Quality Restoration of Point-Cloud-Derived 2D Projections: A Comparative Study of Void-Filling Techniques

Md Rakibul Islam Chowdhury [1], Sang Hyeok Han [1,2], Jong Won Ma [3]

[1] Dept of Building, Civil and Environmental Engineering, Concordia University, Montréal, QC, Canada – mdrakibulislam.chowdhury@mail.concordia.ca
[2] Centre for Innovation in Construction and Infrastructure Engineering and Management (CICIEM), Gina Cody School of Engineering and Computer Science, Concordia University, Montréal, QC, Canada – sanghyeok.han@concordia.ca
[3] School of Civil and Environmental Engineering, Yonsei University, Seoul, South Korea – jongwon.ma@yonsei.ac.kr

Keywords: 3D Point Cloud, 2D Projection, Image Enhancement

Abstract

Point-cloud-derived 2D projections enable generating unlimited virtual views for indoor scene analysis and dataset creation. However, projecting irregular 3D samples onto a dense image grid commonly produces void pixels due to sparsity, occlusions, and incomplete scan coverage. These projection-induced artifacts degrade the visual fidelity of rendered images and limit their usefulness in downstream image-based workflows. This study investigates void-filling strategies tailored to point-cloud-generated RGB projections and provides a comparative evaluation of three representative approaches: (i) K-nearest neighbor (KNN) interpolation with KD-Tree accelerated neighbor search, (ii) a rule-based neighborhood method (NNRule) that adapts filling behavior using local variability to preserve edges, and (iii) a mask-normalized Gaussian-weighted propagation method that diffuses valid color information into void regions. Experiments were conducted on multi-view perspective projections generated from Stanford Large-Scale 3D Indoor Spaces Dataset (S3DIS) Area 3, totalling 5,520 images. Restoration quality was assessed using standard pixel-level metrics such as MAE, RMSE, PSNR, and SSIM. Quantitative results show that Gaussian-weighted propagation achieved the best overall performance, followed by NNRule, while KNN performed weakest numerically. Qualitative comparisons further indicate that KNN produces the most visually realistic texture appearance, whereas diffusion-based filling is softened fine details. Finally, the study establishes a practical baseline that enables both academic researchers to advance point-cloud-to-image restoration without relying on paired RGB datasets and industrial practitioners to deploy light weight void-filling pipelines in real-world applications such as digital twins, indoor robotics, facility management, and augmented reality.

1. Introduction

A point cloud is fundamentally a collection of three-dimensional spatial coordinates that depicts the surface geometry of a 3D scene (Daghigh et al., 2022). 3D point clouds have become a foundational data source for indoor scene understanding, robotics, construction monitoring, and digital twin applications. In recent years, the acquisition of point-cloud data has become widely accessible and cost-effective due to rapid advancements in sensor technologies. In fact, two-dimensional image processing technologies have reached a high level of maturity, the focus of the computer vision community has increasingly expanded to three-dimensional representations that more accurately reflect the real world (Zhang et al., 2023c). Within this domain, point clouds have emerged as a foundational data format due to their rich spatial and geometric expression. A practical advantage of point-cloud data is that, once captured, it enables the generation of an unlimited number of virtual 2D images from arbitrary viewpoints through camera-based projection, providing flexible access to image-like representations without revisiting the physical environment (Lawin et al., 2017). Unlike conventional image acquisition, which requires physically revisiting a space and manually collecting photographs, virtual projections allow flexible data augmentation, multi-view analysis, and the creation of large-scale datasets for downstream tasks. This capability opens opportunities for diverse applications such as inventory management, virtual inspection, 2D semantic segmentation dataset generation, and multi-view learning pipelines that rely on abundant 2D samples (Zhang et al., 2023a).

While projecting point clouds into 2D space offers immense advantages for data augmentation and multi-view analysis, converting irregular 3D spatial samples into a structured, dense 2D pixel grid introduces several significant challenges. Because point clouds are discrete representations rather than continuous surfaces, their spatial fidelity is heavily reliant on the acquisition sensor. 3D scanning inherently involves variable sensor accuracy, resulting in notable density differences and outliers (Huang et al., 2017). Consequently, projected images often contain void regions (black pixels) caused by this inherent sparsity (Qi et al., 2017b), physical occlusions within the scene (Qi et al., 2017a), or limited sensor coverage. When regions of low point density often dictated by the distance from the scanner or the specific imaging mechanism, these are projected onto a high-resolution 2D image plane, the spatial gaps between points directly manifest as these empty pixels. The resulting image quality is fundamentally tied to the resolution and density of the original scan: low-density areas and inherent sensor noise create larger voids, jagged object boundaries, and missing surface textures. Furthermore, uneven sampling and spatial outliers lead to inconsistent or distorted color patterns in the rendered images. Ultimately, degradation is not only a visual issue; it can also negatively affect any workflow that relies on dense image evidence. Void pixels and discontinuities introduce structured missing data that can bias feature extraction and reduce the reliability of projected images when they are used for analysis or as inputs for downstream image-based learning pipelines, including the generation of 2D semantic segmentation datasets from 3D environments.

To address these limitations, this study evaluates several image quality restoration techniques designed specifically for point-cloud–generated 2D projections. We compare three representative methods named K-nearest neighbor (KNN) interpolation, rule-based void filling, and weighted Gaussian smoothing and examine how effectively they enhance projection completeness and structure preservation. The contributions of this study are as follows. **Firstly**, three lightweight void-filling procedures are presented for restoring missing regions in point-cloud–derived 2D projections using local neighborhood information and Gaussian-weighted propagation, enabling effective completion of void pixels without introducing complex model assumptions. **Secondly**, since no training is required, the proposed approaches substantially reduce computational cost and eliminate dependence on large-scale paired RGB datasets or pretrained weight files. As a result, a practical and generalizable baseline is established that can be readily adopted by researchers and practitioners for void restoration in real-world point-cloud-to-image pipelines where paired RGB data are limited or unavailable.

2. Background Research

2.1 Point-Cloud–To–Image Projection

Transforming unstructured 3D point clouds into structured 2D images is a well-established technique, primarily driven by the desire to leverage the maturity, efficiency, and expert architectures of Convolutional Neural Networks (CNNs) for tasks like 2D semantic segmentation. In general, projection-based representations provide a practical bridge between irregular 3D point samples and dense 2D grids, enabling efficient storage, visualization, and large-scale generation of virtual views from arbitrary camera poses. This is particularly valuable for indoor environments where captured point clouds can be re-rendered from unlimited viewpoints, supporting multi-view analysis and synthetic image generation without requiring additional physical image acquisition (Lawin et al., 2017; Zhang et al., 2023a). Projection-based techniques typically follow a common principle: each 3D point is mapped to a 2D pixel location using a predefined projection model, and per-pixel attributes (e.g., RGB, depth, intensity, semantic labels) are accumulated to form a raster image. The subsequent processing depends strongly on how the projection is defined and what information is preserved during the 3D-to-2D conversion. In semantic segmentation pipelines, these projected images are often segmented using 2D CNNs, and pixel-wise predictions are then re-projected back to the original point cloud to obtain point-wise labels. As a result, the quality of the final outcome is highly dependent on the projection choice and the completeness of the generated 2D representation (Su et al., 2015; Lawin et al., 2017; Guo et al., 2021b).

