# $R^3D$：区域引导残差雷达扩散

Hao Li，亚利桑那大学 Xinqi Liu，香港大学 Yaoqing Jin，斯图加特大学 lihao@arizona.edu Xinqi_liu28@connect.hku.hk st186181@stud.uni-stuttgart.de

arXiv:2601.06465v1 [eess.IV] 10 Jan 2026

**摘要—毫米波雷达能够在恶劣条件下为自主系统提供稳健的环境感知能力，但其点云稀疏、噪声较多且角分辨率较低。现有的基于扩散的雷达增强方法，要么通过对完整激光雷达分布建模而带来较高的学习复杂度，要么因采用均匀的区域处理而无法优先处理关键结构。为解决这些问题，我们提出 R3D，一种区域引导残差雷达扩散框架。该框架结合了残差扩散建模——聚焦于集中的激光雷达-雷达残差编码，以补充高频细节并降低学习难度——以及 $\sigma$自适应区域引导：利用雷达特有的信号特性生成注意力图，并仅在低噪声阶段施加轻量级引导，以避免梯度失衡，同时细化关键区域。在 ColoRadar 数据集上的大量实验表明，$R^3D$ 优于当前最先进的方法，为雷达感知增强提供了一种实用方案。我们的匿名代码和预训练模型已在此发布：https://anonymous.4open.science/r/r3d-F836**

索引词—毫米波雷达；点云增强；扩散模型；残差学习；区域引导

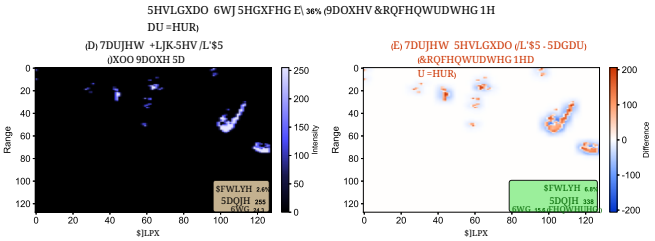


图1. 残差学习将目标分布集中在零附近。左：高分辨率激光雷达目标（有效区域：2.6%，距离：255，完整数值分布）；右：残差（激光雷达－雷达）目标（有效区域：6.8%，距离：338（绝对距离），标准差：15.6，以零为中心）。尽管残差具有更大的绝对距离，但其有效值中有93.2%位于 ±10 以内（高度集中于零附近），从而导致显著更小的标准差。与建模完整激光雷达分布相比，这种集中性使残差学习容易得多，因为模型只需预测偏离零的小幅变化，而不必预测整个数值范围。

## I. 引言

毫米波雷达已成为自主系统（例如微型飞行器（MAV）和高级驾驶辅助系统（ADAS））不可或缺的传感模态，因为它们能够抵御雨、雾和弱光等极端天气条件的影响。然而，由于角分辨率低、信号分布稀疏以及噪声干扰严重等固有限制，雷达点云的稠密度和精度远低于激光雷达点云。这一缺陷阻碍了毫米波雷达在杂乱环境中的应用，而此类环境需要精确的同步定位与建图（SLAM）和环境感知。

为解决毫米波雷达数据稀疏和噪声问题，近期研究探索了跨模态学习框架，利用激光雷达点云作为真实值，引导雷达点云增强。在这些方法中，

扩散模型已成为强大的生成工具，通过对雷达距离-方位角热图（RAHs）到激光雷达鸟瞰图（BEV）图像之间的映射进行建模，在雷达点云生成方面取得了最先进的性能。尽管取得了成功，现有的雷达扩散方法仍面临两个限制其实用性和性能的关键挑战。

首先，现有的雷达扩散模型直接对以雷达数据为条件的激光雷达点云完整分布进行建模，导致学习复杂度较高。这类方法要求模型学习激光雷达的所有空间特征，包括与雷达信号重叠的冗余信息。这不仅增加了计算开销，还分散了模型对高频细节（例如物体边缘和结构边界）的关注，而这些正是雷达本身所缺乏的。在超分辨率领域，人们普遍认为，恢复残差（即高分辨率数据与低分辨率数据之间的差异）要比从头生成高分辨率内容容易得多（如图 I 所示）——这是因为残差主要编码互补的高频细节，而非冗余的基础信息。然而，这一成熟范式尚未扩展到用于点云增强的雷达扩散领域。

其次，现有的扩散驱动雷达增强技术在去噪过程中对输入的所有区域采用统一的处理策略。雷达 RAH 具有明显的空间异质性：强散射区域（例如建筑物、车辆）包含关键的环境结构信息，而弱散射区域则以噪声和杂波为主。尽管在早期去噪阶段，统一处理对于学习全局结构是必要的，但在整个扩散过程中始终对所有区域不加区分地采用相同的优化重点，会在噪声主导的区域浪费计算资源，并可能无法在细节恢复阶段充分优化关键结构区域。

尽管已有研究提出注意力引导扩散方法来解决图像超分辨率中的这一问题，但这些方法依赖面向自然图像的自监督框架（例如 DINO），将其应用于雷达信号时会遭遇领域不匹配，而且复杂的预训练要求也与雷达数据集有限的可用性不相适应。与之形成鲜明对比的是，雷达信号属性——例如强度和局部一致性——无需

