FOTV-HQS: A Fractional-Order Total Variation Model for LiDAR Super-Resolution with Deep Unfolding Network

Huiying Xi[1⋆], Xia Yuan[1⋆⋆], Shiwei Wu[1], Runze Geng[1], Kaiyang Wang[1], Yongshun Liang[1], and Chunxia Zhao[1]

Nanjing University of Science and Technology, Nanjing, China
yuanxia@njust.edu.cn

Abstract. LiDAR super-resolution can improve the quality of point cloud data, which is critical for improving many downstream tasks such as object detection, identification, and tracking. Traditional LiDAR super-resolution models often struggle with issues like block artifacts, staircase edges, and misleading edges. To address these challenges, a novel super-resolution model of LiDAR based on fractional-order total variation (FOTV) is proposed in this paper. We propose a FOTV regularization optimization problem, utilizing an end-to-end trainable iterative network to capture data attributes.This enables the precise reconstruction of fine details and complex structures in point clouds. Specifically, the half-quadratic splitting algorithm divides the problem into data fidelity and prior regularization subproblems. We then propose a deep unfolding network, which iteratively deals with the two subproblems within the FOTV-HQS framework. Numerous experiments have shown that our approach significantly reduces the number of parameters by up to 99.68% and maintains good performance, making it ideal for applications with limited compute and storage resources.

Keywords: Super-resolution · LiDAR · Fractional-order total variation · Deep unfolding

1 Introduction

Light Detection and Ranging (LiDAR) is a cornerstone technology for numerous applications in autonomous systems, including robot navigation[1], autonomous driving [2], and 3D reconstruction [3]. It provides precise 3D representations of the environment by emitting pulsed laser light, capturing fine details crucial for object detection, tracking, and terrain modeling. Unlike traditional imaging sensors that rely on ambient light, LiDAR functions effectively in varied lighting conditions, including nighttime, giving it a significant advantage in applications requiring robust and reliable environmental perception [4].

⋆ Co-first author
⋆⋆ Co-first author, Corresponding author

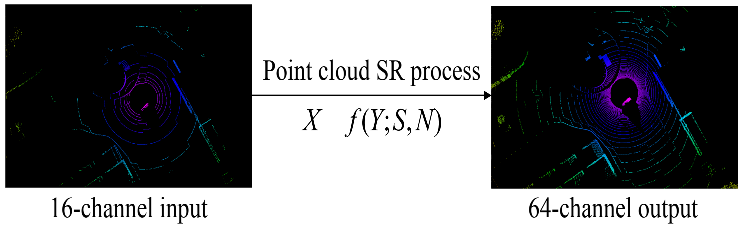


Fig. 1: The input point cloud from a 16-channel LiDAR sensor is reconstructed into a 64-channel point cloud using super-resolution technology, without the need to directly acquire a higher-channel LiDAR sensor.

As application scenarios become increasingly complex, the technological demands on LiDAR performance also escalate. One significant demand is to increase the density of the sensed point clouds, which depends on the number of channels in the sensor. LiDARs with more channels produce denser point clouds, capturing more environmental details and greatly benefiting tasks such as object detection, recognition, tracking, and motion planning. However, pursuing higher resolution in LiDAR technology faces significant challenges. Low-channel LiDAR systems, which are more commonly used, cannot capture high-resolution data directly. These systems often struggle to provide the detailed and dense point clouds necessary for advanced applications in robotics and autonomous navigation.

The advent of super-resolution (SR) techniques offers a promising way to overcome these limitations. By augmenting the resolution of LiDAR data beyond the native capabilities of the sensor (Figure 1), these methods aim to unlock finer environmental details without necessitating the direct acquisition of higher-channel sensors. Yet, traditional approaches to LiDAR super-resolution, including interpolation and conventional learning-based methods, often face challenges such as noise amplification and loss of detail, highlighting the need for more advanced solutions.

In this paper, we propose a novel fractional-order total variation (FOTV) model for LiDAR super-resolution. The energy function of the proposed model consists of two terms: the data fidelity term and the FOTV regularization term. The data fidelity term ensures the similarity between the reconstructed range image and the ground truth range image, while the FOTV regularization term serves to better reconstruct the tiny features and complex structures in the point cloud. We first solve the proposed model by means of the half-quadratic splitting (HQS) algorithm, which divides the model into two subproblems. Additionally, we propose a deep unfolding super-resolution network that iteratively processes the LiDAR super-resolution model based on the FOTV-HQS framework. This network alternates between solving the two sub-problems, one related to the data fidelity term, and the other to the FOTV regularization prior term. We then solve it through the deep unfolding network by replacing the iterative formulas of the two subproblems with neural network modules.

The main contributions of this work are as follows:

1) A novel LiDAR super-resolution model based on fractional-order total variation has been proposed. This model leverages the characteristics of fractional-order to better reconstruct tiny features and complex structures in the point cloud, avoiding issues such as block artifacts, staircase edges and false edge near the edges that are common with traditional regularized super-resolution models.

2) We design a deep unfolding super-resolution network based on FOTV-HQS framework to iteratively process the LiDAR super-resolution model. This network integrates the flexibility of reconstruction-based methods with the advantages of learning-based methods, providing an avenue to bridge the gap between reconstruction-based and learning-based methods.

3) Numerous experiments was conducted on datasets of point cloud range images. Compared to other state-of-the-art deep learning networks, the proposed deep unfolding network based on the FOTV-HQS experiments demonstrated a 99.68% reduction in the number of parameters while also performing well on other quantitative and qualitative metrics.

