STRUCTURE-AWARE DOMAIN ADAPTIVE POINT MAMBA FOR POINT CLOUD COMPLETION

Anonymous authors
Paper under double-blind review

ABSTRACT

Unsupervised domain adaptation for point cloud completion (UDA-PCC) aims to transfer models trained on paired partial-complete source data to unlabeled target point clouds despite substantial domain shifts. Existing approaches use reconstruction, self-supervision, distillation, or feature alignment to address this challenge, but generally lack a unified structural perspective. Their patch-wise, structure-agnostic alignment may associate regions with different structural roles, causing unreliable correspondences and structural drift, while existing Euclidean alignment methods may be limited in modeling multi-level semantic relations across domains. To address these limitations, we propose StructDAMamba, a structure-aware domain adaptive Point Mamba framework that jointly models intra-domain patch dependencies and cross-domain semantic hierarchies. Specifically, Global Structure-Aware Spatial Alignment (GSSA) captures local-to-global patch relationships and aligns compact structural representations at the distribution level, reducing noise-induced mismatches without relying on explicit patch correspondences. Hyperbolic Domain Feature Aggregation (HDFA) projects source- and target-domain features into hyperbolic space for intra-domain integration and cross-domain aggregation, encouraging more transferable representations. Extensive experiments across synthetic and real-world benchmarks demonstrate that StructDAMamba consistently achieves state-of-the-art completion performance while maintaining high computational efficiency.

1 INTRODUCTION

As a fundamental form of 3D multimedia data in the visual domain, point clouds play a vital role in a wide range of real-world applications, including autonomous driving, robotic manipulation, augmented reality, and remote sensing Li et al. (2020b); Qin et al. (2023); Behley et al. (2021). Point cloud completion (PCC) aims to predict the complete geometric shape from a partial input. Previous works Tesema et al. (2023); Fei et al. (2022) have significantly improved the performance of fully-supervised point cloud completion models by utilizing paired partial-complete datasets. However, when transferring these models from source domains to target domains, their performance often degrades substantially due to severe domain shifts, which are induced by sensor-specific scanning patterns, sim-to-real noise variations, and physical occlusion distributions.

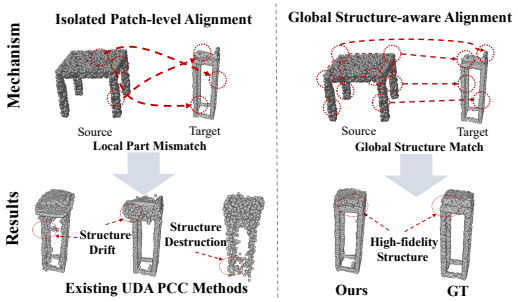


Figure 1: Existing UDA PCC methods align isolated patches while neglecting structural dependencies, leading to local mismatches and structural degradation. Our structure-aware alignment captures global patch relationships for more coherent and faithful completion.

To address this issue, unsupervised domain adaptation (UDA) techniques Wilson & Cook (2020); Liu et al. (2022) have been introduced into PCC to mitigate the domain gap between labeled source domains and unlabeled target domains. Existing UDA PCC methods can be broadly categorized into

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

several paradigms, including adversarial learning Chen et al. (2019); Zhang et al. (2021); Yang et al. (2024b); Li et al. (2025b), data reconstruction Wen et al. (2021); Liu et al. (2024), disentangled learning Gong et al. (2022), self-supervised learning Hong et al. (2023); Yang et al. (2024b); He et al. (2025), knowledge distillation He et al. (2025); Xia et al. (2025), and feature alignment Li et al. (2025c;b). These methods have significantly advanced the domain adaptability of PCC models from different perspectives, enabling more robust completion performance under cross-domain scenarios.

Despite their promising performance, existing methods lack a unified structural perspective that accounts for both intra-domain patch dependencies and cross-domain semantic hierarchies. This deficiency manifests in two critical limitations. *Firstly*, current cross-domain alignment is predominantly patch-wise and structure-agnostic. Prior approaches Li et al. (2025c;b) operate on isolated patch or token units, whether through patch-level similarity, adversarial confusion over index-level features, or MSE-based distance minimization. In unpaired UDA-PCC, however, patch indices do not imply reliable semantic or structural correspondence across domains. Such alignment may therefore pull semantically unrelated or structurally inconsistent regions together while overlooking the intra-domain dependencies that define object geometry, which can induce negative transfer and structural drift. *Secondly*, adaptation methods overlook the hierarchical semantic structure shared across domains. Beyond modeling intra-domain dependencies among patches, effective UDA-PCC should also capture the hierarchical relationships among local patches, object parts, and global shapes across domains. Yet alignment in Euclidean space models features on a flat geometry, which is poorly suited to preserving multi-level semantic relations under domain shifts. As a result, semantically corresponding structures may remain fragmented or be aggregated inconsistently across domains. This limitation motivates hyperbolic representation learning to capture hierarchical cross-domain semantics.

To address these issues, we propose StructDAMamba, a novel structure-aware domain adaptive point Mamba framework for UDA PCC. Specifically, we introduce *Global Structure-Aware Spatial Alignment* (GSSA) to capture structural dependencies among patches and achieve consistent cross-domain alignment, alleviating structural inconsistency and noise-induced mismatches. Complementarily, we develop *Hyperbolic Domain Feature Aggregation* (HDFA), which models source- and target-domain features in hyperbolic space to capture shared semantic hierarchies, thereby reducing domain-specific bias and promoting more domain-invariant and generalizable representations. Extensive experiments demonstrate that StructDAMamba consistently outperforms existing state-of-the-art methods while maintaining high computational efficiency. In summary, our contributions are three-fold:

- We propose StructDAMamba, a unified structure-aware domain adaptive Point Mamba framework for UDA-PCC. From a structural perspective, it jointly models intra-domain patch dependencies and cross-domain semantic hierarchies to learn structurally consistent and domain-invariant representations.
- We develop two complementary modules: *Global Structure-Aware Spatial Alignment* captures global patch dependencies for robust cross-domain alignment, while *Hyperbolic Domain Feature Aggregation* models shared semantic hierarchies to reduce domain-specific bias and improve generalization.
- Extensive experiments on multiple benchmark datasets demonstrate that StructDAMamba consistently achieves state-of-the-art completion performance while maintaining high computational efficiency.

