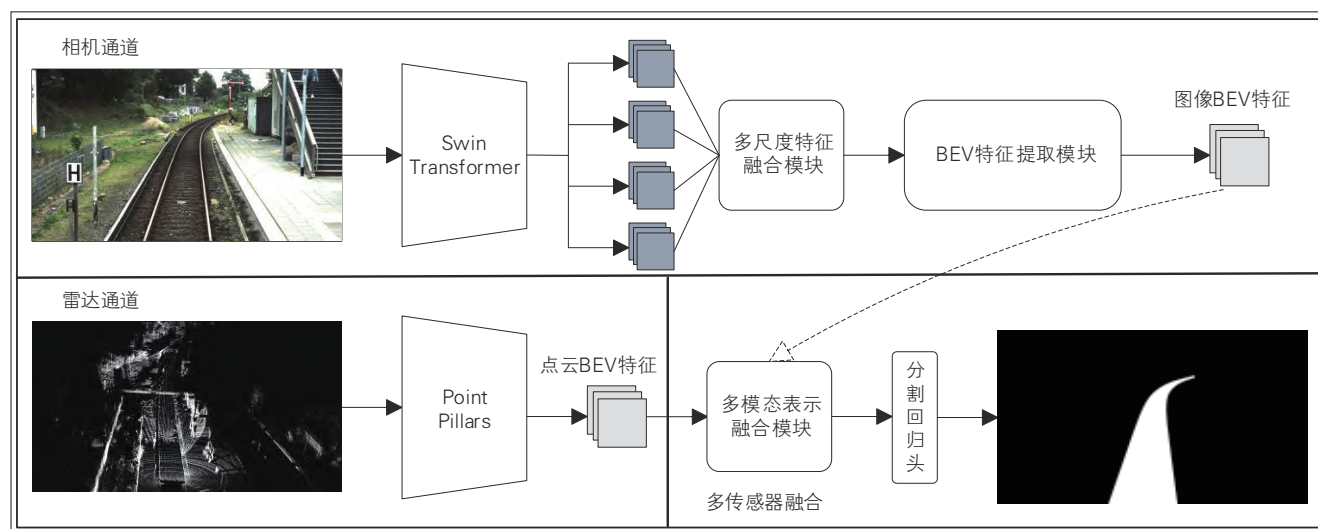


图1 算法架构

1.1 相机通道

1.1.1 特征提取网络

Swin Transformer^[12]作为图像特征提取的主干网络,是1种具有和CNN类似的金字塔结构Transformer主干网络,其采用Transformer解码器代替常规CNN进行图像特征提取。图像是二维结构数据,而Transformer网络面向一维序列数据,因此Swin Transformer将二维图像转换为符合Transformer编码器输入的一维序列数据。Swin Transformer由4个阶段Transformer编码层组

成,各阶段步长分别为4、8、16和32,以此计算不同尺度的局部注意力。不同步长代表不同大小的局部特征聚合,同时也能实现不同尺度的特征提取。Swin Transformer最终输出4种不同尺度特征图,分别为 $[I_1 \in \mathbf{R}^{(H/4) \times (W/4) \times C_1}, I_2 \in \mathbf{R}^{(H/8) \times (W/8) \times C_2}, I_3 \in \mathbf{R}^{(H/16) \times (W/16) \times C_3}, I_4 \in \mathbf{R}^{H/32 \times W/32 \times C_4}]$,其中 I_1 为输出第1尺度特征, I_2 为输出第2尺度特征, I_3 为输出第3尺度特征, I_4 为输出第4尺度特征, C_1 、 C_2 、 C_3 、 C_4 表示各个尺度特征的通道数, H 和 W 分别为图像的长和宽。