2 Related Work

2.1 LiDAR super-resolution

LiDAR super-resolution is a crucial task we are working on. The objective is to enhance the resolution of LiDAR data, particularly by increasing the density of point clouds in the vertical direction, allowing for a more detailed representation of the scanned environment. In this realm, Shan *et al*. [5] propose a LiDAR super-resolution methodology that transforms 3D point clouds into 2D image space for enhancement via a deep convolutional neural network. Gkillas *et al*. [6] explore LiDAR super-resolution from a federated learning perspective, enhancing model robustness and diversity by utilizing private data from autonomous vehicles in varied conditions. Kwon *et al*. [7] propose the Implicit LiDAR Network (ILN) for LiDAR super-resolution, leveraging non-linear weights for pixel interpolation and integrating attention mechanisms from Transformer architecture. Eskandar *et al*. [8] propose a novel Height-Aware Lidar Super-resolution model (HALS), which utilizes a height-aware distribution with a multi-branch generator architecture for LiDAR super-resolution.

2.2 Fractional-order total variation regularization for SR

Another problem related to our work is super-resolution based on fractional-order total variation regularization. Image super-resolution utilizes computer signal processing and algorithms to reconstruct high-resolution (HR) images from low-resolution (LR) frames. Various methods have been developed to enhance super-resolution performance, categorized into interpolation-based [9,10], reconstruction-based [11,12,13], and learning-based [14,15,16,17,18] approaches. Super-resolution, a classic ill-posed inverse problem in low-level computer vision,

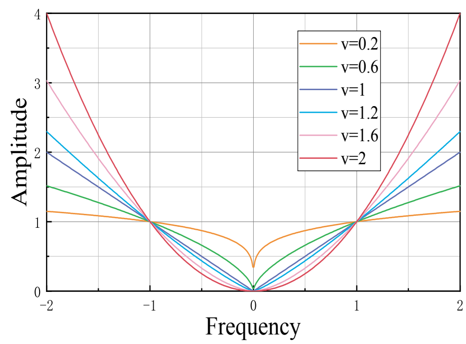


Fig. 2: Differential Amplitude-Frequency Characteristic Curve.

traditionally employs regularization techniques like Tikhonov and total variation (TV) regularization to overcome this ill-posedness. However, these models often lose high-frequency and edge information and exhibit staircasing effects. As shown in Figure 2, fractional differential enhance high-frequency information in signals while preserving mid and low-frequency information better than integer-order differentiation. This is especially beneficial for images with rich textures due to their high self-similarity. Recently, fractional differentiation theory has been widely applied in image processing, yielding significant research achievements. Applying fractional differentiation to super-resolution techniques has become a new trend. Ren *et al.* [19] propose a fractional-order total variation (FOTV) regularization model for super-resolution, effectively handling image texture details while maintaining edge and structure information. Laghrib *et al.* [20] propose a super-resolution method incorporating a nonconvex data fitting term and a FOTV regularization term, effectively reducing complex noises like impulse noise and better preserving image features. Yang *et al.* [21] propose a hybrid single-image super-resolution model that integrates TV and FOTV, utilizing the alternating direction multiplier method for adaptive reconstruction with textural features. Yao *et al.* [22] achieve efficient image reconstruction through a scalar auxiliary variable approach with adaptive time stepping in a hybrid super-resolution framework of TV and FOTV.

3 Proposed Method

LiDAR super-resolution enhances point cloud data quality, essential for tasks such as object detection, identification, and tracking. However, traditional models often face challenges like block artifacts, staircase edges, and misleading edges. To address these issues, we propose a novel fractional-order total variation model for LiDAR super-resolution, as shown in Figure 3, leveraging a dual-component energy function to enhance the reconstruction of vertical resolution in point clouds. Utilizing the HQS algorithm and a deep unfolding network designed around the FOTV-HQS framework, our model iteratively tackles two sub-problems, offering

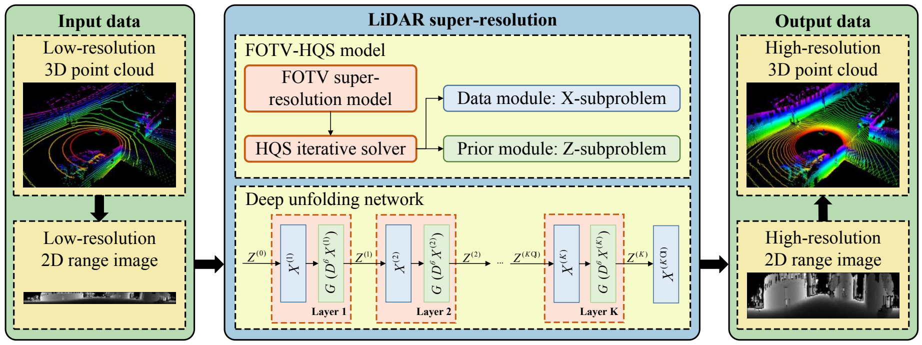


Fig. 3: Overall structure of the proposed method. This method is structured into two distinct modules: LiDAR super-resolution model based on FOTV-HQS and a deep unfolding LiDAR super-resolution network.

a sophisticated solution to traditional super-resolution limitations. This method is structured into two distinct modules: LiDAR super-resolution model based on FOTV-HQS and a deep unfolding LiDAR super-resolution network.

