

基于图像引导的自动驾驶场景下深度补全算法研究

仇毅

西华大学汽车与交通学院，四川 成都

收稿日期：2025年10月14日；录用日期：2025年12月29日；发布日期：2026年1月6日

摘要

深度补全旨在结合稀疏深度测量与彩色图像以恢复出高分辨率的密集深度图像，这一技术在自动驾驶场景中至关重要。大多数现有方法均基于空间传播机制，即对初始估计的密集深度进行迭代精化。然而初始深度估计通常直接通过常规卷积来提取特征，导致在处理空值区域时易产生卷积噪声以及物体边界恢复存在不足。为解决上述问题，本文提出一种双分支多尺度深度补全网络。该方法采用“插值 + 更新”的两阶段策略：在深度插值分支中提出引导双边插值，并结合子流形稀疏卷积与空洞空间金字塔池化生成初始稠密深度图；同时引入通道 - 空间注意力机制，实现跨模态特征的动态加权。在更新分支中，利用迭代传播机制对初始结果进行优化，从而提升全局一致性与局部细节恢复。在KITTI数据集上的验证结果表明，所提出的双分支多尺度深度补全网络优于其他主流方法。

关键词

深度补全，自动驾驶，引导双边插值，稀疏卷积，注意力机制

Research on Image-Guided Depth Completion Algorithms for Autonomous Driving

Yi Qiu

School of Automotive and Transportation, Xihua University, Chengdu Sichuan

Received: October 14, 2025; accepted: December 29, 2025; published: January 6, 2026

Abstract

Depth completion aims to combine sparse depth measurements with RGB images to recover high-

文章引用：仇毅. 基于图像引导的自动驾驶场景下深度补全算法研究[J]. 人工智能与机器人研究, 2026, 15(1): 87-96.
DOI: [10.12677/airr.2026.151010](https://doi.org/10.12677/airr.2026.151010)

resolution dense depth maps. This technology is crucial for autonomous driving scenarios. Most existing approaches rely on spatial propagation mechanisms that iteratively refine the initial dense depth estimates. However, initial depth estimation typically extracts features directly through conventional convolutions, which makes it prone to convolution noise when handling empty regions and less effective in reconstructing object boundaries. To address these issues, this paper proposes a Dual-branch Multi-scale Depth Completion Network (DM-Net). The method adopts a two-stage “interpolation + update” strategy. In the interpolation branch, guided bilateral interpolation is introduced and combined with Submanifold Sparse Convolution and Atrous Spatial Pyramid Pooling to generate the initial dense depth map. Meanwhile, a Convolutional Block Attention Module is incorporated to achieve dynamic weighting of cross-modal features. In the update branch, an iterative propagation mechanism is employed to refine the initial results, thereby enhancing global consistency and local detail recovery. Experiments on the KITTI dataset demonstrate that the proposed dual-branch multi-scale depth completion network outperforms other mainstream methods.

Keywords

Depth Completion, Autonomous Driving, Guided Bilateral Interpolation, Sparse Convolution, Attention Mechanism

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 前言

准确且稠密的场景深度数据在自动驾驶领域[1]中得到了广泛应用。然而受限于现有硬件能力，直接获取密集深度图仍具挑战。例如，使用激光雷达获取的深度过于稀疏，无法直接使用；而深度相机虽然能够获取稠密深度，但其因精度不足，难以满足自动驾驶领域的实际应用需求。因此，补全采集到的深度图中缺失的部分是相当必要的。该任务的关键挑战在于如何从不规则采样的稀疏深度点中高效恢复稠密且准确的深度。

近年来，基于深度学习的深度补全方法表现出了显著优势，根据输入数据的类型可以分为两类：非图像引导法和图像引导法[2]。前者直接基于稀疏 LiDAR 数据生成稠密深度，但当深度数据极为稀疏时，缺乏邻近结构与边缘先验信息，难以恢复完整的深度图。相比之下，图像引导方法通过引入与稀疏深度图配对的彩色图像作为输入的一部分引导深度补全，借助彩色图像中丰富的信息，大幅改善了物体边界预测结果模糊变形、不规则形状物体预测缺失等情况。因此图像引导法能够显著提升深度补全性能，并逐渐成为主流方法[3]。