Broadly, point-cloud–to–image projection can be categorized into multi-view perspective projection, bird’s-eye view (BEV) projection, and spherical (range image) projection. Multi-view projection renders the scene from multiple virtual camera viewpoints to capture complementary perspectives and reduce viewpoint-dependent occlusions. Multi-view strategies have been widely explored because they can represent complex scenes by aggregating evidence from several images, although they require careful selection of camera positions, view angles, and fusion strategies to cover the scene comprehensively (Su et al., 2015; Robert et al., 2022). Bird’s eye view (BEV) projection transforms 3D points into a top-down 2D grid, typically encoding height statistics or multi-channel features. This representation is effective for planar layout reasoning and object footprint analysis, but it may lose information loss such as vertical structure and depth (Zhang et al., 2020; Beltrán et al., 2018). Spherical projection converts point clouds into range images by mapping points according to azimuth and elevation, producing compact 2D representations particularly suited for LiDAR-like scanning patterns; however, spherical images can suffer from distortions and occlusion effects depending on the sensor viewpoint and scene geometry (Wu et al., 2018; Milioto et al., 2019). Spherical Projection involves placing a virtual sensor at the center of the scene and casting the surrounding 360° laser scans onto a cylindrical or spherical surface, which is then unrolled into a flat 2D range image (Jhaldiyal and Chaudhary, 2022). Although spherical imaging delivers comprehensive, wide-field observation, its primary drawbacks include edge warping and the possibility of visual occlusion (Singh and Yadav, 2024).

Unlike BEV or Spherical methods, which rely on fixed, global coordinate transformations, multi-view projection involves rendering the 3D scene from the perspective of multiple, arbitrarily placed virtual cameras. This method generates conventional 2D images that closely mimic human vision and standard photographic datasets. Because multi-view projection allows for complete flexibility regarding virtual camera placemegonzalezt, viewing distance, and angle, it provides the most comprehensive scene representation. This study adopts a multi-view perspective projection setting to generate indoor 2D images from point clouds and focuses specifically on improving projection completeness through void restoration.

2.2 Image Restoration Techniques

Image restoration methods aim to recover missing pixel values while preserving structural boundaries and local appearance consistency. For point-cloud–derived 2D projections, restoration primarily targets void pixels that occur when no 3D samples fall onto the image plane. These missing regions are not random noise; they are structured artifacts driven by sampling density, occlusion, and sensor coverage. In this study, restoration techniques are broadly classified into two categories: training-free methods (e.g., KNN, NNRule, and Gaussian-weighted smoothing) and learning-based generative methods (e.g., GANs and diffusion models), which inherently rely on large-scale training datasets and pretrained weights.

2.2.1 Training-Free Methods

(a) Nearest Neighbour Based Method

Self-similarity based patch search is widely used in tasks such as image restoration because each patch captures local appearance information, so finding similar patches elsewhere in the image provides useful cues for reconstructing missing or degraded regions (Banerjee, 2020; Giraud et al., 2017). Based on this idea, nearest-neighbour based restoration fills missing pixels using values from nearby valid samples as that local neighborhoods provide reliable cues. In this family, nearest-neighbour filling assigns each missing pixel the value of its closest valid sample, while K-nearest neighbour (KNN) interpolation improves robustness by aggregating multiple neighbors, often using inverse-distance weighting so that closer samples contribute more strongly. To make such neighbor retrieval efficient at scale, KD-Tree indexing is commonly used to speed up spatial nearest-neighbour search. By building a KD-Tree over the coordinates of valid pixels, the nearest (or K nearest) valid samples for each void pixel can be retrieved efficiently, enabling fast NN/KNN replacement (Guo et al., 2021a).

Recent image restoration studies also support the usefulness of KNN-based mechanisms beyond classical interpolation. Specifically, KNN-guided designs have been proposed to better

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-133-2026 | © Author(s) 2026. CC BY 4.0 License.

exploit non-local self-similarity and long-range dependencies in degraded images, since purely local CNN processing or restricted attention may miss global correspondences needed to recover fine structures. In this direction, Zeng et al. (2026) introduced a KNN-guided residual learning framework for image restoration and highlight that incorporating KNN-based modules helps capture long-range relationships and improves preservation of high-frequency details and edges, including in cases involving missing regions.

(b) NNrule Based Method

Edge-aware restoration is popular in image denoising and inpainting because, when missing regions are filled using uniform averaging, boundaries tend to get blurred and important structural cues are lost. In particular, edge-aware smoothing is commonly defined as removing fine details or artifacts while preserving large-scale structures and sharp edges, and it has been studied extensively due to its importance in downstream vision tasks (Deng, 2017). Moreover, a central message repeated across the edge-preserving literature is that low-pass smoothing is unreliable at discontinuities; therefore, averaging across edges should be avoided, whereas averaging within homogeneous regions is generally acceptable and beneficial (Kaur et al., 2017). Consistent with this, more recent restoration studies also attribute edge blur to processing edge and non-edge areas in the same manner, emphasizing that explicit attention to boundaries is necessary to prevent structural degradation (Liu et al., 2022). Building on these principles, rule-based neighborhood strategies use simple local statistics to adapt the restoration behavior depending on whether a missing pixel lies in a smooth surface region or near an edge/texture boundary. For example, in low-variation neighborhoods, mean-consistent filling supports continuity and suppresses artifacts, whereas in high-variation neighborhoods (edge-like regions), selecting a representative local value helps reduce cross-edge mixing and better preserves sharp transitions. Overall, this adaptive “switching” logic follows the same design philosophy behind edge-preserving filters (e.g., bilateral and guided filtering), which aim to smooth within regions while maintaining strong discontinuities at boundaries (Deng, 2017; Kaur et al., 2017).

(c) Gaussian Weighted Method

Gaussian-weighted smoothing is long-established, training-free strategies in image denoising and restoration, where missing or degraded pixel values are estimated by spatially aggregating nearby observations with distance-decaying weights. Surveyed image filtering literature consistently highlights Gaussian filtering, anisotropic diffusion, and related smoothing operators as fundamental tools for improving image quality, particularly due to their simplicity and computational efficiency (Patil and Bendre, 2023). However, it is also widely noted that standard Gaussian filtering tends to blur edges because it averages intensities without explicitly modeling regional disparity, motivating edge-aware variants such as bilateral filtering and anisotropic diffusion (Patil and Bendre, 2023). More over, gaussian-weighted propagation produces stable, continuous fills that reduce sparsity artifacts, yet may soften high-frequency texture and sharp boundaries near occlusion edges.

Gaussian weighting is also commonly used as an explicit mechanism for weighted averaging, where contributions are controlled by a Gaussian function of intensity difference or spatial distance. For example, Tan et al. (2017) define Gaussian weights for similarity-weighted averaging and show that the weight decreases as the pixel difference increases, enabling robust aggregation while suppressing inconsistent samples. Also, Ding et al. (2019) propose a Gaussian-weighted nonlocal texture similarity measure to retrieve candidate patches for inpainting, showing that Gaussian weights are useful for emphasizing reliable local information while maintaining natural transitions in the reconstructed area. In addition, a lightweight approximation using down sampling followed by a small Gaussian filter can significantly reduce resource requirements while retaining effective smoothing behavior (Tan et al., 2017). Motivated by these established findings, the Gaussian-weighted method in this paper is adopted as a representative training-free baseline that prioritizes numerical stability and global coherence for filling projection voids, while acknowledging that texture sharpness and edge crispness may be reduced compared with neighbor-sampling strategies.