1.1.2 多尺度特征融合

在图像和视频分析中,轨道轨面大小会因拍摄角度和距离不同而变化。为提高分割效果,需要同时使用不同层

次特征,高层特征有助于准确定位目标,而低层特征有助于识别细节。通过融合不同层次特征,提高对不同尺度轨道的分割能力,设计1种多尺度特征融合模块(见图2)。

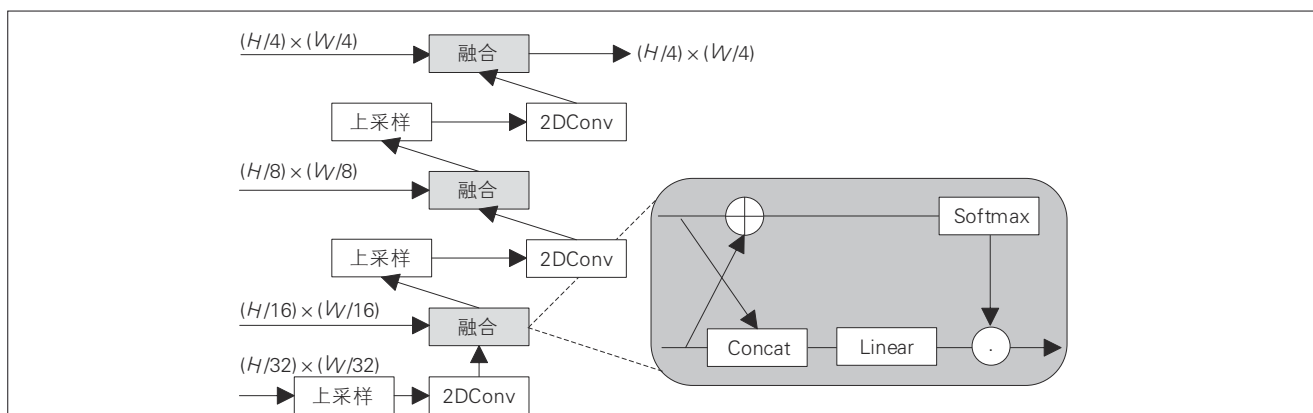


图2 多尺度特征融合模块

整体采用自上而下的金字塔结构,对高层特征图进行上采样与低层特征图逐步融合,最终输出包含多尺度信息的特征表示。以融合 I_4 和 I_3 为例,首先对 I_4 进行双线性插值上采样,将特征图分辨率上采样至与 I_3 的特征图分辨率相同,并利用2D卷积神经网络(2DConv)消除上采样带来的棋盘效应。为进一步融合2种不同层级特征图,设计1种基于通道注意力的融合单元,利用2种特征融合方式的优势和注意力机制选择性关注特征中重要部分。对2种特征进行通道特征级的拼接操作,得到更全面的特征表示,利用线性层自适应融合2种特征,融合过程见式(1)。

$$IC_{34} = \text{Linear}(\text{Concat}(\text{2Dconv}(\text{Up}(I_4, I_3)))) \quad (1)$$

式中: IC_{34} 为 I_4 和 I_3 融合后的特征表示;Up为双线性插值上采样操作;Linear为全连接层;Concat为特征拼接操作。

针对图像特征通道,拼接后通道数为 $2C_3$ 的特征,通过线性变换为通道数为 C_3 的特征。直接拼接后的特征虽然具有更全面的特征表示,但是会导致特征冗余。为提取其中有效的特征表示,进一步计算特征通道注意力权重,对2种层级特征进行特征级相加操作,在每个通道上得到更充分的特征表示 $IA_{34} \in \mathbf{R}^{(H/16) \times (W/16) \times C_3}$ 。对特征通道进行Softmax运算,求得每个通道特征的注意力权重^[13]见式(2)。

$$\text{Att}_{(i,j,k)} = \frac{e^{IA_{34}(i,j,k)}}{\sum_{c=1}^{C_3} e^{IA_{34}(i,j,c)}} \quad (2)$$

式中: $\text{Att}_{(i,j,k)}$ 为特征图中 (i,j) 位置第 k 通道的特征注意力权重; $IA_{34}(i,j,k)$ 为特征图中 (i,j,k) 位置的特征变量。

将注意力权重与拼接融合后的特征进行加权点乘输出最终的多尺度特征表示。模块的整体流程公式见式(3)一式(5)。

$$F_1 = \text{Softmax}(\text{2DConv}(\text{Up}(I_4)) + I_3) \times \text{Linear}(\text{Concat}(\text{2DConv}(\text{Up}(I_4)), I_3)) \quad (3)$$

$$F_2 = \text{Softmax}(\text{2DConv}(\text{Up}(F_1)) + I_2) \times \text{Linear}(\text{Concat}(\text{2DConv}(\text{Up}(F_1)), I_2)) \quad (4)$$

$$F_3 = \text{Softmax}(\text{2DConv}(\text{Up}(F_2)) + I_1) \times \text{Linear}(\text{Concat}(\text{2DConv}(\text{Up}(F_2)), I_1)) \quad (5)$$