图像引导的深度补全方法主要沿着多模态融合与空间传播网络两个方向发展。在多模态融合方面，Tang 等人提出的 GuideNet [4]通过引导卷积在不同尺度上利用图像特征动态更新深度特征，显著提升了彩色图像对深度补全的引导能力。Zhang 等人提出的 Cformer [5]将 CNN 与 Transformer 结合，有效缓解了单一 Transformer 在局部信息提取上的不足。此外，国内学者也开展了大量基于多模态融合的研究[3][6]-[8]，通过分别提取图像与稀疏深度特征来提升深度补全性能。在空间传播网络(SPN [9])方向，Cheng 等人先后提出 CSPN [10]和 CSPN++ [11]，其中 CSPN 使用卷积运算实现高效传播，CSPN++则在此基础上引入预定义窗口实现自适应优化。Park 等人提出的 NLSPN [12]通过非局部邻域实现了灵活的深度传播；Wang 等人提出的 LRRU [13]采用可变形卷积实现亲和度矩阵的内核可变性。

然而，现有图像引导方法仍存在两个主要问题：其一，尽管高分辨率图像蕴含丰富的边界信息，但大多数方法依赖常规卷积提取其特征，导致图像中边界信息挖掘不足，同时对稀疏深度直接卷积容易因空值点传播产生噪声，影响初始深度估计质量；其二，多数方法仅通过通道拼接或逐元素求和等简单操作融合稀疏深度与图像信息，忽略了跨模态特征融合的动态性与有效性。

针对上述问题，本文提出了一种双分支多尺度网络(Dual-branch Multi-scale Network, DM-Net)，整体由深度插值分支和深度更新分支构成。其中，深度插值分支采用引导双边插值对稀疏深度图进行填充，以获得初始的密集深度图，从而有效避免直接对稀疏数据进行卷积时因无效点传播带来的噪声干扰。引导双边插值首先利用引导滤波器提取图像中丰富的边缘与纹理信息，再结合原始图像与稀疏深度的特征，用以生成插值过程中的权重与偏置。具体而言，引导双边插值以稀疏深度图中有效深度点所对应的图像像素为导向，在目标像素的局部邻域内自适应地进行深度插值。这样一来，插值结果不仅考虑空间距离，还能依据图像边缘与深度差异进行动态调整。当像素处于物体边界或深度突变区域时，引导双边插值会自动抑制跨边界的深度传播，强化同一物体内部的结构相关性。得益于这一机制，引导双边插值在边界区域具有更强的结构保持能力，能够在稀疏区域恢复出更加连续和平滑的深度过渡。基于此，深度更新分支进一步引入 CSPN++对初始插值结果进行迭代优化。由于引导双边插值已在前阶段生成了边界友好的初始稠密深度图，CSPN++在传播过程中能够更准确地利用邻域像素间的结构关系，从而实现细节的精细恢复与边缘的锐化。换言之，引导双边插值的边界一致性为后续深度传播提供了更加稳定且可靠的结构先验，显著提升了整体的深度补全质量。此外，针对稀疏深度图的离散特性，网络结合子流形稀疏卷积与空洞空间金字塔池化，以高效提取稀疏深度特征；在跨模态特征融合阶段，引入卷积块注意力模块，实现对彩色图像与稀疏深度特征权重的自适应调整，进一步提升融合特征的辨识能力与深度估计精度。

本文的主要贡献如下：

- 提出一种双分支多尺度深度补全网络，通过引导双边插值生成初始稠密深度，并结合 CSPN++迭代优化，兼顾边界恢复与全局一致性；
- 设计结合子流形稀疏卷积与空洞空间金字塔池化模块，将空洞空间金字塔池化模块中的普通卷积替换为相结合，有效增强稀疏深度的多尺度表征能力；
- 引入卷积块注意模块，在图像与深度特征融合过程中实现跨通道与空间的动态加权。

2. 双分支多尺度深度补全网络

双分支多尺度深度补全网络的整体结构如图 1 所示。首先，利用引导双边插值对稀疏深度图进行填充，生成初始稠密深度图(见第 2.1 节)；其次，结合子流形稀疏卷积与空洞空间金字塔池化模块，用于提取稀疏深度的多尺度特征(见第 2.2 节)；在跨模态特征融合阶段，引入卷积注意力模块，实现对彩色图像与稀疏深度特征的通道 - 空间自适应加权(见第 2.3 节)；最后，在第 2.4 节中介绍了具体的实施细节。