2.2.2 Learning-based methods

(a) Diffusion based methods

Diffusion-based restoration methods fill missing pixels by propagating information from known regions into void regions through iterative smoothing or kernel-based averaging. Diffusion-based methods have emerged as a highly effective paradigm for image restoration and inpainting, demonstrating superior generative capabilities compared to earlier approaches such as traditional PDE-based diffusion techniques (Li et al., 2025). By modeling the gradual denoising process from noise to structured data, these models excel at producing photo-realistic, semantically coherent, and high-fidelity reconstructions, particularly for large or irregular voids (Lugmayr et al., 2022; Saharia et al., 2022a). Notable recent methods include RePaint (Lugmayr et al., 2022) for zero-shot irregular-hole filling, SR3 and Palette (Saharia et al., 2022b) for conditional super-resolution and style-guided inpainting, and posterior-guided frameworks such as DPS and COPAINT (Chung et al., 2022; Zhang et al., 2023b). Key benefits include powerful generative priors for realistic texture recovery, strong handling of structured missing regions, excellent generalization across degradation types, and the ability to produce diverse plausible outputs aligned with human perception (Li et al., 2025).

However, diffusion models heavily depend on extensive pre-training on massive datasets and access to high-quality weight files or checkpoints. Full training or even fine-tuning requires substantial computational resources, making them difficult or impractical to apply in resource-constrained environments (Li et al., 2025; Xia et al., 2023).

(b) Generative Adversarial Network (GAN) based methods

GAN-based methods have emerged as a highly effective paradigm for image restoration and inpainting, demonstrating superior generative capabilities compared to earlier traditional or CNN-based approaches (Feng et al., 2025). By employing adversarial training between a generator and a discriminator, GANs excel at producing sharp, realistic textures and high-frequency details that are often lost in classical smoothing or interpolation methods. Notable recent methods include GENet (Feng et al., 2025), a two-branch ensemble network with pyramid hybrid attention and pretrained transfer-learning for smudged restoration; CycleGAN (Zhu et al., 2017) for unpaired-to-paired synthesis; Pix2Pix (Isola et al., 2017)) for conditional translation; Restore-GAN (Rong et al., 2023), SN PatchGAN (Li et al., 2023) and SA-GAN (Cao et al., 2023). These approaches achieve excellent perceptual quality, edge preservation, and natural style rendering, making them particularly strong for restoring occluded, blurred, or low-quality images on benchmarks such as Helen Face and CelebA datasets.

Key benefits of GAN-based methods include fast single-step inference, strong high-frequency texture generation, and realistic outputs aligned with human perception. However, these advantages come at a significant cost: GAN-based methods heavily depend on extensive adversarial training, large paired or unpaired datasets, and access to high-quality pre-trained weight files. Training is notoriously unstable due to mode collapse and vanishing gradients, while fine-tuning requires substantial computational resources and careful hyperparameter tuning. In resource-constrained environments or when high-quality checkpoints are unavailable, these dependencies make GAN-based methods difficult or impractical to apply effectively.

As a result, KNN-based filling method is particularly appropriate for this study because the input data consists solely of 2D images projected from a 3D point cloud (S3DIS dataset), with no access to corresponding original high-resolution RGB photographs or paired clean–degraded ground truth required for supervised training. Generative approaches such as GANs and diffusion models rely heavily on large-scale paired datasets and pre-trained weight files to learn realistic texture synthesis and style transfer; without these, their adversarial or iterative sampling processes become unstable, computationally prohibitive, or impossible to train from scratch on sparse projected voids alone. In contrast, the KNN-based (Inverse Distance Weighted) interpolation operates in a fully unsupervised, training-free manner: for each void pixel it constructs a KDTree over the normalized coordinates of valid pixels and fills the missing value as a distance-weighted average of its K nearest neighbors, using only the spatial information already present in the projected image itself. This lightweight, single-pass approach preserves local texture and sharp boundaries (e.g., wall–floor transitions and object contours) without introducing the excessive smoothing observed in Gaussian or diffusion-based fills, while requiring negligible computational resources and zero training overhead.

3. Research Methodology

S3DIS dataset with multi-view perspective projection method was used to generate 2D images, and then details the three void-filling techniques and evaluation metrics adopted to assess restoration quality.

After generating multi-view 2D images from the 3D point cloud using a perspective-projection approach, we applied three complementary void-filling techniques to address missing pixel regions (Figure 1). First, a rule-based neighborhood method is employed to infer the replacement color based on the local statistical properties of surrounding valid pixels. A small radius is examined around every void pixel, and the variability of the surrounding colors determines the type of fill. When the neighborhood shows high color variation, which usually indicates edges or textured areas, the method selects the brightest nearby pixel to help maintain sharp details. In smoother, low-variation regions, it instead chooses the pixel whose color is closest to the local average, helping produce clean and natural transitions. The second method, a K-nearest neighbour (KNN) interpolation technique, treats all valid pixels as spatial samples and uses a KD-Tree to efficiently identify the nearest neighbors for each void pixel. The colors of the nearest valid samples are combined using inverse-distance weighting, enabling spatially coherent interpolation across both small and moderately large missing regions. The third technique, a Gaussian-weighted diffusion method, propagates color information into void areas by applying Gaussian filtering to both the RGB channels and the corresponding validity mask. Together, these methods form a comparative framework for evaluating structural preservation, interpolation effectiveness, and smoothness in the restoration of void regions within projected point-cloud imagery.

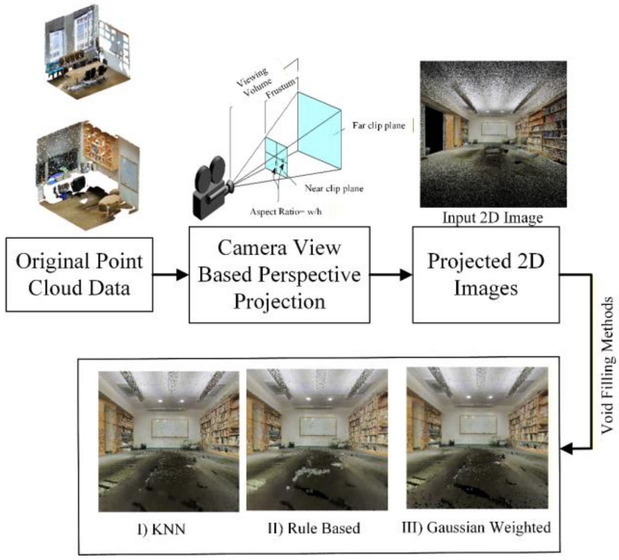

Figure 1. Research Methodology

3.1 Dataset Description

In this study, S3DIS dataset provides 3D point clouds captured from six large indoor areas covering more than 270 rooms (Armeni et al., 2016). Each point is annotated with one of 13 semantic categories including structural and architectural components. The dataset defines 12 semantic categories, covering structural components (ceiling, floor, wall, beam, column, window, and door), furniture (table, chair, and sofa), and office objects (bookcase and board), while all remaining items are labeled as “clutter.” In this study, 5520 images were generated from Area 3 for to justify the void filling techniques. Area 3 of S3DIS was selected as a controlled evaluation subset because room-type diversity is combined with a smaller scene count. This selection enables void-filling methods to be assessed across different indoor configurations and occlusion scenarios without introducing excessive dataset scale.