领域自适应或预训练，便可直接识别目标区域。然而，在整个扩散过程中不加区分地采用激进的区域加权，会破坏噪声级别相关的损失加权（例如 Karras 加权 $w(\sigma_t) = 1/\sigma_t^2$）所形成的梯度平衡，进而导致训练不稳定性。因此，需要一种有原则的方法，使区域引导与扩散模型固有的优化动态相协调。

基于这些观察，我们提出 R$^3$D——区域引导残差雷达扩散（Regional-guided Residual Radar Diffusion），这是一种新颖的跨模态扩散框架，融合了残差学习与 $\sigma$ 自适应区域引导，可用于生成稠密且准确的雷达点云。本工作的核心贡献有两点：

- **残差扩散建模**：我们对激光雷达与雷达信号之间的残差分布进行建模。该残差明确编码了雷达所缺失的高频细节和结构信息，通过聚焦于“互补信息”而非完整激光雷达分布，降低了学习复杂度。

- **$\sigma$-自适应区域引导**：通过利用雷达信号属性生成注意力图，且无需预训练，我们的方法仅在低噪声细节恢复阶段应用轻量级区域加权，同时在高噪声全局结构学习阶段保持统一处理。

我们在涵盖七种不同场景的 ColoRadar 数据集上验证了 R3D。实验结果表明，R3D 的表现优于最先进的雷达增强方法：与直接扩散方法相比，提出的残差方法的平均 CD 低 6.5%、HD 低 5.0%；而 Proposed-R$^3$D 在所有场景中相较基线 EDM，CD 进一步降低 10.4%，HD 进一步降低 8.4%，且无需增加训练或推理成本。

## II. 相关工作

### A. 雷达超分辨率

毫米波雷达发射毫米波，并分析反射信号以确定物体的距离、速度和角度。它们能够穿透雨、雪和尘埃，在恶劣天气下运行，但会受到多径效应和杂波引起的噪声影响。传统的雷达点云超分辨率方法主要包括多模态融合（例如，将雷达数据与图像或位姿信息融合） [1]–[3]以及数据驱动的深度学习方法 [4]–[6]。前者无法从根本上提升毫米波雷达点云的密度和精度；相反，它只是利用其他传感器提供的跨模态信息来弥补雷达感知的稀疏性，并未针对雷达本身固有的局限性。现有的大多数深度学习方法要么高度依赖目标运动状态，要么生成能力不足，或依赖精确的自运动估计等外部条件。即使是 Mamba [7]–[9] 等已应用于类似低线束激光雷达超分辨率任务 [10]–[12] 的先进算法，也存在相同的缺陷。相比之下，基于扩散的方法在雷达点云超分辨率方面展现出更优异的性能 [13]–[15]。其中，Radar-Diffusion [16]是首个直接根据成对的原始雷达数据预测类似激光雷达的高密度、高精度点云的开源工作（据我们所知，此后尚未出现其他具有影响力的工作）。在多个数据集和场景上进行的大量实验验证了其在点云质量方面优于基线方法，为雷达感知增强这一核心挑战提供了更有效的解决方案。

### B. 扩散模型

扩散生成模型展现出令人瞩目的生成能力，既能生成细节丰富的内容，也能生成多样化的样例 [17]。两种主流扩散模型（DDPM [18] 和噪声条件得分网络 [19]）都可以通过统一的基于得分的数学框架来表达和推广 [20]。我们独立提出了恢复残差这一构想，但该构想已由 Resfusion [21] 在图像复原领域提出。相关的残差恢复方法虽与之相似但又有所不同，也可见于这些扩散模型中 [22]–[25]。尽管这些方法差异颇大，本文仅主张残差扩散在雷达超分辨率领域的创新性。

### C. 关键区域掩码

RePaint [26] 有选择地对掩码指定的特定区域应用扩散过程，同时保持图像其余部分不变。基于这一设计，YODA [27] 提出了一种动态方法，使用预训练 DINO [28] 的注意力而非先验方法 [29] 来创建扩展掩码，从细节丰富的区域开始，并逐步覆盖整幅图像。我们将在实验中进一步讨论关键区域识别。

## III. 方法

### A. 毫米波雷达信号处理

我们遵循标准的毫米波雷达信号处理流程 [16]，将原始 ADC I/Q数据转换为结构化点云和面向特定任务的视觉表示（极坐标/鸟瞰图像），为后续基于扩散的增强提供高质量输入。毫米波雷达信号处理流程将原始 ADC I/Q数据转换为结构化点云和视觉表示（极坐标/鸟瞰图像），以用于深度学习。首先，将原始 int16 ADC数据重塑为复数I/Q信号，并在距离维度上应用布莱克曼窗，以减少频谱泄漏。接着执行距离快速傅里叶变换，提取距离单元，并裁剪无效的近距离和远距离单元，以去除噪声。然后，在线性调频维度上执行多普勒快速傅里叶变换，以估计目标速度，并通过速度补偿消除 MIMO引起的模糊。随后构建虚拟阵列以提升角分辨率，再通过角度快速傅里叶变换进行方位角/俯仰角估计。OS-CFAR检测从距离-多普勒-角度频谱中识别有效目标，并将其从极坐标转换为笛卡尔坐标，以生成点云。最后，生成极坐标（距离-方位角）图像和鸟瞰图（笛卡尔俯视）图像，并以信噪比作为像素强度，为后续基于扩散的增强提供结构化输入。