2.1. 引导双边插值

设彩色图像记为 I ，通过标定参数将三维点投影到图像平面上所获得的稀疏深度图记为 S 。彩色图像 I 提供了丰富的场景上下文信息，具有清晰的边界与精确的语义；相比之下，深度图 S 虽然与彩色图 I 具有相同的分辨率，但其分布稀疏且离散，仅包含有限的深度点。但这些深度测量通常具有较高的准确性和确定的几何约束。因此为充分利用两者的互补特性，本文设计了引导双边插值模块，以插值生成初始稠密深度图 D' 。

如图 2 所示，引导双边插值方法首先通过 Sobel 算子提取图像的增强边缘特征，并与原始图像一同输入引导滤波器[14]，得到强化边界特征的引导滤波图，随后，通过卷积得到边界特征 e_j ，用于强化邻域

像素的边缘相关性，接着，将边缘特征与图像和稀疏深度信息融合后，进行加权插值得到目标深度 D'_i ，其计算公式如下：

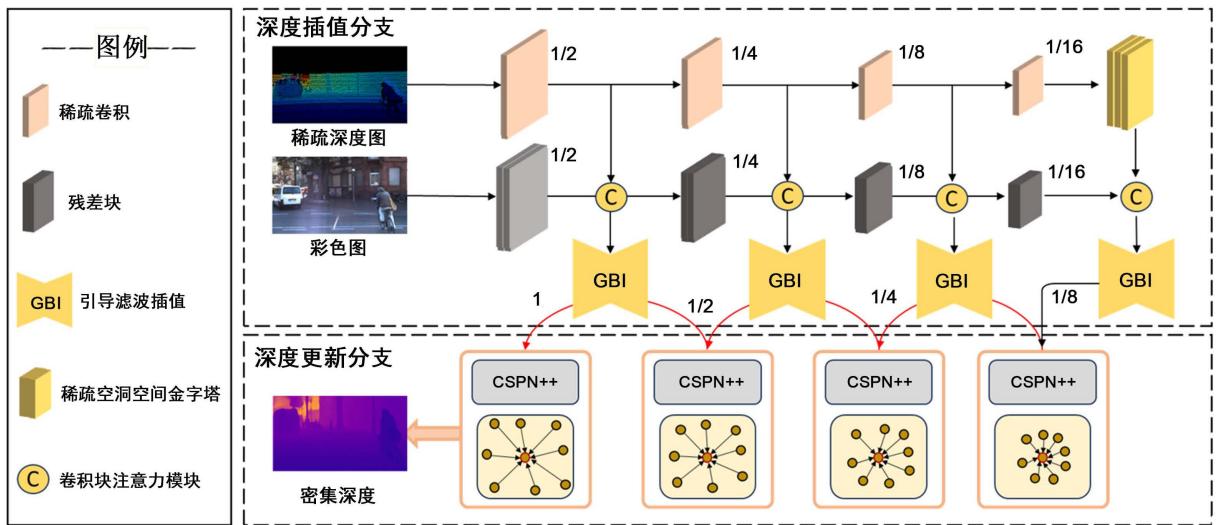


Figure 1. Structure diagram of the dual-branch multi-scale depth completion network (DM-Net)

图 1. 双分支多尺度深度补全网络(DM-Net)结构图

$$D'_i = \frac{\sum_{j \in \mathcal{N}(i)} e_{ij} D_{ij}^c}{\sum_{j \in \mathcal{N}(i)} e_{ij}} \quad (1)$$