2 Related Work

Domain Adaptive Point Cloud Completion. PCC is a core 3D vision task that recovers complete 3D geometry from partial inputs, laying critical geometric foundations for autonomous driving, 3D reconstruction and robotic perception Yang et al. (2024c); Chen et al. (2020). Existing PCC paradigms perform well on in-distribution data, but models based on CNNs Yuan et al. (2018); Tchapmi et al. (2019); Xiang et al. (2021) (suffering from limited receptive fields), Transformers Yu et al. (2021); Zhou et al. (2022); Yan et al. (2025) (burdened by quadratic complexity), or Mamba Zhang et al. (2025); Li et al. (2026); Fu et al. (2025) (tending to overfit domain-specific information Yang et al. (2025); Long et al. (2024)) all degrade drastically in real scenarios due to domain shifts between synthetic training data and real sensor data. To address this, Unsupervised

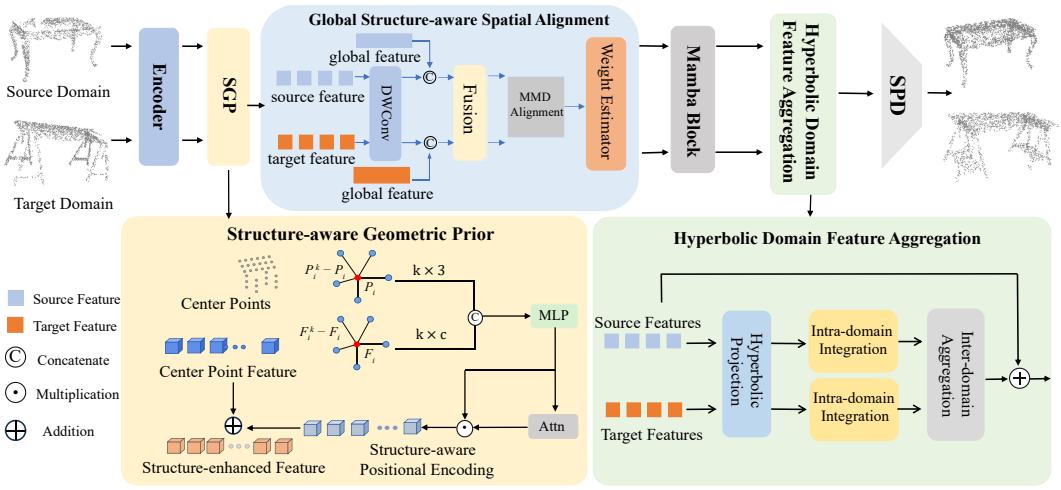


Figure 2: **Overall architecture of StructDAMamba.** Our framework comprises two core modules: (1) *Global Structure-Aware Spatial Alignment (GSSA)*: captures global patch dependencies and learns structure-aware weights for robust cross-domain alignment; (2) *Hyperbolic Domain Feature Aggregation (HDFA)*: projects Euclidean features into hyperbolic space, performs intra-domain integration and inter-domain aggregation via hyperbolic transformation.

Domain Adaptive PCC (UDA-PCC) has emerged Chen et al. (2019); Wen et al. (2021); Gong et al. (2022); Hong et al. (2023); Yang et al. (2024b); Li et al. (2025b;c). However, existing UDA-PCC methods mainly rely on isolated patch-level feature alignment, often ignoring intra-domain structural relationships among point cloud patches. In the unpaired UDA setting, locally similar patches across domains may correspond to different structural roles, and directly pulling their features closer can induce misleading correspondences and negative transfer. This coarse alignment may further lead to local mismatches, structural drift, and severe geometric distortion. Therefore, developing a structure-aware UDA-PCC framework remains a critical challenge.

Hyperbolic Learning in Point Cloud. Hyperbolic learning has attracted growing attention in 3D vision due to its natural suitability for modeling hierarchical structures, which are inherent in 3D objects where local parts at different scales compose the whole shape. This property has motivated the introduction of hyperbolic geometry into various point cloud tasks, including representation learning Liu et al. (2025); Sun et al. (2025), classification Montanaro et al. (2022); Hu et al. (2024), segmentation Sur et al. (2025); Montanaro et al. (2023); Chen et al. (2023), and reconstruction Li et al. (2025a); Leng et al. (2023). It has also shown effectiveness in downstream applications such as matching Xie et al. (2024), pose regression Wang et al. (2023), and point cloud compression Li et al. (2020a). For point cloud completion, recent studies mainly employ hyperbolic loss functions Lin et al. (2023; 2024) to alleviate the limitations of Euclidean metrics. However, hyperbolic learning remains unexplored in unsupervised domain adaptive point cloud completion. To fill this gap, we propose Hyperbolic Domain Feature Aggregation, which exploits the hierarchical modeling capability of hyperbolic space to aggregate source-domain and target-domain structural features, encouraging more generalizable representations and improving cross-domain robustness.

3 Methodology

In this section, we focus on the UDA-PCC task. This task adopts labeled source domain (paired partial-complete point clouds) and unlabeled target domain (partial point clouds) as training data, aiming to mitigate the domain shift between two domains. We propose a tailored UDA-PCC framework named StructDAMamba, which explicitly captures intrinsic structural correlations among point cloud patches. As shown in Figure 2, StructDAMamba integrates two key modules: GSSA and HDFA. GSSA enhances spatial alignment by modeling global spatial dependencies among patches, while HDFA facilitates domain fusion in hyperbolic space.