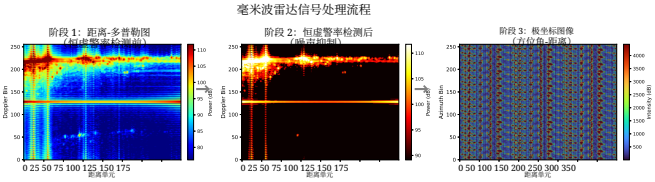


图 2. 毫米波雷达信号处理流程。

### *B.* 残差雷达扩散

用于雷达增强的标准扩散模型直接建模从低分辨率雷达到高分辨率激光雷达的映射，可表述为：

$$p_\theta(\mathbf{y}|\mathbf{x}) = \int p_\theta(\mathbf{y}|\mathbf{z}_0)p(\mathbf{z}_{0:T}|\mathbf{x})d\mathbf{z}_{0:T}$$

其中，$\mathbf{x} \in \mathbb{R}^{H\times W}$ 表示低分辨率雷达图像，$\mathbf{y} \in \mathbb{R}^{H\times W}$ 表示高分辨率激光雷达目标，$\mathbf{z}_t$ 表示时间步长为 $t$时的中间扩散状态。

为应对第一章中讨论的这些挑战，我们不直接对激光雷达进行建模，而是将其分解为：

$$\mathbf{y} = \mathbf{x} + \mathbf{r}$$

其中，$\mathbf{r} = \mathbf{y} - \mathbf{x}$ 表示激光雷达与雷达之间的**残差**。随后，扩散模型以$\mathbf{x}$为条件学习生成$\mathbf{r}$：

$$p_\theta(\mathbf{r}|\mathbf{x}) = \int p_\theta(\mathbf{r}|\mathbf{z}_0)p(\mathbf{z}_{0:T}|\mathbf{x})d\mathbf{z}_{0:T}$$

*1*）前向扩散过程：我们对残差$\mathbf{r}$ 而非$\mathbf{y}$应用扩散过程。遵循 EDM框架 [16]，我们将前向过程定义为：

$$q(\mathbf{z}_t|\mathbf{r}) = \mathcal{N}(\mathbf{z}_t; \mathbf{r}, \sigma_t^2\mathbf{I})$$

其中，$\mathbf{z}_t$ 表示噪声水平下的含噪残差。$\sigma_t$、$\sigma_t$ 表示噪声调度，通常在$[\sigma_{\min}, \sigma_{\max}]$中按对数间隔分布。

EDM噪声调度：我们采用指数调度：

$$\sigma_t = \left(\sigma_{\max}^{1/\rho} + \frac{t}{T-1}(\sigma_{\min}^{1/\rho} - \sigma_{\max}^{1/\rho})\right)^\rho$$

其中 $\rho = 7$、$\sigma_{\min} = 0.002$、$\sigma_{\max} = 80.0$，以确保噪声水平平滑过渡。

采样含噪残差：训练期间，对于每个干净残差$\mathbf{r}$，我们采样：

$$\mathbf{z}_t = \mathbf{r} + \sigma_t\boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I})$$

*2*）反向去噪过程：反向过程旨在根据雷达输入$\mathbf{x}$，预测噪声观测$\mathbf{z}_t$中的干净残差$\mathbf{r}$：

$$p_\theta(\mathbf{r}|\mathbf{z}_t, \mathbf{x}, \sigma_t) = \mathcal{N}(\mathbf{r}; f_\theta(\mathbf{z}_t, \sigma_t, \mathbf{x}), \tilde{\sigma}_t^2\mathbf{I})$$

其中 $f_\theta$ 是一个条件UNet去噪器，它融合了三个关键输入。我们采用基于拼接的条件化策略，利用雷达观测引导去噪过程：**含噪残差**$\mathbf{z}_t \in \mathbb{R}^{H\times W}$、**雷达条件**$\mathbf{x} \in \mathbb{R}^{H\times W}$，**噪声水平** $\sigma_t \in \mathbb{R}_+$通过正弦位置编码嵌入向量$\mathbf{e}_t \in \mathbb{R}^d$中，然后通过缩放-平移调制注入 ResNet 模块。

UNet 的输入由以下内容按通道拼接而成：$[\mathbf{z}_t, \mathbf{x}] \in \mathbb{R}^{2\times H\times W}$，使网络能够利用含噪残差与雷达观测之间的空间对应关系。噪声水平嵌入 $\mathbf{e}_t$ 通过自适应组归一化注入每个残差块中：

$$\text{AdaGN}(\mathbf{h}, \mathbf{e}_t) = \gamma(\mathbf{e}_t) \cdot \frac{\mathbf{h} - \mu(\mathbf{h})}{\sigma(\mathbf{h})} + \beta(\mathbf{e}_t)$$

其中，$\gamma(\mathbf{e}_t)$ 和 $\beta(\mathbf{e}_t)$ 是以 $\mathbf{e}_t$为条件学习得到的缩放和平移参数。