3.1 FOTV LiDAR super-resolution model

Mathematically, super-resolution is an ill-posed inverse problem. We consider a high-resolution point cloud derived from an H-channel LiDAR sensor. To simplify data processing and reduce the computing resource requirements, the high-resolution point cloud is projected into a range image $X \in \mathbb{R}^{H \times W}$, where H denotes the vertical resolution and W denotes the horizontal resolution. The corresponding low-resolution range image $Y \in \mathbb{R}^{\frac{H}{s} \times W}$, where s is the downsampling scale, can be obtained by the following degenerate model:

$$Y = SX + N, \tag{1}$$

where $S \in \mathbb{R}^{\frac{H}{s} \times H}$ denotes the downsampler operator and N is a zero-mean Gaussian noise term. A classical way to overcome this ill-posedness is to add a regularization term to the energy function. Within a regularization-based framework, this ill-posed problem is typically formulated as the following optimization problem [19]:

$$\underset{X}{\operatorname{argmin}} \ \frac{1}{2}\|Y - SX\|_F^2 + \lambda_{tv} TV^{\alpha}(X), \tag{2}$$

where $\frac{1}{2}\|Y - SX\|_F^2$ serves as the data fidelity term, ensuring that the reconstructed point cloud range image corresponds mathematically to the observed image. The notation $\|\cdot\|_F$ denotes the Frobenius norm. $TV^{\alpha}(X)$ is the FOTV regularization term, which acts as a prior to capture and preserve key structural features of range images. It balances edge preservation and noise suppression to optimize the final super-resolution effect. The parameter λ_{tv} is the regularization parameter used to control the balance between these two terms.

For implementation, we consider a discretized version of Equation 2. Consequently, the FOTV LiDAR super-resolution model is articulated as the following optimization problem:

$$\underset{X}{\operatorname{argmin}}\ \frac{1}{2}\|Y - SX\|_F^2 + \lambda_{tv}\|D^\alpha X\|_1, \tag{3}$$

where $D^\alpha X$ denotes the α-order discrete fractional-order gradient of X. In practical numerical computations, the G-L fractional-order gradient is calculated through differential approximation, which mainly relies on the discrete sampling data of the signal. This makes it particularly suitable for processing digital signals or point cloud data. Therefore, we employ the G-L fractional-order gradient for addressing the LiDAR super-resolution problem. The discrete G-L fractional-order gradient ${}^{GL}_{a}D^\alpha_x f(x)$ is defined as

$${}^{GL}_{a}D^\alpha_x f(x) = \lim_{h\to 0}\frac{1}{h^\alpha}\sum_{k=0}^{\left\lfloor\frac{x-a}{h}\right\rfloor}\frac{(-1)^k\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha-k+1)}f(x-kh), \tag{4}$$

where the Gamma function $\Gamma(\tau) = \int_0^\infty e^{-x}x^{\tau-1}\,dx = (\tau-1)!$, and h denotes the step size.

3.2 LiDAR super-resolution based on FOTV-HQS

To solve the cost function proposed in Equation 3, we employ the half-quadratic splitting (HQS) algorithm [23]. This algorithm is instrumental in decomposing the optimization problem into more tractable subproblems, facilitating an iterative approach to finding a solution. The HQS method is particularly advantageous due to its ability to efficiently handle such separable convex programming problems, thereby improving convergence rates and solution accuracy. First, we introduce an auxiliary variable $Z \in \mathbb{R}^{H\times W}$, the optimization problem proposed in Equation 3 is rewritten in the following form:

$$\underset{X,Z}{\operatorname{argmin}}\ \frac{1}{2}\|Y - SX\|_F^2 + \lambda_{tv}\|Z\|_1 \quad \text{s.t. } Z = D^\alpha X. \tag{5}$$

Subsequently, by employing Augmented Lagrange Method (ALM), the optimization problem is transformed into minimizing the following loss function:

$$\mathcal{L}_\beta(X, Z) = \frac{1}{2}\|Y - SX\|_F^2 + \lambda_{tv}\|Z\|_1 + \frac{\beta}{2}\|Z - D^\alpha X\|_F^2, \tag{6}$$

where $\frac{1}{2}\|Y - SX\|_F^2$ denotes the original data fidelity term, $\lambda_{tv}\|z\|_1$ represents the Lagrange multiplier term, and $\frac{\beta}{2}\|Z - D^\alpha X\|_F^2$ is the quadratic penalty term. Here, β is a penalty parameter.

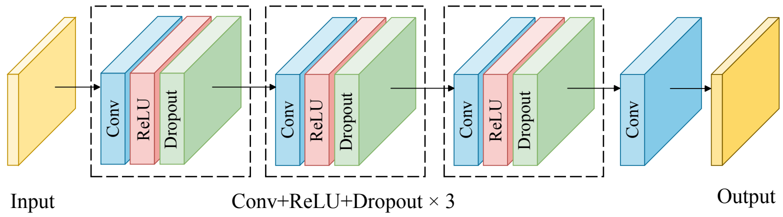


Fig. 4: The architecture of the deep convolutional neural network-based denoiser $G_\theta(\cdot)$ consists of four convolutional layers using spectral normalization. The first three layers are each followed by a ReLU activation function and a Dropout operation. The final layer outputs a denoised image with the same spatial dimensions as the input.

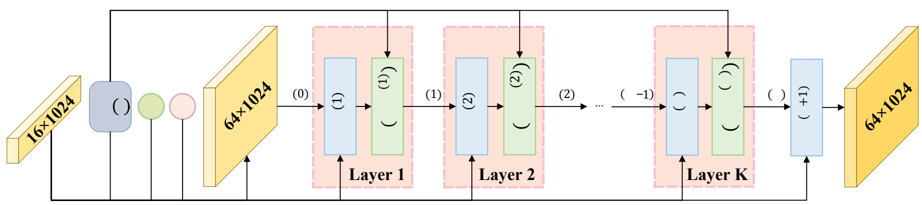


Fig. 5: The overall architecture of the deep unfolding LiDAR super-resolution network with K layers.

By utilizing the HQS algorithm, we decompose Equation 6 into two subproblems, which are addressed by iteratively updating the variables. In the following, k denotes the current iteration step.