3.1 GLOBAL STRUCTURE-AWARE SPATIAL ALIGNMENT

Effective unsupervised domain adaptive point cloud completion requires modeling both cross-domain feature consistency and structural dependencies within each point cloud. In the unpaired setting, source and target patches lack reliable one-to-one semantic correspondences, so direct patch-level alignment may match unrelated regions, causing negative transfer and structural drift. To address this, we propose Global Structure-Aware Spatial Alignment (GSSA). The GSSA module models adaptive inter-patch interactions to construct compact structural representations, aligns them at the distribution level, and further derives structure-aware weights to modulate the original patch features, enabling more robust and structurally consistent domain adaptation.

Structure-aware Geometric Prior. To provide more discriminative structural priors for patch-level alignment, we introduce a Structure-aware Geometric Prior (SGP) module. Rather than relying solely on absolute spatial coordinates or independently embedded patch features, SGP explicitly characterizes the relational structure among neighboring patches by jointly modeling local geometric dependencies and feature-level semantic variations. Specifically, for the i-th center patch, we construct its K-nearest-neighbor set as $\mathcal{N}(i) = \{(p_i^k, x_i^k)\}_{k=1}^K$, where p_i^k and x_i^k denote the spatial coordinate and feature representation of the k-th neighboring patch, respectively. We then compute the relative coordinate and relative feature difference between the center patch and each neighboring patch: $\Delta p_i^k = p_i^k - p_i, \Delta x_i^k = |x_i^k - x_i|$.

The two complementary cues are concatenated to form a local relational descriptor: $h_i^k = [\Delta p_i^k; \Delta x_i^k]$, which is further mapped to a transition feature $T_i^k = \text{MLP}(h_i^k)$ by a shared MLP. To adaptively emphasize informative neighborhood structures, we employ a lightweight attention function to generate channel-wise importance weights:

$$SGP_i = \sum_{k=1}^{K} \text{attn}(T_i^k) \odot T_i^k, \tag{1}$$

where $\text{attn}(\cdot)$ is implemented by shared per-point MLPs with learnable parameters W_{attn}. The final structure-aware feature representation is obtained by injecting the learned relational prior into the original patch feature:

$$F_i = X_i + SGP_i. \tag{2}$$

By characterizing the local structural role of each patch, SGP provides relational priors for subsequent aggregation and cross-domain alignment. Given the structure-enhanced source and target features $\mathbf{F}_s, \mathbf{F}_t \in \mathbb{R}^{B\times G\times D}$, where B, G, and D denote the batch size, number of patches, and feature dimension, respectively, GSSA captures local inter-patch dependencies using a lightweight depthwise one-dimensional convolution and extracts the global structural context through average pooling:

$$\begin{aligned} Z_d^{\text{local}} &= \text{DWConv}(F_d), \qquad d \in \{s,t\}, \\ g_d &= \text{AvgPool}_G(F_d), \qquad g_d \in \mathbb{R}^{B\times 1\times D}. \end{aligned} \tag{3}$$

The global descriptor g_d is then repeated across the patch dimension and concatenated with the local structural features along the channel dimension. A lightweight fusion function integrates the two complementary representations:

$$Z_d = \phi\left(\left[Z_d^{\text{local}}; \text{Repeat}_G(g_d)\right]\right), \qquad d \in \{s,t\}. \tag{4}$$

In this way, each patch representation incorporates both local structural relations captured by DWConv and holistic object-level context aggregated from the entire point cloud, yielding compact local-global structural representations Z_s and Z_t. Since source and target samples are unpaired, we align their structural representations at the distribution level using MMD, avoiding unreliable point- or patch-wise correspondences. The MMD formulation is provided in the *supplementary material*.

The aligned compact structural bases are then fed into a learnable structure-aware weight estimator to generate adaptive patch-wise modulation weights:

$$\begin{aligned} W_{\text{structure}}^s &= \textit{Pool}\left(\text{Sigmoid}(\text{Conv}(\text{GELU}(Z_s)))\right) \in \mathbb{R}^{B\times G}, \\ W_{\text{structure}}^t &= \textit{Pool}\left(\text{Sigmoid}(\text{Conv}(\text{GELU}(Z_t)))\right) \in \mathbb{R}^{B\times G}. \end{aligned} \tag{5}$$

Here, $Pool(\cdot)$ denotes the max-pooling operation along the feature-channel dimension to aggregate structural responses into patch-level importance weights. Finally, the original features are modulated by the learned structure-aware weights through a residual formulation:

$$\tilde{F}_s = X_s \odot W^s_{\text{structure}} + X_s, \quad \tilde{F}_t = X_t \odot W^t_{\text{structure}} + X_t. \tag{6}$$

The residual modulation highlights structurally informative regions while preserving the original geometric details, preventing information loss caused by over-alignment.

3.2 Hyperbolic Domain Feature Aggregation

Effective UDA-PCC requires not only modeling structural dependencies within each point cloud, but also accounting for multi-level semantic relationships across domains. To model cross-domain semantic hierarchies, we introduce *Hyperbolic Domain Feature Aggregation* (HDFA), which projects source- and target-domain features onto a Lorentz manifold for intra-domain integration and cross-domain aggregation. The aggregated features are then mapped back to Euclidean space and residually fused with the original features, reducing domain-specific bias while retaining geometric information.

Hyperbolic Preliminaries. Point cloud features naturally exhibit hierarchical and non-Euclidean structural relations. To model such dependencies, HDFA adopts the Lorentz model, which provides numerically stable and well-defined geometric operations Chen et al. (2022). It represents an n-dimensional hyperbolic manifold with constant negative curvature $\kappa < 0$ as

$$\mathcal{L}^n_\kappa = \left\{ \mathbf{x} \in \mathbb{R}^{n+1} \mid \langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{L}} = 1/\kappa,\ x_0 > 0 \right\}, \tag{7}$$

where $\langle \mathbf{x}, \mathbf{y} \rangle_{\mathcal{L}} = -x_0 y_0 + \sum_{i=1}^{n} x_i y_i$ denotes the Lorentzian inner product. Its tangent space provides a local Euclidean approximation of the manifold, while exponential and logarithmic maps enable transformations between Euclidean features and hyperbolic representations Yang et al. (2024a). HDFA employs these mappings for hyperbolic projection and Euclidean residual fusion; the detailed implementations of intra-domain structural integration (ISI) and inter-domain structural aggregation (ISA) are provided in the supplementary material.