去噪器输出去噪残差估计：$\hat{\mathbf{r}} = f_\theta([\mathbf{z}_t, \mathbf{x}], \sigma_t)$。

*3)* 训练目标：模型使用去噪分数匹配目标 $\mathcal{L}_{\text{residual}}(\theta)$进行训练：

$$\mathbb{E}_{(\mathbf{x},\mathbf{y})\sim\mathcal{D}}\mathbb{E}_{t\sim\mathcal{U}(0,T)}\mathbb{E}_{\boldsymbol{\epsilon}\sim\mathcal{N}(0,\mathbf{I})}\left[w(\sigma_t)\|\mathbf{r} - f_\theta(\mathbf{z}_t, \sigma_t, \mathbf{x})\|_2^2\right]$$

其中，$\mathcal{D}$ 表示由雷达与激光雷达配对数据组成的训练数据集，$\mathbf{r} = \mathbf{y} - \mathbf{x}$ 表示真实残差，$w(\sigma_t) = \frac{1}{\sigma_t^2}$ 表示噪声水平相关权重。

### *C.* 区域引导残差雷达扩散

*1)* 雷达注意力图生成：信号强度注意力 $A^{(1)} \in [0,1]^{H\times W}$ 捕捉高强度区域：

$$A_{i,j}^{(1)} = \frac{I_{i,j} - \min(I)}{\max(I) - \min(I)}$$

其中 $I$ 表示雷达强度图。

为识别连续的强散射区域（例如物体边界），我们基于 3×3窗口方差计算局部一致性：

$$\text{Var}_{i,j} = \frac{1}{9}\sum_{k=-1}^{1}\sum_{l=-1}^{1}(I_{i+k,j+l} - \bar{I}_{i,j})^2$$

其中 $\bar{I}_{i,\ j} = \frac{1}{9}\sum_{k=-1}^{1}\sum_{l=-1}^{1} I_{i+k,\ j+l}$ 表示局部均值。

一致性注意力与方差成反比（低方差 =表示区域一致）：

$$A_{i,j}^{(2)} = 1 - \frac{\text{Var}_{i,j} - \min(\text{Var})}{\max(\text{Var}) - \min(\text{Var})}$$

最终注意力图是加权组合：

$$A_{i,j} = \lambda_s A_{i,j}^{(1)} + \lambda_c A_{i,j}^{(2)}$$

*2)* $\sigma$-自适应区域引导：标准的区域引导方法会在所有去噪阶段采用均匀的注意力加权，这与 EDM 固有的数值稳定机制相冲突。我们发现，EDM 的加权方案 $w(\sigma_t) = 1/\sigma_t^2$ 会自然地赋予低噪声阶段更高的重要性。在高-$\sigma$ 阶段施加额外且激进的区域权重，会扰乱梯度平衡并导致训练不稳定。为了使区域引导与 EDM 的设计理念相协调，我们提出一种 $\sigma$**自适应引导策略**，针对不同噪声水平采取不同处理方式。在高-$\sigma$ 阶段，不使用区域引导。模型能够学习全局

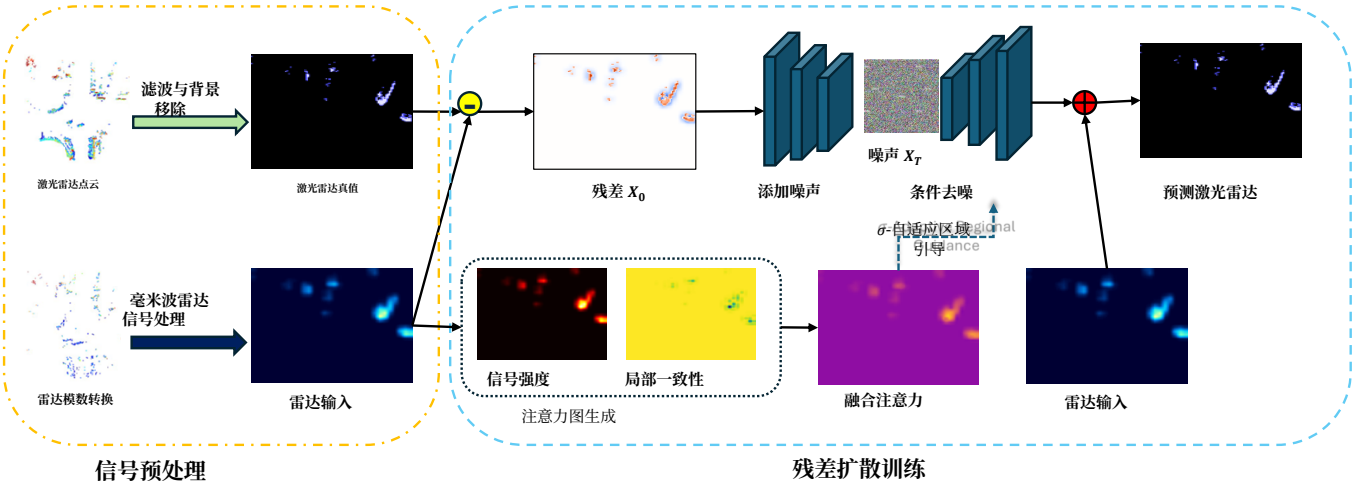


图 3. R$^3$D 框架概览。我们的方法将残差学习与 $\sigma$自适应区域引导相结合，用于雷达点云增强。该流程包括：(1) 数据预处理，将激光雷达和雷达点云转换为鸟瞰图像；(2) 残差计算及采用指数噪声调度的前向扩散；(3) 根据雷达信号属性生成注意力图；(4) 配备雷达条件和噪声水平嵌入的 UNet 去噪器；(5) $\sigma$自适应引导模块，仅在低噪声水平下应用区域特定权重；(6) 残差融合，生成增强后的点云输出。