Data subproblem: Update $X^{(k+1)}$ by minimizing:

$$\begin{aligned} X^{(k+1)} &= \underset{X}{\operatorname{argmin}}\ \mathcal{L}(X, Z^{(k)}) \\ &= \underset{X}{\operatorname{argmin}}\ \frac{1}{2}\|Y - SX\|_F^2 + \frac{\beta^{(k+1)}}{2}\|Z^{(k)} - D^\alpha X\|_F^2. \end{aligned} \tag{7}$$

According to Equation 7, the penalty parameter β is automatically updated during the iterations according to the progress of the algorithm to minimize the loss function. The penalty parameter in the k-th iteration is denoted by $\beta^{(k)}$.For convenience, it is abbreviated as β below. To solve the X-subproblem in Equation 7, we consider the Euler-Lagrange equation of Equation 7 with respect to X, which has the form

$$-S^T(Y - SX) - \beta D^{\alpha T}(Z^{(k)} - D^\alpha X) = 0, \tag{8}$$

and this leads to the closed form solution for $X^{(k+1)}$ as

$$X^{(k+1)} = (S^T S + \beta D^{\alpha T} D^\alpha)^{-1}(S^T Y + \beta D^{\alpha T} Z^{(k)}). \tag{9}$$

Table 1: Super-resolution based on FOTV-HQS

Algorithm Super-resolution based on FOTV-HQS

Require: Low-resolution range image Y

Ensure: High-resolution range image X

1: **Initialization:** $\beta_0 = 0.009$, $\alpha_0 = 1.6$, $Z^{(0)} = \text{Upsample}(Y)$
2: Set the number of unfolding iterations $K = 6$
3: **for** i in range(K) **do**
4: Update X based on Equation 9
5: Update Z based on Equation 11
6: **end for**
7: Update $X^{(K+1)}$ based on Equation 9
8: **return** $X^{(K+1)}$

Prior subproblem: Update $Z^{(k+1)}$ by minimizing:

$$\begin{aligned} Z^{(k+1)} &= \underset{Z}{\operatorname{argmin}}\ \mathcal{L}(X^{(k+1)}, Z) \\ &= \underset{Z}{\operatorname{argmin}}\ \lambda_{tv}\|Z\|_1 + \frac{\beta^{(k+1)}}{2}\|Z - D^{\alpha}X^{(k+1)}\|_F^2. \end{aligned} \tag{10}$$

From a Bayesian perspective, Equation 10 is essentially a denoising task. Building on this and drawing inspiration from the denoising convolutional neural network (DnCNN) [24], we design a deep convolutional neural network-based denoiser, denoted as $G_\theta(\cdot)$. The network structure is shown in Figure 4. Thus, Equation 10 is rewritten as

$$Z^{(k+1)} = G_\theta(D^{\alpha}X^{(k+1)}). \tag{11}$$

The network $G_\theta(\cdot)$ can be effectively pre-trained using pairs of synthetically generated noisy high-resolution images, contaminated with Gaussian noise γ, and their corresponding ground truth images. These pairs are represented as $\{X^j+\gamma, X^j\}_{j=1}^{p}$. The pre-training process leverages a loss function that minimizes the Frobenius norm of the discrepancy between the neural network's output on the noisy images and the ground truth images, expressed as $\sum_{j=1}^{p}\|G_\theta(X^j + \gamma;\theta) - X^j\|_F^2$.

In summary, the FOTV-HQS-based super-resolution model consists of the following two update rules:

$$\begin{cases} X^{(k+1)} = (S^TS + \beta D^{\alpha T}D^{\alpha})^{-1}(S^TY + \beta D^{\alpha T}Z^{(k)}) \\ Z^{(k+1)} = G_\theta(D^{\alpha}X^{(k+1)}) \end{cases}. \tag{12}$$

The process of the super-resolution algorithm based on FOTV-HQS is summarized in Table 1.

3.3 Deep unfolding LiDAR super-resolution network

Inspired by the deep unfolding network for image super-resolution [16], we propose a deep unfolding LiDAR super-resolution network designed to iteratively address the LiDAR super-resolution model based on the FOTV-HQS framework. The network iteratively executes the update rules from Equation 12, with the k-th step of iteration corresponding to the k-th layer of the proposed deep unfolding framework. The overall architecture of the deep unfolding LiDAR super-resolution network with K layers is depicted in Figure 5.

Regarding the loss function, we adopt the L1 loss, as shown in Equation 13, to optimize the model's trainable parameters—namely, the neural network $G_\theta(\cdot)$, the penalty parameter β and the order of the fractional-order gradient α. The choice of L1 loss is motivated by its robustness to outliers and its effectiveness in preserving structural details in range images, providing more accurate reconstruction results in practical applications.

$$l(\theta) = \sum_{i=1}^{p} \|Z_i^{(K)} - W_i\|_F^2, \tag{13}$$

Here, $Z_i^{(K)}$ represents the output from the deep unfolding LiDAR super-resolution network when processing a low-resolution input range image Y_i, while W_i corresponds to the i-th high-resolution ground truth range image.

This end-to-end training approach optimizes super-resolution processing at each iteration step and enhances model efficiency and performance by refining trainable parameters throughout the process. By combining a meticulously designed network architecture with a carefully selected loss function, the model effectively addresses the complexities of LiDAR data processing.

4 Experimental Results and Analyses

To demonstrate the effectiveness of the FOTV-HQS framework, we conducted extensive experiments on a LiDAR dataset. We performed $4\times$ up-sampling from 16 to 64-channel LiDAR, $8\times$ up-sampling from 16 to 128-channel LiDAR, and $16\times$ up-sampling from 16 to 256-channel LiDAR. The capabilities of the FOTV-HQS LiDAR super-resolution architecture were rigorously assessed through both quantitative and qualitative analyses.