3.2 Generating 2D Images

Figure 2 presents the multi-view perspective projection pipeline used to generate 2D images from indoor point-cloud scenes. For each room (e.g., Area3_Lounge1), a set of virtual cameras is positioned around the scene in four principal directions (left, right, front, and back), with the field-of-view configured to capture the interior layout. To improve coverage of occluded regions and height-dependent structures, additional camera poses are placed at multiple elevations within the room. Using these camera configurations, the 3D points are projected onto the image plane to produce a collection of viewpoint-diverse 2D projections, which form the input for subsequent void-filling and image-quality evaluation.

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-133-2026 | © Author(s) 2026. CC BY 4.0 License.

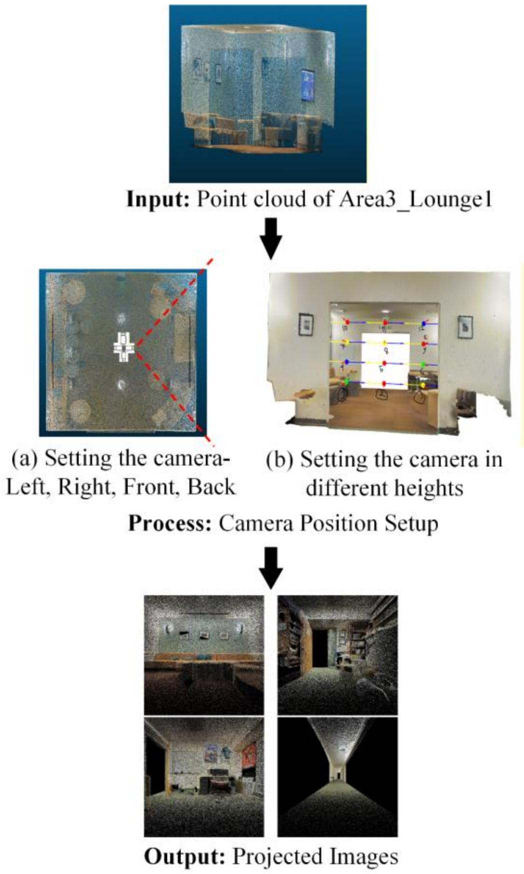

Figure 2. Generating 2D images from point cloud

3.3 Void Filling Strategies

This section provides a description of the three void-filling strategies used in this study and demonstrates detailed explanations of how missing pixels in point-cloud–derived 2D projections are restored using KNN-based interpolation, NNrule-based filling, and Gaussian-weighted propagation.

3.3.1 KNN-Based (fill_voids_knn):

The KNN-based strategy treats all valid pixels as spatial samples and estimates each void pixel from its K nearest valid neighbors in the image plane. A KD-Tree is constructed over the coordinates of valid pixels to enable efficent nearest-neighbor queries. For every void pixel, the K nearest valid pixels are retrieved and their RGB values are aggregated using inverse-distance weighting, so closer samples contribute more strongly. To fill void pixels in the input image, we use a K-nearest neighbors interpolation method based on Inverse Distance Weighting (IDW). This well-established technique is commonly applied for scattered data interpolation and hole filling in images and digital elevation models (Huang et al., 2020; Dong et al., 2025). Figure 3 shows the pseudo code for this method.

Algorithm 1 Nearest-Neighbor Void Filling for RGB Images and Label Maps

Require: RGB image $\mathbf{I} \in \mathbb{R}^{H\times W\times 3}$, label map $\mathbf{M} \in \mathbb{R}^{H\times W}$
Ensure: Filled RGB image $\mathbf{I}_{filled}$, filled label map $\mathbf{M}_{filled}$
Identify RGB void pixels $\mathcal{V}_I$ where $\mathbf{I}(x, y) = (0, 0, 0)$
if $\mathcal{V}_I \neq \emptyset$ **then**
  Compute Euclidean distance transform on $\mathcal{V}_I$
  Find the nearest valid RGB pixel for each void pixel
  Replace each void RGB pixel with the nearest valid RGB value
  $\mathbf{I}_{filled} \leftarrow$ filled RGB image
else
  $\mathbf{I}_{filled} \leftarrow \mathbf{I}$
end if
Identify label void pixels $\mathcal{V}_M$ where $\mathbf{M}(x, y) = 0$
if $\mathcal{V}_M \neq \emptyset$ and $\mathbf{M}$ is not entirely void **then**
  Compute Euclidean distance transform on $\mathcal{V}_M$
  Find the nearest valid labeled pixel for each void pixel
  Replace each void label pixel with the nearest valid label
  $\mathbf{M}_{filled} \leftarrow$ filled label map
else
  $\mathbf{M}_{filled} \leftarrow \mathbf{M}$
end if
return $\mathbf{I}_{filled}, \mathbf{M}_{filled}$

Figure 3. Pseudo code for Nearest-Neighbor based filling

For each void pixel at normalized coordinates $(t_e, j_e) \in [0,1] \times [0,1]$, we build a KDTree over the coordinates of all valid (non-void) pixels to quickly find the K spatially closest valid neighbors. Let $z(i_k, j_k)$ be the color value (e.g., RGB) of the k-th nearest valid neighbor, and let d_k be the Euclidean distance from (t_e, j_e) to (i_k, j_k) in normalized space.

The color value at the void pixel is then computed as the distance-weighted average of these neighbors using equation (1):

$$z(t_e, j_e) = \frac{\sum_{k=1}^{K} w_k\, z(i_k, j_k)}{\sum_{k=1}^{K} w_k} \quad \text{with} \quad w_k = \frac{1}{d_k^p + \epsilon} \tag{1}$$

where the weight is $w_k = 1 / (d_k^p + \varepsilon)$, p is the neighborhood parameter and $\varepsilon = 10^{-6}$ is a small value to avoid division by zero.

3.3.2 NNRule-Based (fill_voids_radius_based):

The rule-based method fills void pixels by analyzing a local radius window around each void pixel and adapting the fill decision based on neighborhood statistics (Gonzalez and Woods, 2018). Specifically, the local color variability (e.g., variance or range) is used as an indicator of whether the void pixel lies near an edge/texture boundary or within a smooth region using the pseudo code illustrated in Figure 4. In high-variation neighborhoods (edge-like regions), the method selects a representative nearby pixel value that helps preserve sharp transitions (e.g., choosing the brightest/most dominant nearby valid pixel). In low-variation neighborhoods (smooth regions), the method assigns a value close to the neighborhood mean to ensure smooth and natural transitions. This adaptive logic reduces edge blurring compared to purely smoothing-based methods.