式中: $F_1 \in \mathbf{R}^{(H/16) \times (W/16) \times C_3}$ 为第3尺度和第4尺度特征融合后的特征; $F_2 \in \mathbf{R}^{(H/8) \times (W/8) \times C_2}$ 为 F_1 和第2尺度融合后输出的特征; $F_3 \in \mathbf{R}^{(H/4) \times (W/4) \times C_1}$ 为模块最终输出,包含4种尺度特征表示。

1.1.3 图像BEV特征提取

图像到BEV的特征投影,采用自下而上的方式构建BEV特征,具体步骤见图3。该模块主要步骤包括3项:一是生成视锥,根据相机内外参将视锥中的点投

影到 ego 坐标系；二是对环视图像完成特征提取，构建图像特征点云；三是运用体素池化（Voxel Pooling）构

建 ego 坐标系点与图像特征点云的 BEV 特征 $B' \in \mathbb{R}^{H \times W \times C_p}$ ，其中 C_p 表示点云 BEV 特征的通道数。

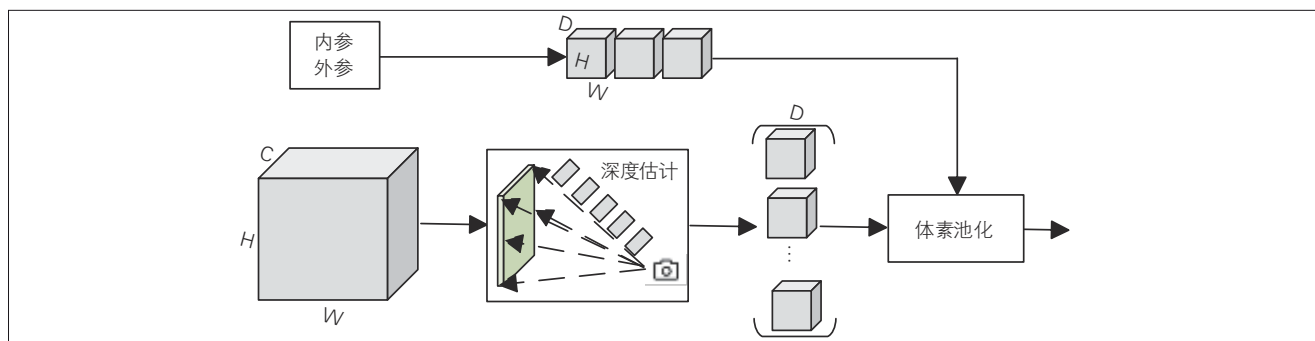


图3 图像BEV特征提取模块

1.2 雷达通道

为了从 BEV 空间解决目标分割问题，采用 Point Pillars^[14] 作为点云特征提取的主干网络。该网络直接将点云从俯视图视角划分为多个立方柱体（Pillars），从而构成伪图片数据，该数据也是 BEV 视角下的点云特征。首先将原始点云平均分配到大小相同的 $x-y$ 平面立方柱体中，每一个柱体被称为 Pillar。当点被放入柱体中，需要对柱体中每个点进行数据增强，计算每个点相对所有点平均值的偏移量、每个点到所对应 Pillar 几何中心点的偏移量。经过数据增强后每个点的特征包含 10 维。进一步堆叠 MLP+BN+Relu 神经网络层，对点云柱体特征进行特征提取，输入高维非空立方柱体特征。经过上述操作编码后的点，再对应 Pillar 放回到原来的位置上，最终生成伪图像数据 $B^p \in \mathbb{R}^{H \times W \times C_p}$ 。

1.3 多传感器 BEV 特征融合模块

相机通道和雷达通道分别输出图像 BEV 特征 B' 和点云 BEV 特征 B^p 。来自不同传感器的 BEV 特征表示各具优势，图像特征提供更丰富的颜色和纹理信息，点云则提供精确的距离信息，这对最终的分割性能具有积极作用，多传感器算法的关键是如何较优地融合来自不同传感器的特征。以往通常将这些特征直接进行相加或者拼接，这种处理方式非常粗糙，易造成特征融合冗余，直接影响算法性能。