Given the source- and target-domain patch features $F_s, F_t \in \mathbb{R}^{B \times G \times D}$, where B, G, and D denote the batch size, number of patches, and feature dimension, respectively, HDFA projects them from the tangent space onto the Lorentz manifold:

$$H_s = \mathcal{P}(F_s), \quad H_t = \mathcal{P}(F_t), \tag{8}$$

where $\mathcal{P}(\cdot)$ denotes the exponential-map-based projection.

Intra-domain Structural Integration. Before performing cross-domain aggregation, HDFA first enhances structural consistency within each domain. Specifically, we apply a hyperbolic self-aggregation operation to the source and target features separately:

$$\bar{H}_s = \text{ISI}(H_s), \quad \bar{H}_t = \text{ISI}(H_t), \tag{9}$$

where $\text{ISI}(\cdot)$ denotes intra-domain structural integration. It captures structural relations among patches within each domain, providing domain-wise structural representations for subsequent inter-domain aggregation.

Inter-domain Structural Aggregation. After obtaining the domain-wise structural representations, HDFA aggregates target-domain structural information into the source-domain feature space. This target-to-source aggregation is designed to prevent the supervised source branch from overfitting to source-specific geometry. Formally, we compute the cross-domain aggregation as:

$$H_{t \to s} = \text{ISA}(\bar{H}_s, \bar{H}_t), \tag{10}$$

where $\text{ISA}(\cdot)$ denotes inter-domain structural aggregation. It aggregates target-domain structural information into the source representation by using source features as queries and target features as keys and values.

The aggregated hyperbolic features are then mapped back to the Euclidean space and fused with the original source features through a residual connection:

$$F'_s = F_s + \alpha \cdot \mathcal{P}^{-1}\left(H_{t \to s}\right), \tag{11}$$

where $\mathcal{P}^{-1}(\cdot)$ denotes the logarithmic map from the hyperbolic manifold to the Euclidean tangent space, and α is a learnable fusion coefficient. The residual design preserves the original geometric details while incorporating cross-domain structural cues, leading to more stable optimization and improved cross-domain transferability.

During inference, only target-domain partial point clouds are available. Therefore, HDFA degenerates into a target-domain structural aggregation module, where the query, key, and value are all derived from the target-domain features. This allows the model to retain the learned structure-aware aggregation behavior without requiring source-domain samples at test time. An ablation validating this inference strategy is provided in the supplementary material.

Optimization. Inspired by HyperCD Lin et al. (2023), we exploit the extra hyperbolic Chamfer distance (HCD) to combat noise and outliers, defined as:

$$\mathcal{L}_{\text{total}} = \sum_{i=1}^{n} \mathcal{L}_{\text{HCD}}(P_i, Q) + \lambda \mathcal{L}_{\text{align}} \tag{12}$$

where P_i denotes the i-th refined output of the adopted Snowflake Point Deconvolution (SPD) decoder Xiang et al. (2021), and Q is the ground truth. Here, n is the total number of refined outputs. $\mathcal{L}_{\text{align}}$ represents the feature alignment loss introduced by GSSA, and λ is a hyperparameter that controls the contribution of this alignment term.

4 Experiments

Dataset. Following the experimental protocols of prior UDA-PCC methods (Gong et al., 2022; Li et al., 2025b;c), we use CRN (Wang et al., 2020) as the source domain, and real-world scans, 3D-FUTURE (Fu et al., 2021), and ModelNet (Wu et al., 2015) as target domains. All models are trained separately for each category, using labeled source data and unlabeled target partial inputs. For CRN, we select 26,863 samples from categories shared with other datasets. On real-world scans, we evaluate on ScanNet (Dai et al., 2017), MatterPort3D (Chang et al., 2017), and KITTI (Geiger et al., 2012) using table, chair, and car categories respectively. We select shared categories between source and target domains, and all input scans are uniformly downsampled to 2048 points for unpaired training and inference.

Comparison Methods. We compare StructDAMamba with seven UDA-PCC methods reported on the same benchmarks: Pcl2Pcl (Chen et al., 2019), ShapeInversion (Zhang et al., 2021), Cycle4Completion (Wen et al., 2021), ACL-SPC (Hong et al., 2023), OptDE (Gong et al., 2022), DAPoinTr (Li et al., 2025b), and DAPointMamba (Li et al., 2025c). Together, these baselines cover established unpaired completion approaches and recent Transformer- and Mamba-based models, providing a comparison across different architectures.

Implementation Details. Experiments are implemented in PyTorch on a single 24GB RTX 4090 GPU. We adopt PointMamba (Liang et al., 2024) blocks as our backbone. Models are trained with a batch size of 32, an initial learning rate of 1×10^{-3}, and a weight decay of 5×10^{-2}. The alignment loss weight λ is set to 0.001. For fair comparison, we reproduced DAPointMamba using its original implementation and settings. In ablation studies, Baseline refers to StructDAMamba explicitly stripped of our proposed GSSA and HDFA modules. All reported quantitative results are averaged over three independent runs with different random seeds.

4.1 Main Results

In this section, we conduct comprehensive experimental comparisons on widely used benchmarks spanning both real-world and synthetic domains. Following standard protocols, we evaluate using Unidirectional Chamfer Distance (UCD), Unidirectional Hausdorff Distance (UHD), and Chamfer Distance (CD). We report CD on the synthetic 3D-FUTURE and ModelNet datasets, and UCD/UHD on the real-world ScanNet, MatterPort3D, and KITTI scans. Lower values indicate better completion under each evaluation protocol.