其中， D'_i 表示目标像素 i 的初始稠密深度， D_{ij}^c 为目标像素 i 的邻近像素 j 的候选深度，由像素 i 邻域内 $N=4$ 个有效深度点的线性组合生成，其计算公式如式 2 所示：

$$D_{ij}^c = \alpha_{ij} S_j + \beta_{ij}, \quad j \in N \quad (2)$$

其中 S_j 为邻近像素 j 所对应的深度值， α_{ij} 为线性组合的权重稀疏系数， β_{ij} 为偏移稀疏。其系数 α_{ij} 与 β_{ij} 由多层感知机(MLP)生成：

$$\alpha_{ij}, \beta_{ij} = \text{MLP}(I_i, I_j, S_j, O_{ij}) \quad (3)$$

通过设定一个两层隐藏层的多层感知机(MLP)实现，MLP 的输入包括目标像素 i 与源像素 j 的图像编码 I ，源像素 j 的深度编码 S ，以及像素 i 到 j 的空间偏移编码 O_{ij} ，MLP 在每层后使用 GELU 激活函数引入非线型性，输出对应系数 α_{ij} 与 β_{ij} 。

引导双边插值巧妙的将目标深度 D'_i 转化为可学习的参数化表述，实现从稀疏深度到稠密深度的自适应估计。与传统插值方法不同，能够根据图像内容和空间位置动态生成目标深度，从而适应稀疏且不规则的深度分布。具体而言，多层感知机会根据目标像素与邻域像素的深度差异、图像特征以及空间偏移动态调整 α_{ij} 与 β_{ij} ，在深度差异较大时会自动抑制邻域深度的影响，从而直接回归更可靠的目标深度。同时，由引导滤波得到的边缘权重系数 e_{ij} 又可进一步强化边缘特征，确保边界结构的一致性。这种机制使得在表达能力和鲁棒性上均优于传统插值方法。

通过上述机制，引导双边插值利用引导滤波提取边缘特征的同时，有效融合稀疏深度的空间信息与彩色图像的纹理信息，从而在保持边缘结构一致性的前提下，提高了全局初始稠密深度的精度。

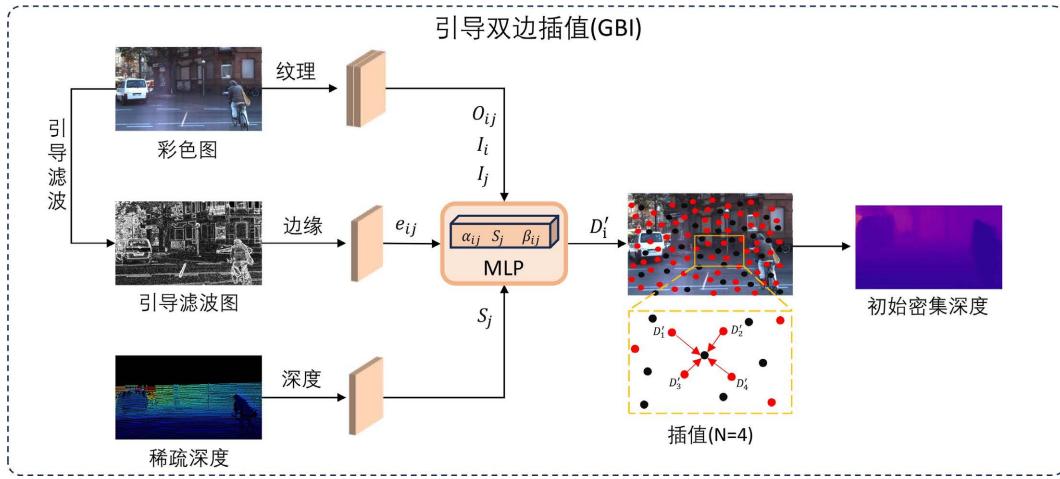


Figure 2. Schematic diagram of the guided bilateral interpolation module
图 2. 引导双边插值模块示意图

2.2. 子流形稀疏卷积与空洞空间金字塔池化

在 KITTI-DC 数据集中，将稀疏点云投影到对应的 RGB 图像后，约有 3% 的像素位置具备有效的深度信息[15]。若直接采用常规卷积对其进行特征提取，类似于对普通图像的处理方式，容易引入大量无效卷积操作，难以有效捕获深度补全任务中所需的关键深度特征。为克服传统卷积在处理稀疏深度图时的不足，本文将传统的普通卷积替换为子流形稀疏卷积[16]，并构建了一种新的特征提取模块子流形稀疏卷积与空洞空间金字塔池化单元。该模块兼具子流形稀疏卷积在稀疏结构保持方面的优势与空洞空间金字塔池化单元的多尺度特征提取能力，可以高效地从稀疏数据中提取深层特征。