面向不同传感器 BEV 特征，设计 1 种融合单元动态融合图像 BEV 特征和点云 BEV 特征的算法（见图 4），使得机器学习能够更有区别地表示。

融合模块从特征本身出发，通过神经网络自适应

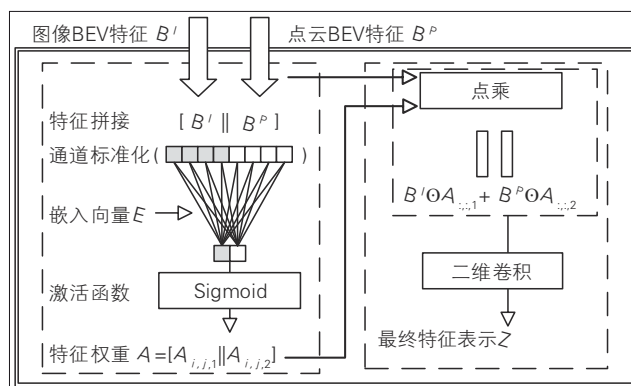


图4 多传感器BEV特征融合模块

地学习融合权重，将 2 种不同传感器的 BEV 特征置于同一空间进行融合。具体来说，为将 2 种不同传感器的 BEV 特征放置到一个空间中，首先通过特征拼接操作对 B' 和 B^p 进行通道特征拼接 $[B' || B^p] \in \mathbb{R}^{H \times W \times 2C_p}$ 。然后，基于拼接后的特征分别计算图像 BEV 特征和点云 BEV 特征的融合权重，不同子空间的 BEV 特征具有差异性，利用通道标准化消除特征之间尺度不一致问题。进一步引入 1 个嵌入向量 $E \in \mathbb{R}^{2C_p \times 2}$ 与拼接后的特征相乘，实现捕获不同特征之间的连接关系，并将维度映射为 2，方便后续计算不同传感器 BEV 特征的权重。最后利用 Sigmoid 激活函数和通道标准化对输入进行归一化，得到权重。计算过程见式（6）。

$$A = \text{Sigmoid}(\text{ChannelNorm}([B' || B^p]E)), \quad (6)$$

式中： $A \in \mathbb{R}^{H \times W \times 2}$ 为各项大于 0 的注意力系数矩阵；ChannelNorm 为通道特征标准化； E 为可学习的参数向量。

对于 B' 和 B^p 的自适应融合见式（7）。

$$Z_{ij} = 2D\text{Conv}(A_{i,j,1} \otimes B'_{i,j} + A_{i,j,2} \otimes B^p_{i,j}), \quad (7)$$

式中: $Z_{ij} \in \mathbf{R}^{1 \times 1 \times C_p}$ 为 (i, j) 位置融合后的通道特征; $Z \in \mathbf{R}^{H \times W \times C_p}$ 为融合后 BEV 空间的最终特征表示; $B'_{i,j}$ 和 $B^p_{i,j}$ 分别为 B' 和 B^p 中位置 (i, j) 对应的特征值; $A_{i,j,1}$ 和 $A_{i,j,2}$ 为 BEV 特征中 (i, j) 位置 B' 和 B^p 对应的融合权重值。

最后, 利用 2 个全连接层和 1 个 Sigmoid 激活函数获取最终的分割概率见式 (8)。

$$B = \text{Sigmoid}(\text{MLP}_2(\text{MLP}_1(Z))), \quad (8)$$

式中: $B \in \mathbf{R}^{H \times W \times 1}$ 为分割概率; MLP_1 为第 1 层全连接层, 特征通道维度变化为 $C_p \rightarrow C_p/2$; MLP_2 为第 2 层全连接层, 特征通道维度变化为 $C_p/2 \rightarrow 1$; Z 为融合后 BEV 空间的最终特征表示。

2 实验分析

在 OSDaR23 数据集上进行充分实验, 验证算法性能。通过多个消融实验, 验证各个模块的有效性。所有的训练和测试实验均在配备有 “Intel (R) Xeon (R) Gold 5218R CPU@2.10 GHz” CPU 和 “NVIDIA A40” GPU 的工作站上进行。

2.1 超参设置

使用 Swin-Transformer 中的 Swin-L 模型, 并下载在 ImageNet-22K 数据集上的预训练模型作为训练前的初始化参数, 以加快模型的收敛速度。不同尺度图像特征维度取值为: I_1 设置为 192, I_2 设置为 384, I_3 设置为 768, I_4 设置为 1 536。BEV 的特征通道 C_p 设置为 384。训练批次大小设置为 24。训练批次为 200, 学习率最初设置为 10^{-5} , 随着训练的增加, 在最后阶段衰减到 10^{-6} 。为扩展训练数据集的多样性, 主要对图像进行特征增强, 包括随机水平翻转和灰度变换, 保持相对应的点云部分。