Table 1: Cross-domain completion on 3D-FUTURE and ModelNet. Chamfer Distance (CD)↓; scale 10^4. Best results are bold.

Methods	3D-FUTURE						ModelNet						
	Avg	Cabinet	Chair	Lamp	Sofa	Table	Avg	Plane	Car	Chair	Lamp	Sofa	Table
Pcl2Pcl	92.83	57.23	43.91	157.86	63.23	141.92	68.14	18.53	17.54	43.58	126.80	38.78	163.62
ShapeInv.	53.21	38.54	26.30	48.57	44.02	108.60	41.61	3.78	15.66	22.25	60.42	22.25	125.31
Cycle4Comp	45.39	32.62	34.08	77.19	43.05	40.00	28.65	5.77	11.85	26.67	83.34	22.82	21.47
ACL-SPC	35.97	70.12	23.87	31.75	28.74	25.38	34.89	5.75	11.73	43.08	106.29	25.62	16.89
OptDE	28.99	28.37	21.87	29.92	37.98	26.81	15.94	2.18	9.80	14.71	39.74	19.43	**9.75**
DAPoinTr	22.35	18.46	17.60	27.91	23.08	24.71	13.79	2.38	8.04	13.83	33.26	12.72	12.51
DAPointMamba	20.92	19.78	16.19	22.79	23.53	22.32	13.90	2.47	8.09	13.01	34.93	12.86	12.03
StructDAMamba	**19.15**	**18.34**	**14.92**	**21.52**	**21.63**	**19.34**	**12.61**	**2.17**	**7.35**	**11.82**	**32.08**	**12.13**	10.12

Results on Synthetic Benchmarks. As shown in Table 1, StructDAMamba achieves the best average CD on both 3D-FUTURE and ModelNet, reaching 19.15 and 12.61, respectively. On 3D-FUTURE, it ranks first across all five categories, improving DAPointMamba from 20.92 to 19.15 on average, with notable gains on Chair (16.19 to 14.92) and Table (22.32 to 19.34). On ModelNet, it improves DAPoinTr and DAPointMamba by 1.18 and 1.29 in average CD and ranks first in five of six categories, with clear improvements on Chair, Lamp, Plane, Car, and Sofa. The only exception is Table, where StructDAMamba ranks second at 10.12 behind OptDE at 9.75, while still achieving a substantially lower overall average than OptDE (12.61 vs. 15.94). These results demonstrate the complementary effectiveness of GSSA and HDFA: GSSA captures local-to-global patch dependencies for robust structural alignment, while HDFA integrates intra- and cross-domain structural cues in hyperbolic space, improving structural coherence and completion fidelity under domain shifts.

Table 2: Cross-domain completion results on real-world scans. [UCD↓/UHD↓] are used as the metrics, and the scale factors are 10^4 and 10^2 for UCD and UHD respectively.

Methods	ScanNet		MatterPort3D		KITTI
	Chair	Table	Chair	Table	Car
Pcl2Pcl	17.3/10.1	9.1/11.8	15.9/10.5	6.0/11.8	9.2/14.1
ShapeInv.	3.2/10.1	3.3/11.9	3.6/10.0	3.1/11.8	2.9/13.8
Cycle4Comp	5.1/6.4	3.6/5.9	8.0/8.4	4.2/6.8	3.3/5.8
ACL-SPC	1.4/4.7	1.8/5.1	1.8/4.8	2.1/4.9	2.0/4.9
OptDE	2.6/5.5	1.9/4.6	3.0/5.5	1.9/5.3	1.6/3.5
DAPoinTr	1.1/2.7	0.96/2.7	1.3/2.9	1.2/2.8	0.45/1.8
DAPointMamba	1.0/2.8	0.91/2.7	1.05/3.0	0.98/2.7	0.49/2.1
StructDAMamba	**0.92/2.7**	**0.71/2.4**	**0.91/2.7**	**0.81/2.5**	**0.35/1.8**

Results on Real-World Scan Benchmarks. As shown in Table 2, StructDAMamba achieves the lowest UCD across all categories on ScanNet, MatterPort3D, and KITTI, demonstrating robust adaptation to noisy and incomplete real-world observations. Compared with DAPointMamba, the UCD is reduced from 1.00 to 0.92 and from 0.91 to 0.71 on ScanNet, while consistent gains are obtained on MatterPort3D. These improvements benefit from GSSA, which models local-to-global patch dependencies and performs distribution-level structural alignment, thereby reducing mismatches caused by irregular sampling and occlusion. HDFA further aggregates intra-domain and cross-domain structural cues in hyperbolic space, improving geometric consistency and boundary preservation. Consequently, StructDAMamba achieves the best UHD on most categories and tied-best results on ScanNet Chair and KITTI. In particular, the KITTI UCD decreases from 0.49 to 0.35, indicating stronger generalization to sparse outdoor LiDAR scans.

4.2 Ablation Study

Effectiveness of Each Component.

As shown in Table 3, GSSA and HDFA provide complementary gains. GSSA reduces average CD from the baseline 22.68 to 20.34 by modeling inter-patch structure; HDFA reaches 20.48 by aggregating hierarchical cross-domain features. GSSA alone has the lower CD on Cabinet, Lamp, Sofa, and Table, whereas HDFA alone performs better on Chair (15.86 versus 16.02). Combining them lowers average CD to 19.15, a further 1.19 below the stronger single-module variant. Relative to the baseline, this is a reduction of 3.53 (15.6%), indicating that the gains from the individual modules remain effective when they are used together. The combined model also improves on the stronger single-module variant in every category, with the largest additional gain on Table (21.43 to 19.34). On Chair, where HDFA alone is stronger, adding GSSA further lowers CD from 15.86 to 14.92. Thus, the joint benefit is visible beyond the average: the two modules address different structural errors rather than merely duplicating one another.