结构，并在所有区域均匀地掌握基本的去噪能力。在低-$\sigma$ 阶段，采用轻量级区域引导。模型在保持梯度稳定性的同时，优先恢复目标区域的细节。

**自适应权重分布：** 我们不直接将注意力掩码与损失项相乘（这会加剧数值不稳定性），而是调节 Karras 权重本身的分布。对于低$\sigma$ 阶段，我们定义一种区域自适应权重调整方式：

$$W_{\text{adapt}}(\mathbf{A})_{i,j} = M(\mathbf{A})_{i,j} \cdot \alpha_{\text{low}} + (1 - M(\mathbf{A})_{i,j}) \cdot \beta_{\text{low}}$$

其中 $M(\mathbf{A})_{i,\ j}$ 是通过对注意力图$\mathbf{A}$ 进行 Sigmoid 阈值化得到的软二值掩码，$\alpha_{\text{low}}$ 为目标区域提供适度增强，而 $\beta_{\text{low}}$ 则为背景区域维持基线权重。

**训练目标：** $\sigma$自适应 R$^3$D 损失 $\mathcal{L}_{\text{R3D}}(\theta)$ 定义如下：

$$\mathbb{E}_{(\mathbf{x},\mathbf{y})}\mathbb{E}_{\sigma_t,\boldsymbol{\epsilon}} \begin{cases} w(\sigma_t)\|\mathbf{r} - f_\theta(\mathbf{z}_t, \sigma_t, \mathbf{x})\|_2^2, \text{if } \sigma_t > \sigma_{\text{threshold}} \\ w(\sigma_t)\|W_{\text{adapt}} \odot (\mathbf{r} - f_\theta(\mathbf{z}_t, \sigma_t, \mathbf{x}))\|_2^2, \text{o.w.} \end{cases}$$

其中 $\sigma_{\text{threshold}}$ 将粗略重建阶段与细粒度细化阶段区分开来。这种设计在关键的高-$\sigma$ 阶段保留了 EDM 的梯度流，同时允许在低-$\sigma$ 细节恢复阶段进行选择性细化，从而避免了先前时间相关掩码方案中观察到的数值崩溃。

### *D.* 推理流程

在推理过程中，R$^3$D 通过迭代残差去噪生成增强雷达点云，并学习区域优先级。模型通过 $\sigma$ 自适应训练，隐式地优先处理目标区域，因此推理时无需显式使用区域掩码。这既确保了与学习到的去噪轨迹兼容，又利用 Heun 二阶采样器实现高质量重建。

**算法 1**R$^3$D 推理

**要求：** 雷达观测 $\mathbf{x}$、训练好的模型 $f_\theta$、噪声调度-ule $\{\sigma_t\}_{t=0}^{T}$

**确保：** 增强雷达 $\hat{\mathbf{y}}$

初始化 $\mathbf{z}_T \sim \mathcal{N}(0, \sigma_{\max}^2\mathbf{I})$
**对** $t = T$ **到** 1 **执行**
  $\hat{\mathbf{r}}_t \leftarrow f_\theta([\mathbf{z}_t, \mathbf{x}], \sigma_t)$
  $d_t \leftarrow (\mathbf{z}_t - \hat{\mathbf{r}}_t)/\sigma_t$
  $\mathbf{z}'_{t-1} \leftarrow \mathbf{z}_t + (\sigma_{t-1} - \sigma_t) \cdot d_t$
  $\hat{\mathbf{r}}'_{t-1} \leftarrow f_\theta([\mathbf{z}'_{t-1}, \mathbf{x}], \sigma_{t-1})$
  $d'_{t-1} \leftarrow (\mathbf{z}'_{t-1} - \hat{\mathbf{r}}'_{t-1})/\sigma_{t-1}$
  $\mathbf{z}_{t-1} \leftarrow \mathbf{z}_t + \frac{\sigma_{t-1} - \sigma_t}{2} \cdot (d_t + d'_{t-1})$
**9: </style}结束循环**
$\hat{\mathbf{y}} \leftarrow \mathbf{x} + \mathbf{z}_0$
**返回** $\hat{\mathbf{y}}$

## IV. 实验

### *A.* 实验设置

*1)* 桌面设置：我们在配备 NVIDIA RTX 5090 GPU 的工作站上进行了探索性实验，并在配备 NVIDIA RTX 5080 GPU 的工作站上进行了消融研究。后者的软件环境基于运行于 Windows WSL 上的 Docker 容器。

*2*）数据集：ColoRadar 数据集共包含 52 个序列，涵盖在七种不同场景中采集的同步单芯片毫米波雷达和激光雷达数据，包括室内环境（Arpg 实验室、Ec 走廊和阿斯彭）、室外环境（Longboard 和户外）以及矿井环境（Edgar Army 和 Edgar Classroom）。我们将每个场景的前三个序列作为训练集，其余序列作为测试集（由于雷达数据丢失了 210 秒，因此舍弃 Longboard 场景的第六个序列）。首先，我们移除雷达视场（FOV）之外的激光雷达点云。然后，