2.2 实验数据

采用 OSDaR23 数据集进行实验, OSDaR23 数据集是 1 个专为无人驾驶列车在干线铁路上运行而设计的多传感器数据集, 2021 年 9 月在德国汉堡进行了数据采集, 并于 2023 年发布。数据集的传感器配置包括多个校准和同步的红外与视觉摄像机、激光雷达以及安装在列车前部的加速度传感器。这些传感器的组合能够提供丰富的数据源, 以支持算法的训练和测试, 实验

过程中, 采用高分辨率图像数据和其对应的点云数据, 训练测试数据集的数量比例为 8 : 2。

2.3 评价指标

在模型性能评估阶段, 选择平均交并比 (mIOU) 作为度量指标, 计算过程见式 (9)。

$$\text{mIOU} = \frac{1}{N} \sum_{n=1}^N \frac{\text{TP}_n}{\text{TP}_n + \text{FP}_n + \text{FN}_n}, \quad (9)$$

式中: mIOU 为平均交并比; N 为类别的数量, N 取 2 表示包括轨道和非轨道区域; TP_n 为类别 n 的真正例 (预测正确且实际为正的样本数); FP_n 为类别 n 的假正例 (预测为正但实际为负的样本数); FN_n 为类别 n 的假负例 (预测为负但实际为正的样本数)。

2.4 实验结果与分析

2.4.1 与主流方法性能比较

融合算法与其他主流基线方法在公共数据集上的分割性能对比见表 1。表中 “×” 表示未使用该模块, “√” 表示使用了该模块。可以看出, 该方法取得了最优的分割性能, 这主要得益于多传感器信息的有效融合。具体来说, 摄像头提取的 BEV 特征提供了丰富的语义信息, 而雷达提取的 BEV 特征则补充了精确的位置信息, 这种信息的互补性显著提高了模型的分割准确率。与次优方法 LSS 相比, 多传感器融合方法在 mIOU 指标上实现了 5.7% 的显著提升。这种性能差异主要归因于多模态信息的融合, 它增强了模型在复杂环境下的鲁棒性并提高了分割精度, 表明通过有效融合多源信息, 可有效解决单一传感器方案中常见问题, 比如视觉遮挡、光照变化以及恶劣天气条件等, 从而使得整体的分割性能更加稳定和可靠。但多传感器融合方法也存在一些潜在局限性, 例如计算成本较高、数据同步问题以及传感器故障可能带来的影响。尽管如此, 通过多传感器融合在分割任务中取得了显著进步, 未来的研究将着重于解决这些局限性, 以实现更高效、更可靠的分割性能。

2.4.2 消融实验

为验证各个模块的有效性, 进行了丰富的消融实验 (见表 2)。具体来说, 没有使用多传感器特征融合模块的消融实验是指仅使用图像通道完成分割任务, 即不考虑雷达提供的信息。从表 2 中可以看出, 采用 Swin-L 作为主干网络相较于传统的 ResNet 主干网络取

表1 轨道分割实验结果对比

方法	相机	雷达	mIOU
OFT ^[15]	√	×	0.834
LSS ^[16]	√	×	0.859
CVT ^[17]	√	×	0.846
Point Pillars ^[14]	×	√	0.801
Centerpoints ^[18]	×	√	0.854
融合算法	√	√	0.908

得了较高的性能提升，验证了Swin-L的全局注意力机制能够较好地注意到轨道区域，能够有效地区分轨面和轨旁。此外，多尺度特征融合模块的引入也带来了显著的性能提升。由于在不同的视角下，轨面面积会有所不同，因此多尺度信息对于捕捉不同大小的目标至关重要。最后，加入激光雷达特征后，模型的分割准确率提升了3.68%，证明了不同传感器的数据能够实现信息互补，进而为最终的分割模块提供更高质量的特征表示。激光雷达提供的精确位置信息与摄像头提供的丰富语义信息相结合，能够提高模型对复杂场景的理解能力。实验表明，多传感器融合方法在实际操作中具有优越性。

表2 消融实验结果

图像主干网络	多尺度特征融合模块	多传感器融合模块	mIOU
ResNet	×	×	0.821
Swin-L	×	×	0.847
Swin-L	√	×	0.876
Swin-L	√	√	0.908