具体而言，针对输入深度图的稀疏与离散特性，模块采用更适合稀疏数据的子流形稀疏卷积替代传统卷积。该卷积机制通过限制输出点的激活范围，仅在有效深度点上执行卷积运算，从而准确定位有效点并避免冗余计算。在保证深度点空间特征完整性的同时，有效降低了计算复杂度[16]。此外，在网络的深层部分引入空洞卷积以扩展感受野(如图 3 所示)，并通过设置膨胀率为 2、4 和 8 的子流形卷积来在不同尺度上捕获局部特征，同时减少模型参数量。随后，采用全局平均池化提取全局语义信息，最终通过跳跃连接将稀疏深度特征与 Conv2D 层进行融合，以防止信息丢失。

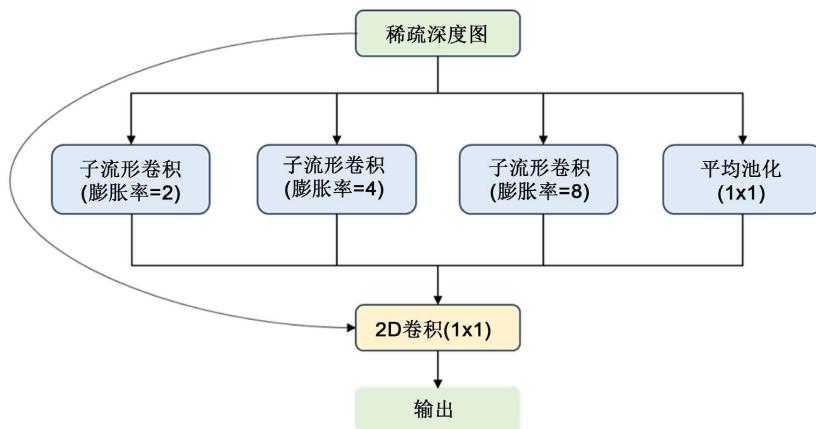


Figure 3. Schematic diagram of sparse convolution and atrous spatial pyramid pooling
图 3. 稀疏卷积与空洞空间金字塔池化示意图

2.3. 卷积块注意力模块

在现有大多数图像引导的深度补全方法中，多数方法仅通过通道拼接或逐元素求和等简单操作融合稀疏深度与图像信息，这种融合方式不仅没有考虑到不同场景下不同模态特征的置信度，还难以充分利用有效深度点的信息，引入大量无效特征，从而降低深度预测的准确性。为解决这一问题，本研究在稀疏深度与图像特征融合阶段引入卷积块注意力模块，该模块通过自适应调整通道和空间权重，在有效深度点区域增强点云特征的重要性，同时在无效深度点区域提升图像特征的贡献，从而整体上提高深度补全的精度与鲁棒性。

具体而言，卷积块注意力模块由通道注意力和空间注意力两个顺序子模块串联构成[17]（如图4所示），通道注意力模块利用特征间的通道关系生成通道注意力图，为高效计算通道注意力，采用最大池化与平均池化操作对输入特征图的空间维度进行压缩。空间注意力模块则利用特征间的空间关系生成空间注意力图。与通道注意力不同，空间注意力更关注特征的“位置”信息，与通道注意力形成互补。二者结合后，能够根据像素点与深度点的综合信息动态调整各通道的重要性。在有效深度区域强化点云特征，在空缺区域增强图像引导，从而兼顾几何精度与纹理细节，该自适应重标定机制有效提升了模型在有限且关键的深度信息上的关注度，从而增强深度补全任务的预测性能。

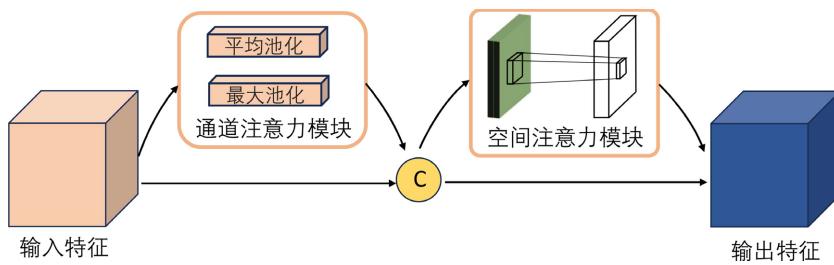


Figure 4. Diagram of convolutional block attention module (CBAM)
图 4. 卷积块注意力模块示意图(CBAM)