Algorithm 2 Radius-Based Void Filling for RGB Images and Label Maps

Require: RGB image $\mathbf{I} \in \mathbb{R}^{H\times W\times 3}$, label map $\mathbf{M} \in \mathbb{R}^{H\times W}$, radius r, color threshold τ
Ensure: Filled RGB image $\mathbf{I}_{filled}$, filled label map $\mathbf{M}_{filled}$
Identify void pixels $\mathcal{V}$ where RGB$=(0,0,0)$
if $\mathcal{V}$ is empty **then**
  return $\mathbf{I}, \mathbf{M}$
end if
Initialize $\mathbf{I}_{filled} \leftarrow \mathbf{I}$
Initialize $\mathbf{M}_{filled} \leftarrow \mathbf{M}$
for all void pixels p in $\mathcal{V}$ **do**
  Extract neighborhood $\mathcal{N}_r(p)$ within radius r
  $\mathcal{N}_I \leftarrow$ valid RGB neighbors (non-zero)
  $\mathcal{N}_M \leftarrow$ valid label neighbors (non-zero)
  if $\mathcal{N}_I \neq \emptyset$ **then**
    Compute RGB standard deviation σ_{rgb} over $\mathcal{N}_I$
    $\mu_{std} \leftarrow \text{mean}(\sigma_{rgb})$
    if $\mu_{std} > \tau$ **then**
      // High variation $\rightarrow$ preserve edges
      Select pixel with maximum RGB intensity
    else
      // Low variation $\rightarrow$ smooth region
      Compute mean RGB μ
      Select pixel closest to μ
    end if
    Assign selected RGB value to $\mathbf{I}_{filled}(p)$
  end if
  if $\mathcal{N}_M \neq \emptyset$ **then**
    Assign modal label of $\mathcal{N}_M$ to $\mathbf{M}_{filled}(p)$
  end if
end for
return $\mathbf{I}_{filled}, \mathbf{M}_{filled}$

Figure 4. Pseudo code for Rule-based filling

3.3.3 Gaussian Weighted (fill_voids_weighted):

In the third method, void regions in the projected RGB images were filled using a Gaussian-weighted normalized smoothing scheme (Figure 5). Specifically, void pixels were first identified as zero-valued RGB locations. For each color channel, Gaussian smoothing was applied to the valid pixel intensities, and the result was normalized by the Gaussian-smoothed valid-pixel mask to avoid intensity attenuation near missing regions. For the label maps, each non-zero class was converted into a one-hot channel and processed using the same normalized Gaussian smoothing strategy. The final label at each pixel was assigned by selecting the class with the maximum smoothed response, while original non-zero labels were preserved to avoid altering already valid annotations.

Algorithm 3 Gaussian-Weighted Void Filling for RGB Images and Label Maps

Require: RGB image $\mathbf{I} \in \mathbb{R}^{H\times W\times 3}$, label map $\mathbf{M} \in \mathbb{R}^{H\times W}$, Gaussian parameter σ
Ensure: Filled RGB image $\mathbf{I}_{filled}$, filled label map $\mathbf{M}_{filled}$
Identify RGB void mask Ω_v where $\mathbf{I}(x,y)=(0,0,0)$
Define valid RGB mask $\Omega = \neg\Omega_v$
Initialize $\mathbf{I}_{filled} \leftarrow \mathbf{I}$
for $c = 1$ to 3 **do**
  Extract channel $\mathbf{I}^{(c)}$
  Set void pixels in $\mathbf{I}^{(c)}$ to zero
  $\mathbf{S}^{(c)} \leftarrow \text{GAUSSIANFILTER}(\mathbf{I}^{(c)}, \sigma)$
  $\mathbf{N} \leftarrow \text{GAUSSIANFILTER}(\Omega, \sigma)$
  $\mathbf{I}^{(c)}_{filled} \leftarrow \mathbf{S}^{(c)}/(\mathbf{N}+\epsilon)$
end for
Extract valid class set $\mathcal{C} = \{\text{unique labels in } \mathbf{M}\} \setminus \{0\}$
if $\mathcal{C} = \emptyset$ **then**
  return $\mathbf{I}_{filled}, \mathbf{M}$
end if
Define valid label mask Ψ where $\mathbf{M} \neq 0$
$\mathbf{N}_L \leftarrow \text{GAUSSIANFILTER}(\Psi, \sigma)$
for all class $k \in \mathcal{C}$ **do**
  Construct one-hot map $\mathbf{C}_k$ from $\mathbf{M}$
  $\mathbf{S}_k \leftarrow \text{GAUSSIANFILTER}(\mathbf{C}_k, \sigma)$
  $\mathbf{P}_k \leftarrow \mathbf{S}_k/(\mathbf{N}_L+\epsilon)$
end for
For each pixel, find class $k^* = \arg\max_k \mathbf{P}_k$
Assign class k^* to $\mathbf{M}_{filled}$
Restore original non-zero labels from $\mathbf{M}$ into $\mathbf{M}_{filled}$
return $\mathbf{I}_{filled}, \mathbf{M}_{filled}$

Figure 5. Pseudo code for Gaussian-based filling

4. Experimental Results

This section presents the experimental validation of the proposed void-filling framework on point-cloud–derived 2D projections. The experimental setup and evaluation metrics are first outlined, and quantitative comparisons of the three restoration strategies are then reported using standard image-quality metrics.

4.1 Experimental Setup

Experiments were conducted using the S3DIS, focusing on Area 3 scenes. Multi-view RGB images were generated from the original point clouds using the perspective-projection configuration described in Section 3.2, producing a total of 5520 projected images. Three void-filling methods were evaluated: (i) KNN-based interpolation (fill_voids_knn), (ii) NNrule-based filling (fill_voids_radius_based), and (iii) Gaussian weighted filling (fill_voids_weighted). All methods were applied to the same projected images under identical conditions to ensure a fair comparison.

4.2 Evaluation Metrics

To quantitatively evaluate the quality of the restored 2D projections, we computed several standard image-quality metrics, including Mean Absolute Error (MAE) to capture overall intensity reconstruction accuracy, Root Mean Squared Error (RMSE) was used alongside MAE to capture reconstruction errors, Peak Signal-to-Noise Ratio (PSNR) for measuring pixel-level fidelity, Structural Similarity Index Measure (SSIM) for assessing perceptual structural consistency. The equations used for the metrices are below:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x_i - yi| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \quad (3)$$

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad (4)$$

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (5)$$

Where, x_i= ground truth pixel, y_i= predicted pixel N= total number of pixels × channels, MAX_I=255, MAX=255 for 8-bit images, μ_x, μ_y = mean intensity, σ_x^2,σ_y^2= variance, σ_{xy}= covariance, C1 ,C2 = stability constants

4.3 Quantitative Results

Table 1 summarizes the quantitative comparison among the evaluated void-filling methods. Overall, the Gaussian Weighted method achieved the best performance across all reported metrics, producing the lowest reconstruction error (MAE 68.54, RMSE 96.45) and the highest fidelity and structural similarity (PSNR 8.582 dB, SSIM 0.1796). The NNRule method ranked second, with MAE 76.12, RMSE 105.05, PSNR 7.88 dB, and SSIM 0.1528, indicating improved structural preservation compared with basic neighbor interpolation. The KNN baseline showed the weakest quantitative performance (MAE 87.69, RMSE 109.97, PSNR 7.454 dB, SSIM 0.0926), suggesting that purely neighbor-driven interpolation is more sensitive to boundary mixing and local inconsistencies. Gaussian-weighted propagation yields more accurate and structurally consistent restoration than direct KNN filling, while NNRule provides a competitive trade-off through edge-aware local decisions.