2.4.3 轨道分割结果可视化

为了全面验证轨道分割方法的有效性，采用多种可视化技术直观地展示分割结果。具体来说，对不同场景下的轨道进行以下3种类型的可视化演示（见图5）。

（1）轨道分割结果的可视化，展示了轨道分割的直观结果。通过对比分析，可以明显观察到此算法在轨道识别和分割方面的优势，它能够有效地区分轨道和轨旁区域，确保分割结果的准确性和清晰度。

（2）图像与点云数据的坐标映射，便于进一步探索图像与点云数据之间的空间关系。通过坐标映射技术将点云数据准确地投影到对应的图像上，不仅增强了数据可视化效果，而且为后续的分析 and 处理提供了更加丰富的上下文信息。

（3）点云数据上的轨道分割结果可视化。为了更

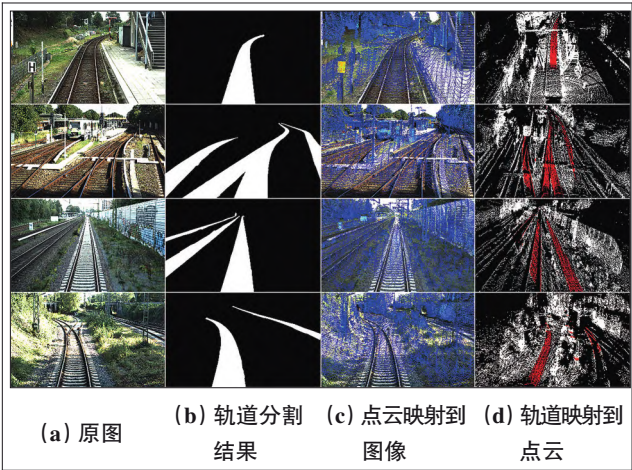


图5 轨道分割结果可视化

深入地展示分割效果，将分割结果直接映射到原始点云数据上，有效实现从二维图像到三维空间的转换。这种三维可视化不仅提供更为立体的视角，而且使得分割结果的验证更加直观和全面，也为实际应用提供更直观的可视化效果。

3 实际应用与系统集成探讨

主动障碍物检测是当下轨道交通的重点技术发展方向，能够为列车提供更安全的保障信息。融合算法在实际工程中主要面临2个问题：一方面是如何进行算法的边缘部署，另一方面是如何与现有系统交互集成。针对算法的边缘部署问题，基于TensorRT深度学习推理工具，算法已经在NVIDIA Drive Orin上测试通过，能够满足实时推理，因此选择Orin边缘计算模块。对于与现有系统交互集成问题，车载级别的主动障碍物感知能够为列车自动驾驶（ATO）和列车超速防护（ATP）提供障碍物信息，进而根据接收到的障碍物信息向司机发送降速运行和采取制动措施。目前用于ATO更合适，随着算法软件和硬件不断提升，如果主动障碍物检测技术达到SIL4级别，也可直接用于ATP。

4 结束语

针对当下单传感器方法信息局限的问题，提出1种融合算法。基于OSDaR23公开数据集与主流的神经网络进行对比，结果展示模型具有较好的分割效果。

通过研究提出1种多传感器融合的轨道分割算法结

构,有效解决数据信息局限的问题,提高了分割精度;设计1种多尺度特征融合模块,能够自适应的融合多尺度图像特征,性能得到明显提升;另外设计1种基于注意力机制的多传感器特征融合模块,实现图像和点云特征的自适应融合。

轨道分割是主动障碍物检测的基础,同时也弥补被动障碍物检测主动性不足的问题。在未来,将研究多传感器融合方法在不同实际轨道线路的分割性能,并集成到ATO中。通过扩大数据库,提高模型的泛化性能,保障轨道交通列车更加安全可靠运行。