2.4. 实施细节

2.4.1. 损失函数

本次实验采用的是 $\ell_1 + \ell_2$ 组合计算损失[12]，这种方式可以同时发挥 ℓ_1 损失函数和 ℓ_2 损失函数的优势，在深度补全网络模型训练时的快速收敛和取得优良性能之间取得了相当的平衡。真实深度为 D_{gt} ，为了避免无效深度对损失的干扰，仅在有效像素处计算误差，定义掩码函数：

$$\mathbf{1}_{\{D_{gt} > 0\}}(x) = \begin{cases} 1, & D_{gt}(x) > 0 \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

在每一层，计算预测与真值之间的 L1 与 L2 误差，并通过逐像素相乘 \odot 与掩码结合：

$$\ell_{\sigma}^{(i)} = \left\| (\hat{D}^i - D_{gt}) \odot \mathbf{1}_{\{D_{gt} > 0\}} \right\|_{\sigma}, \quad \sigma \in \{1, 2\} \quad (5)$$

其中：

- $\sigma = 1$ 表示 L1 范数，对应 MAE；
- $\sigma = 2$ 表示 L2 范数，对应 RMSE；

为了强调深层预测结果的重要性，同时保证浅层预测的约束作用，我们为第 i 层引入衰减系数 γ ，本文实验中设置 $\gamma = 0.8$ ，因此最后一层的权重最大，更加突出最终结果的优化。整体损失函数定义为：

$$L = \sum_{\sigma=1}^2 \sum_{i=1}^N \gamma^{N-i} \left\| \left(\hat{D}^i - D_{gt} \right) \odot \mathbf{1}_{\{D_{gt} > 0\}} \right\|_{\sigma} \quad (6)$$

其中：

- N 表示网络的输出层数，对应为 4；
- $\mathbf{1}_{\{D_{gt} > 0\}}$ 表示仅在有效深度像素点处计算损失。

2.4.2. 训练设置

本文基于 PyTorch 框架实现所提出的方法，并在配备 2 张 Nvidia RTX 3090 GPU 的工作站上进行训练。网络结构主要由残差块与稀疏卷积堆叠而成，并在残差块的跳跃连接前引入 Dropout 作为正则化手段，以增强模型训练的稳定性。优化器采用 AdamW，权重衰减系数设为 0.05，同时对梯度进行裁剪，当梯度的 L2 范数超过 0.1 时进行约束。模型输入为 KITTI-DC 训练集，从零初始化开始训练，总训练轮次(epoch)为 40，学习率按照 OneCycle 策略调整，最终降低至最大学习率的 25%，批量大小设为 32，最大学习率为 0.001。

3. 验证结果及分析

3.1. 验证数据集

为了全面验证所提方法的有效性，本实验选取了广泛应用于自动驾驶研究的 KITTI-DC 数据集[15]，该数据集由车载平台在城市、乡村及高速公路等多样化场景下采集，包含真实的 RGB 图像、原始稀疏深度图以及对应的半密集深度图，图像分辨率为 1216×352 ，每张图像最多包含约 15 辆车辆和 30 个行人，并具有不同程度的遮挡和截断情况。KITTI 数据集共提供超过 93,000 组样本，按照训练集、验证集和测试集分别划分为 86,000、7000 和 1000 组，其中稀疏深度由 Velodyne HDL-64E 激光雷达采集，半密集深度图通过对原始数据进行半全局匹配(semi-global matching, SGM)纯化并聚合 11 帧连续数据生成，有效点约占 30%，在保证真实感的同时兼顾了稀疏性与稠密性。

3.2. 评估指标

KITTI 中关于深度补全数据集包含真值作为监督信息用于训练模型和评估，评价指标通常使用真实深度图 GT 中的有效深度值与预测深度图 P 中对应位置的深度值进行比较，根据不同的距离衡量标准，可以分为均方误差(RMSE)、平均绝对误差(MAE)、逆深度均方误差(iRMSE)和逆深度平均绝对误差(iMAE)四个指标，其中 RMSE 为主要的评估指标。