Table 3: Component ablation on 3D-FUTURE (CD↓, scale 10^4). The single-module variants add GSSA or HDFA to the baseline; Ours uses both.

Category	Baseline	+GSSA	+HDFA	Ours
Avg	22.68	20.34	20.48	**19.15**
Cabinet	22.45	19.12	19.24	**18.34**
Chair	17.62	16.02	15.86	**14.92**
Lamp	24.92	22.47	22.83	**21.52**
Sofa	24.54	22.68	22.73	**21.63**
Table	23.87	21.43	21.75	**19.34**

4.3 Analysis and Visualization

Effectiveness of GSSA-based Structure-aware Spatial Alignment.

Figure 3 visualizes the structure-aware spatial alignment effect of GSSA. Before alignment, patches are scattered without explicit structural organization, making it difficult to distinguish structurally important regions. The previous spatial alignment method produces fragmented responses and tends to emphasize local patch similarities, which may introduce mismatched structural links under unpaired cross-domain settings. In contrast, GSSA generates more coherent and continuous high-response regions along key object structures, such as boundaries and supporting parts. Its geometric prior combines relative positions and feature differences among neighboring patches, while local-to-global aggregation gives each patch a broader object context. The resulting structure-aware weights reinforce informative regions through residual feature modulation, preserving local geometric details. The more connected response is consistent with the average CD reduction from 22.68 to 20.34 when GSSA is added alone in Table 3. By aligning compact structural representations at the distribution level rather than enforcing hard patch-wise correspondence, GSSA reduces local mismatches and structural drift.

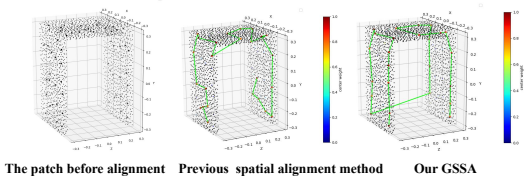


Figure 3: Comparison of spatial alignment responses produced by GSSA and the previous alignment method.

Visualization Results on 3D-FUTURE. Figure 4 presents qualitative cross-domain completion results on the 3D-FUTURE dataset, where StructDAMamba is compared with representative state-of-the-art PCC methods across diverse object categories, including cabinet, chair, sofa, lamp, and table. As highlighted by the red dashed circles, existing CNN/Transformer-based methods often suffer from severe structural degradation under domain shifts, such as incomplete local parts, distorted object boundaries, and missing or fragmented thin structures. Although the Mamba-based DA-PointMamba improves the overall shape consistency to some extent, it still tends to produce local mismatches and lose fine-grained geometric details, especially around thin structures and supporting parts. In contrast, StructDAMamba generates more faithful and structurally coherent completions. Benefiting from GSSA and HDFA, our method better preserves overall shape organization and local details, producing completions consistent with the ground truth across categories.

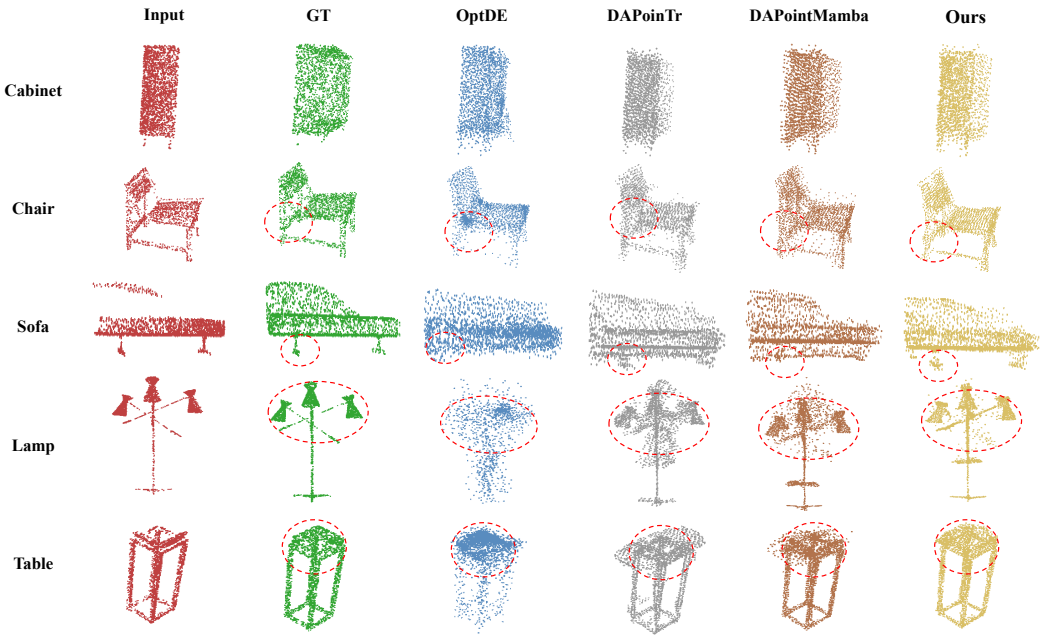


Figure 4: Qualitative comparisons of our method with state-of-the-art PCC methods on the 3D-FUTURE dataset.

t-SNE Visualization of Hyperbolic Domain Feature Aggregation.

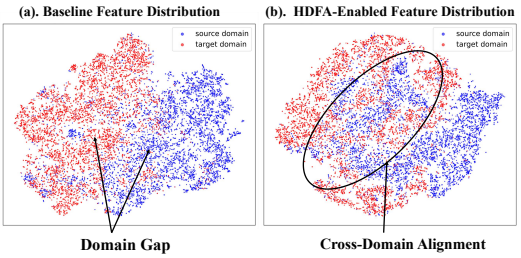


Figure 5: t-SNE visualization comparing the feature distributions of the baseline and HDFA.