4.1 Datasets

We conduct experiments on the Ouster dataset [5], which simulates a 64-channel LiDAR OS1-64 within the CARLA Town 2 scenario, with a vertical field of view (FOV) of 33.2° and a horizontal FOV of 360°. Corresponding low-resolution point clouds were generated using a simulated 16-channel LiDAR OS1-16 in the same scenario. Projecting these point clouds onto range images yielded 7000 pairs of high and low-resolution range images with resolutions of 64×1024 and 16×1024, respectively, to serve as the training set. For the test set, real-world Ouster LiDAR data from 8825 scans provided high-resolution point clouds projected to range images of 64×1024 resolution. By extracting 16 rows from these high-resolution range images, low-resolution range images of 16×1024 were generated, resulting in 8825 pairs of test high and low-resolution range image.

4.2 Experimental settings

Neural Network Architecture: The deep convolutional neural network $G_\theta(\cdot)$, based on a denoiser, incorporates four convolutional layers with spectral normalization. Each layer has 64 filters of size 3×3. The first three layers employ the ReLU activation function and feature a Dropout operation with a dropout rate of 0.05. The final layer outputs a denoised image that retains the same spatial dimensions as the input.

Proposed Model-Parameter Setting: For the FOTV-HQS framework, we set the number of unfolding iterations, K, to 6, constructing a six-layer deep unfolding network architecture. In the end-to-end training process, we employed the Adam optimizer. Specifically, in the experiments with $4\times$ up-sampling (from 16 to 64-channels LiDAR), $8\times$ up-sampling (from 16 to 128-channel LiDAR) and $16\times$ upsampling experiments (from 16 to 256-channel LiDAR), we started with a learning rate of 1e-03 for 50 epochs, then adjusted it to 1e-04 for an additional 50 epochs.

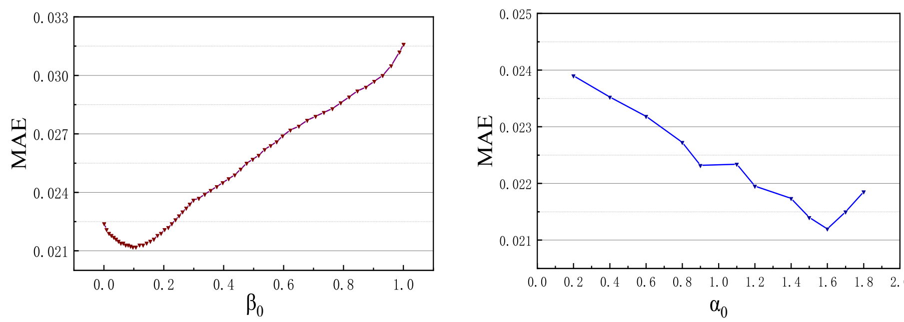
Compared Methods: To thoroughly evaluate our proposed method, we selected six distinct super-resolution techniques designed to reconstruct high-resolution 3D point clouds from sparse LiDAR data. These methods include: (i) interpolation-based techniques: bilinear [9] and bicubic [10] interpolation; (ii) deep learning-based approaches: the avant-garde SR-ResNet [15] model renowned in classical image super-resolution, the USRNet [16] model leveraging deep unfolding networks, the leading-edge LiDAR-SR [5] model designed for LiDAR super-resolution, and the FL-SR [6] model, which utilizes federated learning for LiDAR super-resolution. For these deep learning methods, we used the Adam optimizer, starting with a learning rate of 1e-03 for 50 epochs, then adjusting it to 1e-04 for another 50 epochs.

Evaluation Metrics: To quantitatively assess our method's performance, we employ five quality metrics: peak signal-to-noise ratio (PSNR), structural similarity (SSIM) [25], edge preservation index (EPI) [26], mean squared error (MSE) [27] and mean absolute error (MAE) [28]. Higher values of PSNR, SSIM, and EPI, and lower values of MSE and MAE, indicate superior super-resolution performance.

4.3 Sensitivity tests for α and β

Given that our proposed model incorporates two principal parameters—α, the order of the fractional-order gradient, and β, the penalty parameter, which act as coupling parameters in the regularized inverse problem— it is crucial to examine their sensitivity to super-resolution performance. During the training process, our model dynamically adjusts and learns these parameters to optimize performance. To further understand their impact, we conducted sensitivity tests focusing on their initial value in a $4\times$ up-sampling experiment which increases LiDAR channels from 16 to 64.

We varied β_0 (the initial value of β) within the range of 0.0001 to 1, while keeping α_0 (the initial value of α) fixed at 1.6. The experiment showed that the



(a) MAE Variation with β_0 (b) MAE Variation with α_0

Fig. 6: Sensitivity tests of our proposed model to parameters β_0 (with the fixed $\alpha_0 = 1.6$) and α_0 (with the fixed $\beta_0 = 0.009$).

MAE reached its minimum when β_0 was set to 0.1, as depicted in Figure 6(a). This suggests that the model's super-resolution performance is highly sensitive to this particular setting of β_0. The dynamic adjustment during training further fine-tunes this parameter to achieve optimal results.

We then held β_0 constant at 0.009 and varied α_0 within the interval (0, 2) to assess its impact on the up-sampling performance of point clouds. Figure 6(b) illustrates that the MAE reached its lowest value when α_0 was set to 1.6. This observation underscores the critical role of α_0 in enhancing the resolution quality of the control point cloud. The model's learning process fine-tunes α during training, but selecting an optimal α_0 is essential for maximizing the efficacy of the super-resolution process.