参考文献

- [1] 张伟. CBTC与CTCS-2兼容互通方案[J]. 铁路通信信号工程技术, 2023, 20(6): 93-97.
- [2] 郑亚宏, 何家玉. 铁路线路障碍监测报警系统研究[J]. 中国铁路, 2019(9): 111-117.
- [3] 戎志立, 于苾健. 城市轨道交通主动障碍物检测预警防护系统研究[J]. 工程建设与设计, 2022(10): 113-115.
- [4] 侯智雄, 王昊, 郝蕊. 铁路轨道检测数据智能传输系统研制及应用研究[J]. 中国铁路, 2020(11): 133-138.
- [5] SHI J, MALIK J. Normalized cuts and image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905.
- [6] LATEEF F, RUICHEK Y. Survey on semantic segmentation using deep learning techniques[J]. Neurocomputing, 2019, 338(21): 321-348.
- [7] 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: Springer Nature, 2015.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014.
- [9] 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.
- [10] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [11] HE K, GKIOXARI G, DOLLÁR P, et al. Mask r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision: IEEE. Venice, 2017.
- [12] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021.
- [13] LI X, WANG W, HU X, et al. Selective kernel networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019.
- [14] LANG A H, VORA S, CAESAR H, et al. Pointpillars: fast encoders for object detection from point clouds[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019.
- [15] THOMAS R, ALEX K, ROTBERTO C. Orthographic feature transform for monocular 3D object detection[C]// In British Machine Vision Conference. Cardiff: BMVA, 2019.
- [16] PHILION J, FIDLER S. Lift, splat, shoot: encoding images from arbitrary camera rigs by implicitly unprojecting to 3d[C]// European Conference on Computer Vision. Online, 2020.
- [17] ZHOU B, PHILIPP K. Cross-view transformers for real-time map-view semantic segmentation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022.
- [18] YIN T, ZHOU X, PHILIPP K. Center-based 3D object detection and tracking[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Kuala Lumpur: IEEE, 2021.

责任编辑 白箬

收稿日期 2024-07-09

Track Segmentation Algorithm Based on Multi-sensor BEV Feature Fusion

SHI Haifeng, MIAO Zhuangzhuang, YAN Zhengxiang, HAN Gaitang, YANG Lianbao, WANG Xiaoliang
(Beijing HollySys Co., Ltd., Beijing 100176, China)

Abstract: Track segmentation technology serves as the core support for the active obstacle detection by trains. Currently, the prevalent track segmentation methods primarily rely on camera sensors, which are unable to capture depth information of track positions, thereby impairing segmentation accuracy. To address this issue, we propose a track segmentation algorithm based on multi-sensor bird's eye view (BEV) feature fusion, which enhances the perception of track data by integrating input from both cameras and LiDAR. First, features are extracted from the image data captured by cameras and the point cloud data generated by LiDAR. Next, a fusion module utilizing an attention mechanism is designed to integrate these two types of data features. Finally, the fused features are processed through a linear layer to achieve precise track segmentation. To validate the effectiveness of the algorithm, training and testing are conducted using the OSDaR23 dataset. The results demonstrate that this method exhibits outstanding segmentation performance.

Keywords: track segmentation; radar; camera; bird's eye view; multi-sensor fusion; attention mechanism

(上接第 73 页)

Research on Standardization of Recyclable Multi-point Support Swivel Equipment for Railway Girder Bridges

XIONG Tao, LI Qianming, MA Xingchuan

(China Railway Major Bridge Reconnaissance & Design Institute Co., Ltd., Wuhan Hubei 430074, China)

Abstract: For the horizontal swivel construction of bridges with structural forms such as T structure and continuous beams (rigid frames), the multi-point support swivel scheme with auxiliary legs at the edge spans offers advantages such as good stability against overturning and minimal impact on existing traffic. This approach has been successfully applied in several overpass bridge projects in China. However, the diverse range of one-time-use swivel equipment is not conducive to the promotion of advanced technologies and is inconsistent with the concepts of green and intelligent construction. In order to address the issues of standardized manufacturing, installation, and recycling of multi-point support swivel equipment for railway girder bridges, this paper takes the cross-line swivel bridge of the Shenzhen-Nanning HSR as an example. Using comparative analysis, theoretical approaches, and numerical calculations, the research focuses on the selection of detachable central spherical hinges, the rational arrangement and structural dimensions of auxiliary legs, and the mechanical performance and structural requirements of the drive devices. The goal is to explore and establish the form and standardized parameters of swivel equipment suitable for railway girder bridges with a swivel mass of less than 10 000 tons. The main conclusions are as follows: the central spherical hinge should use a support-type spherical hinge, with its planar dimensions determined by the concrete strength of the lower turntable. The auxiliary legs should be placed 25 m from the swivel center of the edge span, adopting a steel pipe column form and making contact with the slideway via a roller trolley. The counterweight for cantilever spans of 40 m to 80 m corresponds to values between 360 to 600 kN. The motor gear rack drive system can meet the power requirements for normal swivel.

Keywords: railway girder bridge; horizontal swivel construction; multi-point support; recyclable type; swivel equipment; gear rack; standardization