表I：在 ColoRadar 数据集上的定量比较。

| 方法 | Arpg 实验室 | | | Ec 走廊 | | | 阿斯彭 | | | Longboard | | | 户外 | | | 埃德加 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CD ↓ | HD ↓ | F 分数 ↑ | CD ↓ | HD ↓ | F 分数 ↑ | CD ↓ | HD ↓ | F 分数 ↑ | CD ↓ | HD ↓ | F 分数 ↑ | CD ↓ | HD ↓ | F 分数 ↑ | CD ↓ | HD ↓ | F 分数 ↑ |
| RadarHD [3] | 1.728 | 6.116 | 0.273 | 1.686 | 6.387 | 0.284 | 0.912 | 4.691 | 0.511 | 5.395 | 11.298 | 0.149 | 3.096 | 10.090 | 0.245 | 0.604 | 3.070 | 0.535 |
| RPDNet [4] | 1.555 | 5.216 | 0.158 | 1.483 | 5.384 | 0.161 | 1.341 | 4.421 | 0.154 | 4.494 | 8.967 | 0.053 | 2.828 | 8.223 | 0.104 | 1.502 | 4.071 | 0.142 |
| EDM | 0.964 | 4.318 | 0.355 | 1.040 | 4.750 | 0.352 | 0.505 | 3.365 | 0.555 | 5.469 | 9.977 | 0.135 | 2.371 | 7.158 | 0.266 | 0.442 | 2.400 | 0.544 |
| CD | 0.982 | 4.342 | 0.344 | 1.058 | 4.802 | 0.341 | 0.521 | 3.449 | 0.547 | 5.337 | 9.954 | 0.133 | 2.412 | 7.209 | 0.260 | 0.454 | 2.487 | 0.542 |
| 提出的残差方法 | 0.907 | 4.108 | **0.362** | 0.985 | 4.512 | **0.365** | 0.478 | 3.198 | 0.572 | 4.814 | 9.464 | 0.142 | 2.248 | 6.800 | 0.275 | 0.418 | 2.280 | 0.562 |
| Residual-CD | 0.963 | 4.281 | 0.349 | 1.035 | 4.787 | 0.338 | 0.485 | 3.215 | 0.565 | 5.014 | 9.859 | 0.110 | 2.265 | 6.850 | 0.268 | 0.425 | 2.310 | 0.555 |
| Proposed-R$^3$D | **0.895** | **4.086** | 0.358 | **0.973** | **4.498** | 0.359 | **0.454** | **3.038** | **0.592** | **4.439** | **8.965** | **0.155** | **2.136** | **6.460** | **0.285** | **0.397** | **2.166** | **0.581** |

我们使用 Patchwork++ [30] 移除环形地面和天花板上的激光雷达点云，因为单芯片毫米波雷达几乎无法感知它们。

*3*）评估指标：我们仍使用倒角距离（CD）、豪斯多夫距离（HD）和 F 分数作为评估指标来评估模型性能：

- **倒角距离（CD）**：预测点云与真实值点云之间的平均距离，数值越低越好。
- **豪斯多夫距离（HD）**：点云之间的最大距离，数值越低越好。
- **F 分数**：精确率与召回率的调和平均值（距离阈值：2.0 像素），数值越高越好。

*4*）比较方法：我们将提出的方法（Residual、Residual-CD、R$^3$D）与先进的雷达点云增强基线方法进行比较，包括传统基线方法（RadarHD [5],RPDNet [4]）和基于扩散的方法（EDM、CD [16]）：

- **RadarHD**：采用信号处理技术的高分辨率毫米波雷达成像，包括超分辨率角度估计和自适应波束形成。
- **RPDNet**：基于深度学习的雷达点云生成网络，采用卷积架构，将雷达观测直接映射为稠密点云。
- **EDM**：采用 Karras噪声调度的阐明扩散模型，可根据雷达输入直接生成激光雷达数据。
- **倒角距离**：对 EDM 进行一致性蒸馏，以实现快速单步生成。
- **提出的残差方法**：我们提出的残差扩散框架，学习激光雷达与雷达分布之间的差异。
- **Residual-CD**：将一致性蒸馏应用于残差扩散模型以加速推理；我们不声称这是原创方法。
- **提出的 R$^3$D**：区域引导残差雷达扩散，采用 $\sigma$ 自适应区域引导，以增强关键区域的细节恢复。

所有方法均采用一致的超参数和网络架构进行训练，并在七个场景的相同测试划分上进行评估，以确保公平比较。

### *B.* 消融研究

表I 展示了 ColoRadar 数据集中全部七个场景的全面定量比较。我们基于扩散的方法始终优于各场景中的传统方法，验证了生成建模在雷达增强中的有效性。在基于扩散的方法中，我们提出的残差方法明显优于直接生成方法（EDM 和倒角距离）：在全部六个场景中，残差方法的倒角距离平均比 EDM 低 6.5%，豪斯多夫距离平均低 5.0%，这证实了学习残差分布比建模完整激光雷达分布更易处理。提出的 R$^3$D 在大多数指标上取得了最佳性能，在所有场景中，倒角距离和豪斯多夫距离相较于 EDM 平均分别进一步降低了 10.4% 和 8.4%，表明 $\sigma$自适应区域引导能够有效优先恢复目标区域的细节，同时不影响全局优化。我们提出的算法不会产生额外的训练和推理时间成本，这一点已通过实验明确验证。