均方误差如式(7)：

$$RMSE = \sqrt{\frac{1}{|V|} \sum_{v \in V} |GT_v - P_v|^2} \quad (7)$$

平均绝对误差如式(8)：

$$MAE = \frac{1}{|V|} \sum_{v \in V} |GT_v - P_v| \quad (8)$$

逆深度均方误差如式(9)：

$$iRMSE = \sqrt{\frac{1}{|V|} \sum_{v \in V} \left| \frac{1}{GT_v} - \frac{1}{P_v} \right|^2} \quad (9)$$

逆深度平均绝对误差如式(10):

$$\text{iMAE} = \frac{1}{|V|} \sum_{v \in V} \left| \frac{1}{GT_v} - \frac{1}{P_v} \right|^2 \quad (10)$$

其中, 其中 V 是 GT 中有效点的集合。 GT_v 为真值中有效点的深度值, P_v 即预测值中对应位置的深度值。

3.3. 验证结果分析

表 1 展示了本文方法与当前主流深度补全方法在 KITTI-DC 基准上的定量评估结果, 所有方法均按照 RMSE 指标进行排名。由于从稀疏深度图和 RGB 图直接预测稠密深度图是一个高度病态的问题[13], 类似与 GuideNet [4]这种直接回归类方法往往模型参数规模庞大, 推理时间较长, 而其性能往往难以达到最优水平, 为提升直接回归类方法的表现, 研究者提出了一系列空间传播网络作为后处理模块对预测结果进行细化, 例如 CSPN [10]、CSPN++ [11] 和 NLSPN [12]。这些方法在一定程度上提升了结果精度, 并逐渐成为深度补全的主流方向。然而, 它们仍难以避免普通卷积在稀疏深度图中处理空值区域时产生的卷积噪声问题, 同时对彩色图像边界信息的利用不足, 导致物体边界的恢复效果有限。

Table 1. Comparative test results of various methods on the KITTI dataset

表 1. 在 KITTI 数据集上各方法对比试验结果

Methods	RMSE↓ [mm]	MAE↓ [mm]	iRMSE↓ [1/km]	iMAE↓ [1/km]
CSPN [10]	1019.64	279.46	2.93	1.15
GuideNet [4]	736.24	218.83	2.25	0.99
CSPN++ [11]	743.69	209.28	2.07	0.9
NLSPN [12]	741.68	199.59	1.99	0.84
CFormer [5]	708.87	203.45	2.01	0.88
DM-Net	698.26	195.54	1.97	0.82

注: 最佳结果以黑色加粗标出。

相比之下, 本文提出的双分支多尺度深度补全网络展现出更高的有效性。具体而言, 该网络在 KITTI-DC 的所有评估指标(RMSE、MAE、iRMSE 与 iMAE)上均取得了优异表现。该性能提升的核心在于 DM-Net 的双分支协同设计, 尤其是深度插值分支中引导双边插值模块的创新应用: 通过强化图像的边界信息并结合稀疏深度特征以生成高质量初始稠密深度图, 不仅为后续 CSPN++的迭代优化提供了可靠先验, 更从根本上规避了直接对稀疏深度图进行普通卷积所带来的无效点传播、噪声扩散及结构模糊等问题。同时, 引导双边插值模块与 CSPN++的协同机制有效提升了深度补全的细节恢复能力与边缘锐化效果, 为稀疏深度补全任务提供了高效可行的解决方案。

图 5 展示了 KITTI-DC 验证集上的可视化结果对比。本文采用各主流方法作者提供的公开代码及其最佳模型进行复现, 并保持输入稀疏深度一致。结果显示, 本文方法在物体边界及细节区域的深度预测更为清晰, 而其他方法在这些复杂区域仍存在明显不足, 从而进一步验证了本方法在边界恢复上的有效性。

4. 结束语

本文针对自动驾驶场景中深度补全过程中因深度点稀疏导致空值区域易产生卷积噪声以及物体边界



Figure 5. Visualization comparison of KITTI-DC dataset test results
图 5. KITTI-DC 数据集测试结果可视化对比