Overall, these sensitivity analyses highlight that while the model dynamically learns the optimal α and β through training, the initial values (α_0 and β_0) play a significant role in the performance of the model. Therefore, careful selection of these initial values is crucial for achieving superior super-resolution performance.

4.4 LiDAR super-resolution performance

In this segment, we compare our proposed model with several other methods, including bilinear [9] and bicubic [10] interpolation, the SR-ResNet [15] model, the USRNet [16] model, the LiDAR-SR [5] model, and the FL-SR [6] model. Figures 7 show the super-resolution results achieved by our approach and these methods on the Ouster dataset, with upscaling factors of 4, 8, and 16. Visually, our method appears closest to the ground truth in the 64-channel 3D point clouds. Our results demonstrate excellent detail preservation and sharpness, with evenly distributed LiDAR scan lines. This indicates that our method can more accurately reconstruct the fine structure and edges of point clouds without the common problems of regularization methods, such as block artifacts, step edges, and false edges, which is important for high-quality super-resolution of point clouds.

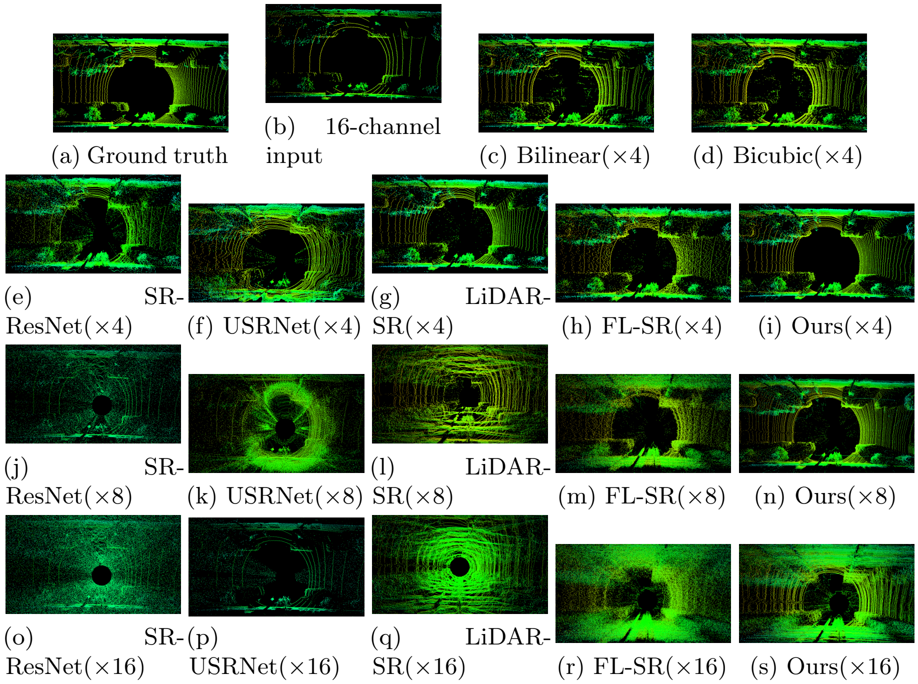


Fig. 7: The LiDAR super-resolution results obtained by our method and the comparison methods on the Ouster dataset, with upscaling factors of 4× (from 16 to 64-channel LiDAR), 8× (from 16 to 128-channel LiDAR), and 16× (from 16 to 256-channel LiDAR), respectively.

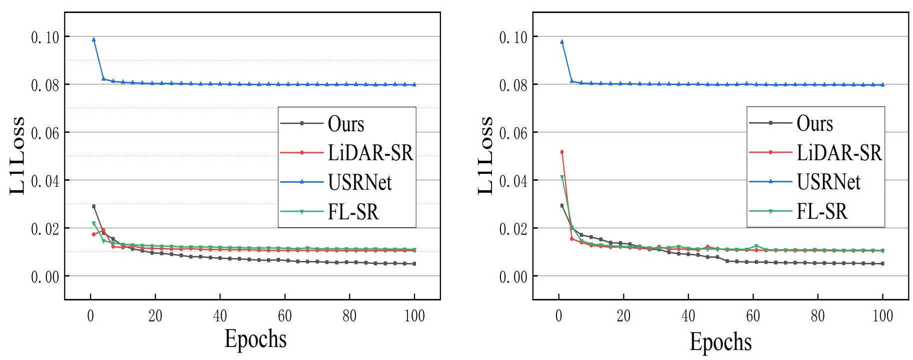
Table 2 summarizes the reconstruction results of different methods. At 64-channel resolution, the qualitative evaluation index is calculated by comparing each method's super-resolution results with the 64-channel ground truth. For 128-channel and 256-channel resolutions, due to the lack of corresponding ground truth, we use the bicubic interpolation results of the original 64-channel data as a baseline. Our proposed FOTV-HQS method outperforms other methods in various quantitative evaluation indicators. Additionally, our method requires fewer parameters compared to deep learning methods, particularly USRNet [16] and LiDAR-SR [5]. Our model has a 99.68% reduction in the number of parameters, making it suitable for practical applications with limited computing and storage resources.

4.5 Analysis and discussion

As shown in Figure 8, our proposed model demonstrates significant convergence behavior in super-resolution tasks. In the 64-channel task, the model quickly reached a low L1 loss value and remained stable, indicating its superior learning ability and stability. Compared to USRNet [16], LiDAR-SR [5] and FL-SR [6],

Table 2: Quantitative comparison of LiDAR super-resolution on Ouster dataset. The bold texts represent the best performance for each metric.