Figure 5 compares source- and target-domain feature distributions with and without HDFA to assess cross-domain alignment. In Figure 5(a), the baseline without HDFA exhibits a clear separation between source- and target-domain features, indicating that the learned representations remain strongly entangled with domain-specific factors. In contrast, Figure 5(b) shows that HDFA produces more compact and overlapping feature distributions by projecting features into hyperbolic space and performing intra-domain integration and inter-domain aggregation. This hierarchical aggregation facilitates the interaction of semantically consistent structural cues across domains while suppressing domain-specific biases. Consequently, HDFA narrows the domain gap and promotes more domain-invariant representations, which contributes to more stable point cloud completion.

5 Conclusion

We propose StructDAMamba, an efficient structure-aware domain adaptive Point Mamba framework for unsupervised point cloud completion under domain shifts. Unlike existing UDA-PCC methods that mainly rely on local feature matching or distribution-level alignment, StructDAMamba explicitly models structural correlations and hierarchical geometric relationships among point cloud patches. Specifically, GSSA captures global structural dependencies for robust cross-domain spatial alignment, while HDFA leverages hyperbolic geometry to aggregate category-consistent cross-domain features and learn domain-invariant representations. Extensive experiments demonstrate state-of-the-art performance while preserving geometric coherence.

AI USE STATEMENT

The authors used generative AI tools to assist with language polishing and grammar checking. All AI-assisted content was reviewed and revised by the authors. The authors take full responsibility for the final content of the paper, including all text, claims, experiments, and code.

ETHICS STATEMENT

This work uses publicly available datasets for point cloud completion research and does not involve human subject experiments, private personal data, or sensitive attributes. We do not foresee direct negative societal impacts beyond the general dual-use considerations associated with point cloud reconstruction technologies.

REPRODUCIBILITY STATEMENT

We provide implementation details, training settings, dataset configurations, evaluation protocols, and ablation studies in the main paper and appendix. Code and configuration files will be released upon publication to facilitate reproducibility.

REFERENCES

Jens Behley, Martin Garbade, Andres Milioto, Jan Quenzel, Sven Behnke, Jürgen Gall, and Cyrill Stachniss. Towards 3d lidar-based semantic scene understanding of 3d point cloud sequences: The semantickitti dataset. *The International Journal of Robotics Research*, 40(8-9):959–967, 2021.

Angel Chang, Angela Dai, Thomas Funkhouser, Maciej Halber, Matthias Niessner, Manolis Savva, Shuran Song, Andy Zeng, and Yinda Zhang. Matterport3d: Learning from rgb-d data in indoor environments. *arXiv preprint arXiv:1709.06158*, 2017.

Bike Chen, Wei Peng, Xiaofeng Cao, and Juha Röning. Hyperbolic uncertainty aware semantic segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 25(2):1275–1290, 2023.

Siheng Chen, Baoan Liu, Chen Feng, Carlos Vallespi-Gonzalez, and Carl Wellington. 3d point cloud processing and learning for autonomous driving. *arXiv preprint arXiv:2003.00601*, 2020.

Weize Chen, Xu Han, Yankai Lin, Hexu Zhao, Zhiyuan Liu, Peng Li, Maosong Sun, and Jie Zhou. Fully hyperbolic neural networks. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 5672–5686, 2022.

Xuelin Chen, Baoquan Chen, and Niloy J Mitra. Unpaired point cloud completion on real scans using adversarial training. *arXiv preprint arXiv:1904.00069*, 2019.

Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5828–5839, 2017.

Ben Fei, Weidong Yang, Wen-Ming Chen, Zhijun Li, Yikang Li, Tao Ma, Xing Hu, and Lipeng Ma. Comprehensive review of deep learning-based 3d point cloud completion processing and analysis. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):22862–22883, 2022.

Huan Fu, Rongfei Jia, Lin Gao, Mingming Gong, Binqiang Zhao, Steve Maybank, and Dacheng Tao. 3d-future: 3d furniture shape with texture. *International Journal of Computer Vision*, 129 (12):3313–3337, 2021.

Zhiheng Fu, Jiehua Zhang, Longguang Wang, Lian Xu, Hamid Laga, Yulan Guo, Farid Boussaid, and Mohammed Bennamoun. Completionmamba: Taming state space model for point cloud completion. *IEEE Transactions on Image Processing*, 2025.

Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3354–3361, 2012.

Jingyu Gong, Fengqi Liu, Jiachen Xu, Min Wang, Xin Tan, Zhizhong Zhang, Ran Yi, Haichuan Song, Yuan Xie, and Lizhuang Ma. Optimization over disentangled encoding: Unsupervised cross-domain point cloud completion via occlusion factor manipulation. In *European Conference on Computer Vision*, pp. 517–533, 2022.

Xing He, Zhe Zhu, Liangliang Nan, Honghua Chen, Jing Qin, and Mingqiang Wei. Pointsfda: Source-free domain adaptation for point cloud completion. *arXiv preprint arXiv:2503.15144*, 2025.

Sangmin Hong, Mohsen Yavartanoo, Reyhaneh Neshatavar, and Kyoung Mu Lee. Acl-spc: Adaptive closed-loop system for self-supervised point cloud completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9435–9444, 2023.

Naiwen Hu, Haozhe Cheng, Yifan Xie, Pengcheng Shi, and Jihua Zhu. Hyperbolic image-and-pointcloud contrastive learning for 3d classification. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4973–4979, 2024.

Zhiying Leng, Shun-Cheng Wu, Mahdi Saleh, Antonio Montanaro, Hao Yu, Yin Wang, Nassir Navab, Xiaohui Liang, and Federico Tombari. Dynamic hyperbolic attention network for fine hand-object reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 14894–14904, 2023.

Li Li, Zhu Li, Shan Liu, and Houqiang Li. Rate control for video-based point cloud compression. *IEEE Transactions on Image Processing*, 29:6237–6250, 2020a.