Method	MAE	RMSE	PSNR (dB)	SSIM
KNN	87.69	109.97	7.454	0.0926
NNRule	76.12	105.05	7.88	0.1528
Gaussian Weighted	**68.54**	**96.45**	**8.582**	**0.1796**

Table 1. Quantitative results of the restored images

Figure 6 presents a visual comparison of the restored projections produced by the three void-filling strategies across a sample of five representative indoor scenes: (a) lounge_1, (b) office_3, (c) conference room_1, (d) WC_2, and (e) hallway_2. The original projected images contain severe void regions and salt-and-pepper–like sparsity artifacts caused by occlusions and incomplete point sampling. Among the restoration methods, the KNN-based approach produces the most visually realistic results across all scene types (Figure 6), as it fills missing pixels using locally consistent samples and preserves the natural appearance of surfaces without introducing excessive smoothing (e.g., wall–floor boundaries and object contours). In contrast, the weighted Gaussian-based method generates continuous fills, but fails to preserve edges and suppress fine texture, which reduces visual realism in cluttered indoor environments such as office_3 and hallway 2. The rule-based method improves completeness while attempting to preserve edges through local variability analysis than the weighted gaussian method; however, its heuristic selection can occasionally yield less natural transitions in complex regions, particularly around occlusion boundaries in WC_2 and along long planar structures in hallway_2. Overall, the qualitative results indicate that KNN better maintains plausible local texture and scene realism, while the other methods exhibit stronger smoothing or heuristic artifacts.

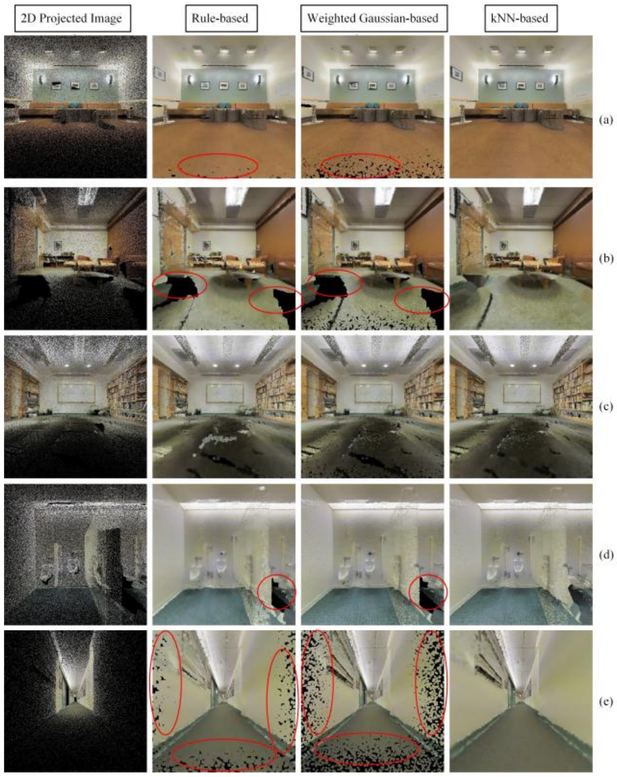

Figure 6. Visual assessment of the restored images (a) lounge_1, (b) office_3, (c) conference room_1, (d) WC_2, (e) hallway_2

5. Discussions & Limitations

The key novelty of this study is the demonstration of lightweight, training-free classical methods that omits the training hassle of GANs and diffusion models for void restoration. Generative approaches that require paired RGB ground-truth and pre-trained weights (unavailable here), the proposed methods operate unsupervised on the sparse projection alone, preserving sharp boundaries and local texture with zero training overhead. This establishes a practical baseline for real-world point-cloud-to-image restoration without external RGB data. This quantitative evidence highlights the contribution of this study by establishing an effective and generalizable restoration strategy tailored for 2D projections derived from 3D point clouds.

5.1 Discussions

The comparative results reveal that void filling in point-cloud–derived projections involve a fundamental trade-off between numerical fidelity and perceptual realism, driven by how each method propagates information into missing regions. Gaussian-weighted propagation performs best in the reported image-quality metrics because it acts as a regularized smoother in pixels by aggregating evidence from a wider neighborhood and normalizing with the validity mask, it produces a smooth and spatially consistent solution that reduces high-frequency variation inside void regions (Dong et al., 2024). Pixel-wise metrics such as MAE/RMSE and fidelity measures such as PSNR/SSIM tend to reward this behavior because smoothing suppresses local deviations and increases structural agreement at a coarse scale. In other words, the Gaussian approach behaves like a low-pass, stability-oriented completion that aligns well with the assumptions implicit in these metrics.

In contrast, the KNN method shows the most visually realistic RGB restorations because it fills void pixels using actual observed samples from the nearest valid neighbors rather than generating an averaged surface over the entire missing area. This sampling-based behavior preserves natural texture granularity and avoids the “washed-out” appearance that diffusion can introduce on cluttered indoor surfaces. However, KNN is also more sensitive to the geometry of the projection at near depth discontinuities and occlusion boundaries, the nearest valid pixels in the 2D image plane may belong to different 3D surfaces. Even small cross-surface mixing can introduce localized inconsistencies that are penalized by pixel-wise metrics, despite appearing visually plausible. Therefore, KNN can score lower numerically while still producing outputs that look more realistic to human observers.

From an application perspective, these findings imply that restoration should be selected according to the intended use. If the goal is metric-driven fidelity or stable reconstruction in larger void regions, Gaussian-weighted propagation is advantageous. If the goal is visual realism for RGB projections, KNN may be preferable because it preserves texture using real samples. This also motivates hybrid designs such as KNN for RGB realism combined with edge-aware or discrete-consistent strategies for semantic masks so that both appearance and structural integrity are maintained where they matter most.

5.2 Limitations and Challenges

A key limitation of projection-based restoration is that the achievable 2D image quality is fundamentally bounded by the completeness of the original point cloud. Many void regions in projected images are not only rasterization artifacts; they originate from true coverage gaps and sparse sampling in the original scans (Figure 7). As a result, the proposed 2D void-filling algorithms cannot reliably reconstruct missing geometry, texture, or sharp boundaries when the underlying 3D evidence is absent.

A critical challenge is the inability of 2D filling to compensate for substantial structural data loss. While the evaluated methods effectively resolve micro-voids caused by sampling sparsity and projection discretization, they degrade when confronted with macro-voids that create large-scale missing geometry due to severe occlusion, incomplete scanning trajectories, or coverage gaps (e.g., uncovered hallway sides or partially scanned storage rooms). In such cases, localized methods (such as KNN) may lack sufficient valid context, leading to stalled propagation or overly smooth continuations. As a result, KNN force completion by drawing samples from distant, unrelated surfaces, producing cross-surface mixing and distorted appearance (Figure 7 column 4). These failure modes emphasize that 2D-only restoration can interpolate appearance but cannot recover missing structure without additional geometric constraints.

Additional limitations include the strong viewpoint dependency of void patterns in multi-view projections and persistent difficulty near depth discontinuities, where neighbor-based methods may mix surfaces with blur boundaries. Finally, standard pixel-level metrics (MAE, RMSE, PSNR, SSIM) emphasize fidelity to a reference and may not fully capture perceptual realism, especially in large void regions where multiple completions may be plausible. Future work will therefore consider geometry-aware constraints (e.g., depth discontinuity handling and multi-view fusion), evaluation on additional datasets, and task-driven validation to better align restoration quality with downstream use cases.