Scale	Method	↑PSNR	↑SSIM	↑EPI	↓MSE	↓MAE	↓Parameters (millions)
×4	Bilinear [9]	21.7155	0.7138	0.2868	0.0078	0.0306	-
	Bicubic [10]	21.4153	0.7124	0.2617	0.0083	0.0318	-
	SR-ResNet [15]	22.3703	0.7682	0.3385	0.0066	0.0250	1.264M
	USRNet [16]	22.1984	0.7085	0.3311	0.0069	0.0379	17.023M
	LiDAR-SR [5]	22.2991	0.7703	0.3329	0.0068	0.0223	31.042M
	FL-SR [6]	22.4562	**0.9382**	0.3445	0.0066	0.0223	0.149M
	Ours	**22.7895**	0.9346	**0.3691**	**0.0061**	**0.0212**	**0.111M**
×8	SR-ResNet [15]	12.3333	0.5095	0.1054	0.0589	0.1685	1.264M
	USRNet [16]	14.2922	0.6582	0.1520	0.0414	0.1248	17.023M
	LiDAR-SR [5]	21.4592	0.8840	0.2302	0.0078	0.0528	31.042M
	FL-SR [6]	24.1320	0.9300	0.4331	0.0044	0.0275	0.149M
	Ours	**24.7658**	**0.9364**	**0.4730**	**0.0038**	**0.0219**	**0.111M**
×16	SR-ResNet [15]	14.5289	0.5656	0.1809	0.0357	0.1311	1.264M
	USRNet [16]	4.3081	0.3227	0.0839	0.3710	0.5504	17.023M
	LiDAR-SR [5]	16.5945	0.7319	0.2011	0.0233	0.0987	31.042M
	FL-SR [6]	22.8832	0.9098	0.3462	0.0057	0.0369	0.149M
	Ours	**24.2554**	**0.9205**	**0.4482**	**0.0042**	**0.0298**	**0.111M**

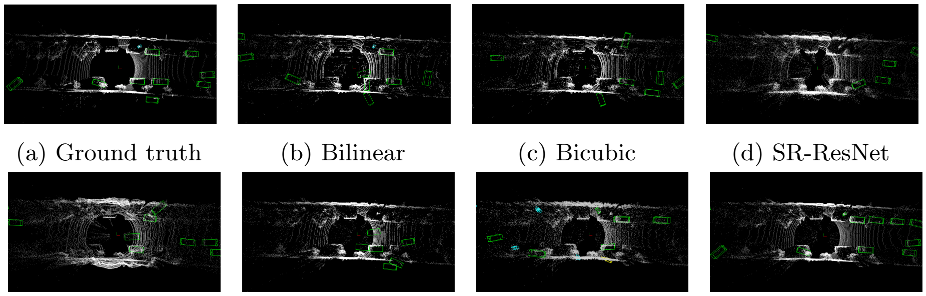


(a) L1Loss - 64 (b) L1Loss - 128

Fig. 8: L1Loss during training. We report loss results every 3 epochs. Compared with other methods, the proposed method has fast convergence speed and stable performance.

FOTV-HQS shows a faster initial convergence rate and a lower final loss value, further validating its effectiveness in processing high-dimensional data.

Compared to existing LiDAR super-resolution technologies, FOTV-HQS not only outperforms other methods in qualitative evaluation but also provides clearer and more detailed reconstructed point cloud images. This advantage is crucial



(a) Ground truth (b) Bilinear (c) Bicubic (d) SR-ResNet

(e) USRNet (f) LiDAR-SR (g) FL-SR (h) Ours

Fig. 9: Object detection results after 4× super-resolution (from 16 to 64-channel LiDAR) on Ouster dataset.

Table 3: Quantitative comparison of Object detection performance after 4× super-resolution (from 16 to 64-channel LiDAR) on the Ouster dataset. Bold text indicates the best performance.

Method	↓MAE	↑Precision	↑Recall	↑F1-score
Bilinear [9]	13.1917	0.9106	0.8962	0.9033
Bicubic [10]	14.3596	**0.9842**	0.7933	0.8785
SR-ResNet [15]	16.0655	0.8325	0.7709	0.8005
USRNet [16]	14.9459	0.9565	0.7316	0.8291
LiDAR-SR [5]	**11.5380**	0.9223	0.7650	0.8363
FL-SR [6]	13.5474	0.9167	0.8348	0.8738
Ours	12.1844	0.9600	**0.9478**	**0.9539**

for downstream applications requiring high-precision visual information. We used the PV-RCNN [29] model for 3D object detection on the super-resolution reconstructed point cloud. The results, shown in Figure 9 and Table 3, demonstrate that our method achieves good performance in 3D object detection. Notably, while our method does not achieve the lowest MAE or the highest Precision, it excels in Recall and F1-score, indicating a significant reduction in missed detections and a more balanced overall performance. This performance is particularly critical for applications where high Recall is essential to ensure the detection of as many true targets as possible.

Additionally, the FOTV-HQS model excels in parametric efficiency. Compared to conventional deep learning methods like USRNet [16] and LiDAR-SR [5], FOTV-HQS reduces the number of parameters by 99.68%. This significant reduction not only lowers the computational burden but also optimizes storage requirements. This makes FOTV-HQS particularly suitable for applications with limited computing power and storage space, such as mobile devices and real-time online processing systems.

5 Conclusions

LiDAR super-resolution enhances point cloud data quality, crucial for tasks like object detection, identification, and tracking. Traditional LiDAR super-resolution models often struggle with block artifacts, staircase edges, and misleading edges. To address these challenges, we propose the FOTV-HQS model for LiDAR super-resolution. Our approach formulates an FOTV regularization optimization problem and utilizes an end-to-end trainable iterative network to capture data attributes, enabling precise reconstruction of fine details and complex structures. Specifically, the HQS algorithm divides the problem into data fidelity and prior regularization subproblems. We then introduce a deep unfolding network that iteratively addresses these subproblems within the FOTV-HQS framework. Experiments show our approach significantly reduces parameters by up to 99.68% while maintaining good performance, making it ideal for applications with limited computing and storage resources.