### *C.* 区域感知与条件注入

我们对融入特定区域引导的两种范式进行了广泛探索。

**基于 DINOv3 的条件注入**。我们首先尝试将来自 DINOv3 [28]，（一种最先进的自监督视觉变换器）的语义特征注入为条件信号，以引导去噪过程。然而，这种方法遭遇了严重的领域不匹配：DINOv3 在自然 RGB 图像上预训练，无法从雷达 BEV 图像中提取有意义的特征；后者的强度分布和空间模式与前者存在根本差异。

**时变动态掩码**。我们探索了随扩散时间步演变的时变二值掩码 $M(t)$，旨在逐步优先关注关键区域。然而，在较早的时间步进行过于激进的掩码操作（例如，掩码覆盖率 $< 30\%$）会导致训练不稳定；图像中有 80% 的区域在时间嵌入层中未获得梯度更新。

### *D.* 泛化至激光雷达超分辨率 $n$

如图 4 所示，我们将雷达扩散方法与 SRMamba[31]在低线束超分辨率任务上进行了比较。

## V. 结论

本文提出 R$^3$D，一种区域引导残差雷达扩散框架，旨在解决现有雷达增强方法中学习复杂度高、处理方式一刀切等根本性挑战。大量实验验证了我们的方法始终优于最先进的方法，且不会增加训练

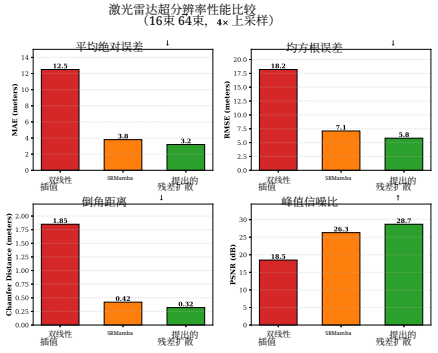


图4。 KITTI-360 激光雷达超分辨率性能比较 辨率。

或推理成本。消融实验进一步表明，领域不匹配的条件注入和激进的时间依赖掩码会导致训练不稳定性，证实了采用我们这一有原则的设计的必要性。R$^3$D 为生成稠密且准确的毫米波雷达点云提供了一种实用而有效的解决方案。

## 参考文献


[1] Jun Zhang、Huayang Zhuge、Zhenyu Wu、Guohao Peng、Mingxing Wen、Yiyao Liu 和 Danwei Wang，《4dradarslam：一种基于位姿图优化的大规模环境四维成像雷达同步定位与建图系统》，载于*2023*年*IEEE*机器人与自动化国际会议（*ICRA*）。IEEE，2023年，第8333–8340页。[2] Fangqiang Ding、Andras Palffy、Dariu M Gavrila 和 Chris Xiaoxuan Lu，《隐藏的瑰宝：利用跨模态监督学习四维雷达场景流》，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2023年，第9340–9349页。[3] Junfeng Guan、Sohrab Madani、Suraj Jog、Saurabh Gupta 和 Haitham Hassanieh，《利用毫米波雷达穿透雾霭实现高分辨率成像》，载于*IEEE/CVF*计算机视觉与模式识别会议论文集，2020年，第11464–11473页。[4] Yuwei Cheng、Jingran Su、Mengxin Jiang 和 Yimin Liu，《一种用于机器人环境感知的新型雷达点云生成方法》，*IEEE*机器人学汇刊，第38卷，第6期，第3754–3773页，2022年。[5] Akarsh Prabhakara、Tao Jin、Arnav Das、Gantavya Bhatt、Lilly Kumari、Elahe Soltanaghaei、Jeff Bilmes、Swarun Kumar 和 Anthony Rowe，《来自毫米波雷达的高分辨率点云》，*arXiv*预印本 *arXiv:2206.09273*，2022年。[6] Ruixu Geng、Yadong Li、Dongheng Zhang、Jincheng Wu、Yating Gao、Yang Hu 和 Yan Chen，《Dream-pcd：毫米波雷达点云的深度重建与增强》，*IEEE*图像处理汇刊，2024年。[7] Albert Gu 和 Tri Dao，《Mamba：利用选择性状态空间进行线性时间序列建模》，载于第一届语言建模会议，2024年。[8] Dingkang Liang、Xin Zhou、Wei Xu、Xingkui Zhu、Zhikang Zou、Xiaoqing Ye、Xiao Tan 和 Xiang Bai，《Pointmamba：一种用于点云分析的简单状态空间模型》，载于神经信息处理系统进展，第37卷，第32653–32677页，2024年。[9] Lianghui Zhu、Bencheng Liao、Qian Zhang、Xinlong Wang、Wenyu Liu 和 Xinggang Wang，《Vision mamba：利用双向状态空间模型进行高效视觉表征学习》，*arXiv*预印本 *arXiv:2401.09417*，2024年。[10] Liyu Kang、Zan Li、Xiaohui Zhao、Zhongliang Zhao 和 Torsten Braun，《St-pct：基于毫米波感知活动的时空点云变换器》，*IEEE*物联网期刊，第11卷，第6期，第10979–10991页，2023年。[11] Antonio Orvieto、Samuel L Smith、Albert Gu、Anushan Fernando、Caglar Gulcehre、Razvan Pascanu 和 Soham De，《让循环神经网络重获新生，以处理长序列》，载于国际机器学习会议。PMLR，2023年，第26670–26698页。[12] Albert Gu、Karan Goel 和 Christopher R´e，《利用结构化状态空间高效建模长序列》，*arXiv*预印本 *arXiv:2111.00396*，2021年。