恢复不足等问题，提出了一种双分支多尺度深度补全网络。该方法通过“插值 + 更新”的两阶段策略，在插值分支中结合引导双边插值、子流形稀疏卷积与空洞空间金字塔池化，实现初始稠密深度估计；同时引入通道 - 空间注意力机制，动态增强跨模态特征融合。在更新分支中，利用迭代传播机制进一步优化结果，从而在保持全局一致性的同时提升了局部细节恢复能力。最后，在 KITTI 数据集的实验结果表明，DM-Net 相较于现有主流方法在 RMSE、MAE、iRMSE 和 iMAE 等指标上均取得了较好性能，显著优于 CSPN、GuideNet、NLSPN 等方法。

参考文献

- [1] Song, Z., Lu, J., Yao, Y. and Zhang, J. (2022) Self-Supervised Depth Completion from Direct Visual-Lidar Odometry in Autonomous Driving. *IEEE Transactions on Intelligent Transportation Systems*, **23**, 11654-11665. <https://doi.org/10.1109/tits.2021.3106055>
- [2] Hu, J., Bao, C., Ozay, M., et al. (2022) Deep Depth Completion from Extremely Sparse Data: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 8244-8264.
- [3] 付志超. 基于彩色图像语义信息引导的深度补全研究[D]: [博士学位论文]. 上海: 华东师范大学, 2024.
- [4] Tang, J., Tian, F., Feng, W., Li, J. and Tan, P. (2021) Learning Guided Convolutional Network for Depth Completion. *IEEE Transactions on Image Processing*, **30**, 1116-1129. <https://doi.org/10.1109/tip.2020.3040528>
- [5] Zhang, Y., Guo, X., Poggi, M., Zhu, Z., Huang, G. and Mattoccia, S. (2023) CompletionFormer: Depth Completion with Convolutions and Vision Transformers. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, 17-24 June 2023, 18527-18536. <https://doi.org/10.1109/cvpr52729.2023.01777>
- [6] 徐杰杰. 基于特征融合的单目深度估计和深度补全研究[D]: [硕士学位论文]. 南京: 南京信息工程大学, 2023.
- [7] 陈思远. 基于激光雷达与摄像头融合的深度补全算法研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2022.
- [8] 王海廷. 基于多视角融合和注意力机制的深度补全算法研究[D]: [硕士学位论文]. 大连: 大连理工大学, 2023.
- [9] Liu, S., De Mello, S., Gu, J., et al. (2017) Learning Affinity via Spatial Propagation Networks. arXiv: 1710.01020.
- [10] Cheng, X., Wang, P. and Yang, R. (2020) Learning Depth with Convolutional Spatial Propagation Network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 2361-2379. <https://doi.org/10.1109/tpami.2019.2947374>

-
- [11] Cheng, X., Wang, P., Guan, C. and Yang, R. (2020) CSPN++: Learning Context and Resource Aware Convolutional Spatial Propagation Networks for Depth Completion. *Proceedings of the AAAI Conference on Artificial Intelligence*, **34**, 10615-10622. <https://doi.org/10.1609/aaai.v34i07.6635>
 - [12] Park, J., Joo, K., Hu, Z., Liu, C. and So Kweon, I. (2020) Non-Local Spatial Propagation Network for Depth Completion. In: Vedaldi, A., Bischof, H., Brox, T. and Frahm, J.M., Eds., *Computer Vision—ECCV 2020*, Springer, 120-136. https://doi.org/10.1007/978-3-030-58601-0_8
 - [13] Wang, Y., Li, B., Zhang, G., Liu, Q., Gao, T. and Dai, Y. (2023) LRRU: Long-Short Range Recurrent Updating Networks for Depth Completion. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, 1-6 October 2023, 9388-9398. <https://doi.org/10.1109/iccv51070.2023.00864>
 - [14] Li, Y., Huang, J., Ahuja, N. and Yang, M. (2019) Joint Image Filtering with Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**, 1909-1923. <https://doi.org/10.1109/tpami.2018.2890623>
 - [15] Uhrig, J., Schneider, N., Schneider, L., Franke, U., Brox, T. and Geiger, A. (2017) Sparsity Invariant CNNs. *2017 International Conference on 3D Vision (3DV)*, Qingdao, 10-12 October 2017, 11-20. <https://doi.org/10.1109/3dv.2017.00012>
 - [16] Graham, B. and Van der Maaten, L. (2017) Submanifold Sparse Convolutional Networks. arXiv: 1706.01307.
 - [17] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1