References

1. Fiqri Ahmad Agung, H Herizon, Era Madona, Muhammad Rohfadli, et al. Implementation of lidar sensor for mobile robot delivery based on robot operating system. *JECCOM*, 1(2):67–79, 2023.
2. Ying Li, Lingfei Ma, Zilong Zhong, Fei Liu, Michael A Chapman, Dongpu Cao, and Jonathan Li. Deep learning for lidar point clouds in autonomous driving: A review. *IEEE TNNLS*, 32(8):3412–3432, 2020.
3. Weikun Zhen, Yaoyu Hu, Huai Yu, and Sebastian Scherer. Lidar-enhanced structure-from-motion. In *ICRA*, pages 6773–6779, 2020.
4. Michiel Vlaminck, Hiep Quang Luong, Werner Goeman, Peter Veelaert, and Wilfried Philips. Towards online mobile mapping using inhomogeneous lidar data. In *IV*, pages 845–850, 2016.
5. Tixiao Shan, Jinkun Wang, Fanfei Chen, Paul Szenher, and Brendan Englot. Simulation-based lidar super-resolution for ground vehicles. *RAS*, 134, 2020.
6. Alexandros Gkillas, Gerasimos Arvanitis, Aris S Lalos, and Konstantinos Moustakas. Federated learning for lidar super resolution on automotive scenes. In *DSP*, pages 1–5, 2023.
7. Youngsun Kwon, Minhyuk Sung, and Sung-Eui Yoon. Implicit lidar network: Lidar super-resolution via interpolation weight prediction. In *ICRA*, pages 8424–8430, 2023.
8. George Eskandar, Sanjeev Sudarsan, Karim Guirguis, Janaranjani Palaniswamy, Bharath Somashekar, and Bin Yang. Hals: A height-aware lidar super-resolution framework for autonomous driving.
9. Thierry Blu, Philippe Thévenaz, and Michael Unser. Linear interpolation revitalized. *IEEE TIP*, 13(5):710–719, 2004.
10. Robert Keys. Cubic convolution interpolation for digital image processing. *IEEE TASSP*, 29(6):1153–1160, 1981.
11. Henry Stark and Peyma Oskoui. High-resolution image recovery from image-plane arrays, using convex projections. *JOSA A*, 6(11):1715–1726, 1989.
12. Feng Shi, Jian Cheng, Li Wang, Pew-Thian Yap, and Dinggang Shen. Lrtv: Mr image super-resolution with low-rank and total variation regularizations. *IEEE TMI*, 34(12):2459–2466, 2015.
13. Jiahao Zhang, Shengrong Zhao, Hu Liang, Changchun Wen, and Chen Liang. Image super resolution via multi-regularization combining hybrid tikhonov-tv prior and deep denoiser prior. In *ICTAI*, pages 829–835, 2023.
14. Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE TPAMI*, 38(2):295–307, 2015.
15. Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *CVPR*, pages 4681–4690, 2017.
16. Kai Zhang, Luc Van Gool, and Radu Timofte. Deep unfolding network for image super-resolution. In *CVPR*, pages 3217–3226, 2020.
17. Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *ICCV*, pages 1833–1844, 2021.
18. Xiangyu Chen, Xintao Wang, Jiantao Zhou, Yu Qiao, and Chao Dong. Activating more pixels in image super-resolution transformer. In *CVPR*, pages 22367–22377, 2023.

19. Zemin Ren, Chuanjiang He, and Qifeng Zhang. Fractional order total variation regularization for image super-resolution. *SP*, 93(9):2408–2421, 2013.
20. Amine Laghrib, Anouar Ben-Loghfyry, Aissam Hadri, and Abdelilah Hakim. A nonconvex fractional order variational model for multi-frame image super-resolution. *Signal Processing: Image Communication*, 67:1–11, 2018.
21. Xiaomei Yang, Jiawei Zhang, Yanan Liu, Xiujuan Zheng, and Kai Liu. Super-resolution image reconstruction using fractional-order total variation and adaptive regularization parameters. *The Vis. Comput.*, 35:1755–1768, 2019.
22. Wenjuan Yao, Jie Shen, Zhichang Guo, Jiebao Sun, and Boying Wu. A total fractional-order variation model for image super-resolution and its sav algorithm. *JSC*, 82:1–18, 2020.
23. Manya V Afonso, José M Bioucas-Dias, and Mário AT Figueiredo. Fast image recovery using variable splitting and constrained optimization. *IEEE TIP*, 19:2345–2356, 2010.
24. Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE TIP*, 26(7):3142–3155, 2017.
25. Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, 2004.
26. Justin Joseph, Sivaraman Jayaraman, Rajadurai Periyasamy, and Simi V Renuka. An edge preservation index for evaluating nonlinear spatial restoration in mr images. *CMI*, 13(1):58–65, 2017.
27. Bartosz Sakowicz, Marek Kamiński, Robert Ritter, Rafał Kotas, Paweł Marciniak, Joanna Kupis, Przemysław Sękalski, and Andrzej Napieralski. Methods of 3d images quality assesment. In *MIXDES*, pages 123–128, 2015.
28. Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *CR*, 30(1):79–82, 2005.
29. Shaoshuai Shi, Chaoxu Guo, Li Jiang, Zhe Wang, Jianping Shi, Xiaogang Wang, and Hongsheng Li. Pv-rcnn: Point-voxel feature set abstraction for 3d object detection. In *CVPR*, pages 10529–10538, 2020.