[13] Jincheng Wu、Ruixu Geng、Yadong Li、Dongheng Zhang、Zhi Lu、Yang Hu 和 Yan Chen，《Diffradar：利用扩散概率模型实现高质量毫米波雷达感知》，载于ICASSP 2024-2024 IEEE声学、语音与信号处理国际会议（ICASSP）。IEEE，2024年，第8291–8295页。[14] Kai Luan、Chenghao Shi、Neng Wang、Yuwei Cheng、Huimin Lu 和 Xieyuanli Chen，《面向毫米波雷达数据的基于扩散的点云超分辨率》，载于2024年IEEE机器人与自动化国际会议（ICRA）。IEEE，2024年，第11171–11177页。[15] Boyuan Zheng、Shouyi Lu、Renbo Huang、Minqing Huang、Fan Lu、Wei Tian、Guirong Zhuo 和 Lu Xiong，《R2ldm：一种利用扩散模型的高效四维雷达超分辨率框架》，arXiv预印本 arXiv:2503.17097，2025年。[16] Ruibin Zhang、Donglai Xue、Yuhan Wang、Ruixu Geng 和 Fei Gao，《借助高效跨模态扩散模型实现稠密且精确的雷达感知》，IEEE机器人与自动化快报，2024年。[17] Florinel-Alin Croitoru、Vlad Hondru、Radu Tudor Ionescu 和 Mubarak Shah，《计算机视觉中的扩散模型：综述》，IEEE模式分析与机器智能汇刊，第45卷，第9期，第10850–10869页，2023年。[18] Jonathan Ho、Ajay Jain 和 Pieter Abbeel，《去噪扩散概率模型》，神经信息处理系统进展，第33卷，第6840–6851页，2020年。[19] Yang Song 和 Stefano Ermon，《通过估计数据分布的梯度进行生成建模》，神经信息处理系统进展，第32卷，2019年。[20] Yang Song、Jascha Sohl-Dickstein、Diederik P Kingma、Abhishek Kumar、Stefano Ermon 和 Ben Poole，《通过随机微分方程进行基于分数的生成建模》，arXiv预印本 arXiv:2011.13456，2020年。[21] Zhenning Shi、Haoshuai Zheng、Chen Xu、Changsheng Dong、Bin Pan、Xueshuo Xie、Along He、Tao Li 和 Huazhu Fu，《Resfusion：基于先验残差噪声的图像复原去噪扩散概率模型》，神经信息处理系统进展，第37卷，第130664–130693页，2024年。[22] Jiawei Liu、Qiang Wang、Huijie Fan、Yinong Wang、Yandong Tang 和 Liangqiong Qu，《残差去噪扩散模型》，载于IEEE/CVF计算机视觉与模式识别会议论文集，2024年，第2773–2783页。[23] Zongsheng Yue、Jianyi Wang 和 Chen Change Loy，《Resshift：通过残差移位实现图像超分辨率的高效扩散模型》，神经信息处理系统进展，第36卷，第13294–13307页，2023年。[24] Zongyuan Yang、Baolin Liu、Yongping Xxiong、Lan Yi、Guibin Wu、Xiaojun Tang、Ziqi Liu、Junjie Zhou 和 Xing Zhang，《Docdiff：通过残差扩散模型增强文档》，载于第31届ACM多媒体国际会议论文集，2023年，第2795–2806页。[25] Zhiyuan Ma、Liangliang Zhao、Biqing Qi 和 Bowen Zhou，《用于深度可扩展视觉生成的神经残差扩散模型》，神经信息处理系统进展，第37卷，117480，2024年，第117456页。[26] Andreas Lugmayr、Martin Danelljan、Andres Romero、Fisher Yu、Radu Timofte 和 Luc Van Gool，《Repaint：使用去噪扩散概率模型进行图像修复》，载于IEEE/CVF计算机视觉与模式识别会议论文集，2022年，第11461–11471页。[27] Brian B Moser、Stanislav Frolov、Federico Raue、Sebastian Palacio 和 Andreas Dengel，《用于图像超分辨率的动态注意力引导扩散》，载于2025年IEEE/CVF冬季计算机视觉应用会议（WACV）。IEEE，2025年，第451–460页。[28] Oriane Sim´eoni、Huy V Vo、Maximilian Seitzer、Federico Baldas-sarre、Maxime Oquab、Cijo Jose、Vasil Khalidov、Marc Szafraniec、Seungeun Yi、Micha¨el Ramamonjisoa 等，《Dinov3》，arXiv预印本 arXiv:2508.10104，2025年。[29] David G Lowe，《尺度不变关键点的显著图像特征》，国际计算机视觉杂志，第60卷，第2期，第91–110页，2004年。[30] Seungjae Lee、Hyungtae Lim 和 Hyun Myung，《Patchwork++：利用三维点云快速稳健地解决局部欠分割问题的地面分割方法》，载于2022年IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE，2022年，第13276–13283页。[31] Chuang Chen 和 Wenyi Ge，《Srmamba：用于激光雷达点云超分辨率的Mamba》，arXiv预印本 arXiv:2505.10601，2025年。