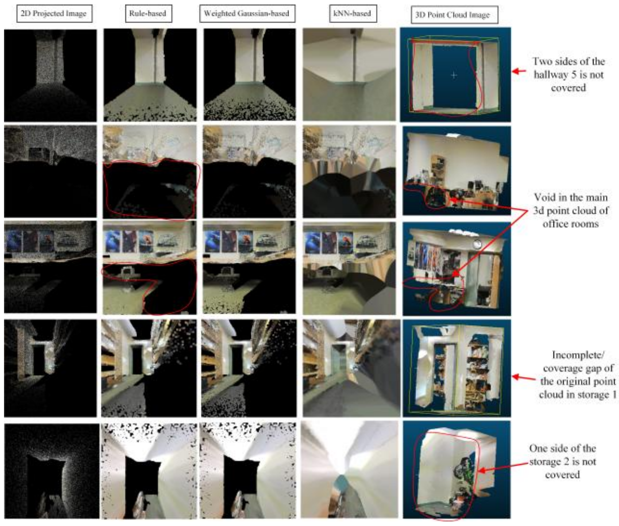

Figure 7. Void filling in projected images with reference to 3D point-cloud sparsity and missing regions

6. Conclusions and Recommendations

6.1 Conclusions

This study investigated void filling for point-cloud–derived multi-view 2D projections generated from S3DIS Area 3, where missing pixels arise from sparsity, occlusion, and incomplete sensor coverage. Three deterministic restoration strategies were evaluated named: KNN interpolation, NNRule (rule-based neighborhood filling), and Gaussian-weighted propagation by using standard image-quality metrics and qualitative visual inspection. The quantitative results show that Gaussian-weighted propagation provides the best overall restoration performance (lowest MAE/RMSE and highest PSNR/SSIM), indicating that mask-normalized diffusion produces the most stable and globally consistent reconstructions across sparse projections. At the same time, qualitative comparisons demonstrate that KNN often yields the most visually realistic RGB appearance, because it fills missing pixels using locally consistent observed samples and better preserves natural texture compared with smoothing-driven approaches. These findings highlight an important and expected trade-off in projection restoration: gaussian based methods tend to optimize pixel-level fidelity by smoothing void regions, while neighbor-based methods tend to preserve perceptual realism by sampling real local textures.

6.2 Recommendations for Practical Application

Based on the comparative evaluation of deterministic void-filling strategies, the selection of an appropriate algorithm should be explicitly guided by the requirements of the downstream application:

When the primary objective is numerical stability, structural coherence, or preparing data for semantic segmentation where generalized shapes are more critical than granular texture, Gaussian-weighted propagation is highly recommended. Its mask-normalized smoothing effectively minimizes pixel-wise error and provides the most mathematically consistent continuous fills.

On the ither hand, if the projected images are intended for human visual inspection, virtual reality, or applications demanding high-fidelity texture realism, KD-Tree-driven KNN interpolation is the recommended baseline. Despite its sensitivity to depth discontinuities, its reliance on actual observed neighbor samples prevents the artificial "washing out" of complex textures in cluttered indoor scenes.

6.3 Future Study

Future studies must extend these restoration frameworks to outdoor Terrestrial Laser Scanning (TLS) and Mobile Laser Scanning (MLS) data, where extreme spatial density gradients and vast un-sampled regions (e.g., the sky) will require the development of dynamically scaling, depth-aware search radiuses rather than the fixed-parameter windows utilized in this study. Furthermore, to address the full spectrum of data loss across these complex environments, future work will explore the development of hybrid restoration pipelines. These hybrid frameworks could theoretically employ computationally lightweight KNN interpolation to efficiently handle micro-voids and preserve valid local texture, while selectively triggering heavier generative models (like GANs) exclusively in regions where macro-voids dictate that semantic synthesis is strictly necessary.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT/MOE) (RS-2026-25484641).

References

Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., and Savarese, S., 2016: 3D Semantic Parsing of Large-Scale Indoor Spaces, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1534-1543, 10.1109/CVPR.2016.170.

Banerjee, A., 2020: SimPatch: A Nearest Neighbor Similarity Match between Image Patches. arXiv preprint arXiv:2008.03085.

Beltrán, J., Guindel, C., Moreno, F. M., Cruzado, D., García, F., and Escalera, A. D. L., 2018: BirdNet: A 3D Object Detection Framework from LiDAR Information, 21st International Conference on Intelligent Transportation Systems (ITSC), 3517-3523, 10.1109/ITSC.2018.8569311.

Cao, J., Hu, X., Cui, H., Liang, Y., and Chen, Z., 2023: A generative adversarial network model fused with a self-attention mechanism for the super-resolution reconstruction of ancient murals. IET Image Processing, 17(8), 2336-2349. https://doi.org/10.1049/ipr2.12795.

Chung, H., Kim, J., Mccann, M. T., Klasky, M. L., and Ye, J. C., 2022: Diffusion posterior sampling for general noisy inverse problems. arXiv preprint arXiv:2209.14687.

Daghigh, H., Tannant, D. D., Daghigh, V., Lichti, D. D., and Lindenbergh, R., 2022: A critical review of discontinuity plane extraction from 3D point cloud data of rock mass surfaces. Computers & Geosciences, 169, 105241. https://doi.org/10.1016/j.cageo.2022.105241.

Deng, G., 2017: Guided Wavelet Shrinkage for Edge-Aware Smoothing. IEEE Transactions on Image Processing, 26(2), 900-914. 10.1109/TIP.2016.2633941.

Ding, D., Ram, S., and Rodríguez, J. J., 2019: Image Inpainting Using Nonlocal Texture Matching and Nonlinear Filtering. IEEE Transactions on Image Processing, 28(4), 1705-1719. 10.1109/TIP.2018.2880681.

Dong, W., Zeng, L., Ji, S., and Yang, Y., 2024: Gaussian error loss function for image smoothing. Image Vision Comput., 152, 105300. https://doi.org/10.1016/j.imavis.2024.105300.

Dong, Z., Tao, J., Liu, Y., Feng, Y., Chen, Y., and Wang, Y.: Shallow Sea Bathymetric Inversion of Active–Passive Satellite Remote Sensing Data Based on Virtual Control Point Inverse Distance Weighting, 10.3390/rs17213621, 2025.

Feng, C.-B., Liu, K., Sun, J., Lai, Q., Jin, J.-P., Su, H., Wang, G., and Vong, C.-M., 2025: Unpaired to paired data synthesis and generative ensemble network for smudged image restoration. Applied Soft Computing, 183, 113648. https://doi.org/10.1016/j.asoc.2025.113648.

Giraud, R., Ta, V. T., Bugeau, A., Coupé, P., and Papadakis, N., 2017: SuperPatchMatch: An Algorithm for Robust Correspondences Using Superpixel Patches. IEEE Transactions on Image Processing, 26(8), 4068-4078. 10.1109/TIP.2017.2708504.

Gonzalez, R. C. and Woods, R. E.: Digital Image Processing, 4th, 2018.

Guo, Q., Zhang, Y., Qiu, S., and Zhang, C., 2021a: Accelerating patch-based low-rank image restoration using kd-forest and Lanczos approximation. Information Sciences, 556, 177-193. https://doi.org/10.1016/j.ins.2020.12.066.

Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., and Bennamoun, M., 2021b: Deep Learning for 3D Point Clouds: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(12), 4338-4364. 10.1109/TPAMI.2020.3005434.

Huang, S.-C., Peng, Y.-T., Chang, C.-H., Cheng, K.-H., Huang, S.-W., and Chen, B.-H., 2020: Restoration of images with high-density impulsive noise based on sparse approximation and ant-colony optimization. IEEE Access, 8, 99180-99189.

Huang, X., Zhang, J., Fan, L., Wu, Q., and Yuan, C., 2017: A Systematic Approach for Cross-Source Point Cloud Registration by Preserving Macro and Micro Structures. IEEE Transactions on Image Processing, 26(7), 3261-3276. 10.1109/TIP.2017.2695888.

Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A., 2017: Image-to-image translation with conditional adversarial networks, Proceedings of the IEEE conference on computer vision and pattern recognition, 1125-1134.

Jhaldiyal, A. and Chaudhary, N., 2022: Semantic segmentation of 3D LiDAR data using deep learning: a review of projection-based methods. Appl Intell, 53(6), 6844-6855. 10.1007/s10489-022-03930-5.

Kaur, A., Singh, H., and Arora, D., 2017: An efficient approach for image denoising based on edge-aware bilateral filter, 4th International Conference on Signal Processing, Computing and Control (ISPCC), 56-61, 10.1109/ISPCC.2017.8269649.

Lawin, F. J., Danelljan, M., Tosteberg, P., Bhat, G., Khan, F. S., and Felsberg, M., 2017: Deep projective 3D semantic segmentation, International conference on computer analysis of images and patterns, 95-107, https://doi.org/10.48550/arXiv.1705.03428.

Li, H.-a., Hu, L., and Zhang, J., 2023: Irregular mask image inpainting based on progressive generative adversarial networks. The Imaging Science Journal, 71(3), 299-312. 10.1080/13682199.2023.2180834.

Li, J., Wang, H., Li, Y., and Zhang, H.: A Comprehensive Review of Image Restoration Research Based on Diffusion Models, 10.3390/math13132079, 2025.

Liu, Y., Yin, H., Wan, J., Liu, Z., and Chong, A., 2022: Edge Aware Network for Image Dehazing. IEEE Signal Processing Letters, 29, 174-178. 10.1109/LSP.2021.3130014.

Lugmayr, A., Danelljan, M., Romero, A., Yu, F., Timofte, R., and Van Gool, L., 2022: Repaint: Inpainting using denoising diffusion probabilistic models, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 11461-11471.

Milioto, A., Vizzo, I., Behley, J., and Stachniss, C., 2019: RangeNet ++: Fast and Accurate LiDAR Semantic Segmentation, 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 4213-4220, 10.1109/IROS40897.2019.8967762.

Patil, K. and Bendre, V., 2023: A Comprehensive Survey on Image Filtering & Inpainting for Improved Image Quality, OITS International Conference on Information Technology (OCIT), 514-520, 10.1109/OCIT59427.2023.10430651.

Qi, C. R., Su, H., Mo, K., and Guibas, L. J., 2017a: Pointnet: Deep learning on point sets for 3d classification and segmentation, Proceedings of the IEEE conference on computer vision and pattern recognition, 652-660, https://doi.org/10.48550/arXiv.1612.00593.

Qi, C. R., Yi, L., Su, H., and Guibas, L. J., 2017b: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Advances in neural information processing systems, 30. https://doi.org/10.48550/arXiv.1706.02413.

Robert, D., Vallet, B., and Landrieu, L., 2022: Learning Multi-View Aggregation in the Wild for Large-Scale 3D Semantic Segmentation, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

Rong, R., Wang, S., Zhang, X., Wen, Z., Cheng, X., Jia, L., Yang, D. M., Xie, Y., Zhan, X., and Xiao, G., 2023: Enhanced Pathology Image Quality with Restore–Generative Adversarial Network. The American Journal of Pathology, 193(4), 404-416. https://doi.org/10.1016/j.ajpath.2022.12.011.

Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D. J., and Norouzi, M., 2022a: Image super-resolution via iterative refinement. IEEE transactions on pattern analysis and machine intelligence, 45(4), 4713-4726.

Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., and Norouzi, M., 2022b: Palette: Image-to-Image Diffusion Models, ACM SIGGRAPH 2022 Conference Proceedings, Vancouver, BC, Canada, 10.1145/3528233.3530757.

Singh, D. P. and Yadav, M., 2024: Deep learning-based semantic segmentation of three-dimensional point cloud: a comprehensive review. International Journal of Remote Sensing, 45(2), 532-586. 10.1080/01431161.2023.2297177.

Su, H., Maji, S., Kalogerakis, E., and Learned-Miller, E., 2015: Multi-View Convolutional Neural Networks for 3D Shape Recognition, Proceedings of the IEEE International Conference on Computer Vision (ICCV)

Tan, X., Liu, Y., Zuo, C., and Zhang, M., 2017: A real-time video denoising algorithm with FPGA implementation for Poisson–Gaussian noise. Journal of Real-Time Image Processing, 13(2), 327-343. 10.1007/s11554-014-0405-2.

Wu, B., Wan, A., Yue, X., and Keutzer, K., 2018: SqueezeSeg: Convolutional Neural Nets with Recurrent CRF for Real-Time Road-Object Segmentation from 3D LiDAR Point Cloud, 2018 IEEE International Conference on Robotics and Automation (ICRA), 1887-1893, 10.1109/ICRA.2018.8462926.

Xia, B., Zhang, Y., Wang, S., Wang, Y., Wu, X., Tian, Y., Yang, W., and Van Gool, L., 2023: Diffir: Efficient diffusion model for image restoration, Proceedings of the IEEE/CVF international conference on computer vision, 13095-13105.

Zeng, Y., Wang, H., and Zhou, X., 2026: K-nearest neighbor-enhanced Residual Learning Framework for image restoration. Engineering Applications of Artificial Intelligence, 167, 113816. https://doi.org/10.1016/j.engappai.2026.113816.

Zhang, A., Li, S., Wu, J., Li, S., and Zhang, B., 2023a: Exploring Semantic Information Extraction From Different Data Forms in 3D Point Cloud Semantic Segmentation. IEEE Access, 11, 61929-61949. 10.1109/ACCESS.2023.3287940.

Zhang, G., Ji, J., Zhang, Y., Yu, M., Jaakkola, T. S., and Chang, S., 2023b: Towards coherent image inpainting using denoising diffusion implicit models, Proceedings of the 40th International Conference on Machine Learning, 41164 - 41193.

Zhang, H., Wang, C., Tian, S., Lu, B., Zhang, L., Ning, X., and Bai, X., 2023c: Deep learning-based 3D point cloud classification: A systematic survey and outlook. Displays, 79, 102456. https://doi.org/10.1016/j.displa.2023.102456.

Zhang, Y., Zhou, Z., David, P., Yue, X., Xi, Z., Gong, B., and Foroosh, H., 2020: PolarNet: An Improved Grid Representation for Online LiDAR Point Clouds Semantic Segmentation, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)

Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A., 2017: Unpaired image-to-image translation using cycle-consistent adversarial networks, Proceedings of the IEEE international conference on computer vision, 2223-2232.