Wenrui Li, Zhe Yang, Wei Han, Hengyu Man, Xingtao Wang, and Xiaopeng Fan. Hyperbolic-constraint point cloud reconstruction from single rgb-d images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 4959–4967, 2025a.

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 Transactions on Neural Networks and Learning Systems*, 32(8):3412–3432, 2020b.

Yinghui Li, Qianyu Zhou, Jingyu Gong, Ye Zhu, Richard Dazeley, Xinkui Zhao, and Xuequan Lu. Dapointr: Domain adaptive point transformer for point cloud completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 5066–5074, 2025b.

Yinghui Li, Qianyu Zhou, Di Shao, Hao Yang, Ye Zhu, Richard Dazeley, and Xuequan Lu. Dapointmamba: Domain adaptive point mamba for point cloud completion. *arXiv preprint arXiv:2511.20278*, 2025c.

Yixuan Li, Lipeng Ma, Weidong Yang, and Ben Fei. 3dmambacomplete: Structured state space model for high-efficiency point cloud completion. *ACM Transactions on Multimedia Computing, Communications and Applications*, 22(1):1–24, 2026.

Dingkang Liang, Xin Zhou, Wei Xu, Xingkui Zhu, Zhikang Zou, Xiaoqing Ye, Xiao Tan, and Xiang Bai. Pointmamba: A simple state space model for point cloud analysis. *Advances in Neural Information Processing Systems*, 37:32653–32677, 2024.

Fangzhou Lin, Yun Yue, Songlin Hou, Xuechu Yu, Yajun Xu, Kazunori D Yamada, and Ziming Zhang. Hyperbolic chamfer distance for point cloud completion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 14595–14606, 2023.

Fangzhou Lin, Songlin Hou, Haotian Liu, Shang Gao, Kazunori D Yamada, Haichong K Zhang, and Ziming Zhang. Hyperbolic chamfer distance for point cloud completion and beyond. *arXiv preprint arXiv:2412.17951*, 2024.

Fengqi Liu, Jingyu Gong, Qianyu Zhou, Xuequan Lu, Ran Yi, Yuan Xie, and Lizhuang Ma. Cloud-mix: Dual mixup consistency for unpaired point cloud completion. *IEEE Transactions on Visualization and Computer Graphics*, 31(4):2182–2195, 2024.

Xiaofeng Liu, Chaehwa Yoo, Fangxu Xing, Hyejin Oh, Georges El Fakhri, Je-Won Kang, and Jonghye Woo. Deep unsupervised domain adaptation: A review of recent advances and perspectives. *arXiv preprint arXiv:2208.07422*, 2022.

Yingjie Liu, Pengyu Zhang, Ziyao He, Mingsong Chen, Xuan Tang, and Xian Wei. Hyperbolic contrastive learning for hierarchical 3d point cloud embedding. *arXiv preprint arXiv:2501.02285*, 2025.

Shaocong Long, Qianyu Zhou, Xiangtai Li, Xuequan Lu, Chenhao Ying, Yuan Luo, Lizhuang Ma, and Shuicheng Yan. Dgmamba: Domain generalization via generalized state space model. In *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 3607–3616, 2024.

Antonio Montanaro, Diego Valsesia, and Enrico Magli. Rethinking the compositionality of point clouds through regularization in the hyperbolic space. *Advances in Neural Information Processing Systems*, 35:33741–33753, 2022.

Antonio Montanaro, Diego Valsesia, and Enrico Magli. Towards hyperbolic regularizers for point cloud part segmentation. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1–5, 2023.

Yuzhe Qin, Binghao Huang, Zhao-Heng Yin, Hao Su, and Xiaolong Wang. Dexpoint: Generalizable point cloud reinforcement learning for sim-to-real dexterous manipulation. In *Conference on Robot Learning*, pp. 594–605, 2023.

Yiding Sun, Haozhe Cheng, Chaoyi Lu, Zhengqiao Li, Minghong Wu, Huimin Lu, and Jihua Zhu. Hyperpoint: Multimodal 3d foundation model in hyperbolic space. *Pattern Recognition*, pp. 112800, 2025.

Tanuj Sur, Samrat Mukherjee, Kaizer Rahaman, Subhasis Chaudhuri, Muhammad Haris Khan, and Biplab Banerjee. Hyperbolic uncertainty-aware few-shot incremental point cloud segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11810–11821, 2025.

Lyne P Tchapmi, Vineet Kosaraju, Hamid Rezatofighi, Ian Reid, and Silvio Savarese. Topnet: Structural point cloud decoder. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 383–392, 2019.

Keneni W Tesema, Lyndon Hill, Mark W Jones, Muneeb I Ahmad, and Gary KL Tam. Point cloud completion: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 30(10): 6880–6899, 2023.

Sijie Wang, Qiyu Kang, Rui She, Wei Wang, Kai Zhao, Yang Song, and Wee Peng Tay. Hypliloc: Towards effective lidar pose regression with hyperbolic fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5176–5185, 2023.

Xiaogang Wang, Marcelo H Ang Jr, and Gim Hee Lee. Cascaded refinement network for point cloud completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 790–799, 2020.

Xin Wen, Zhizhong Han, Yan-Pei Cao, Pengfei Wan, Wen Zheng, and Yu-Shen Liu. Cycle4completion: Unpaired point cloud completion using cycle transformation with missing region coding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13080–13089, 2021.

Garrett Wilson and Diane J Cook. A survey of unsupervised deep domain adaptation. *ACM Transactions on Intelligent Systems and Technology*, 11(5):1–46, 2020.

Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1912–1920, 2015.

Zhiqian Xia, Haifeng Xia, Shichao Jin, Wei Wang, Zhengming Ding, and Xiaochun Cao. Dspf: Dual-stage preservation and fusion for source-free domain adaptive point cloud completion. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pp. 8428–8437, 2025.

Peng Xiang, Xin Wen, Yu-Shen Liu, Yan-Pei Cao, Pengfei Wan, Wen Zheng, and Zhizhong Han. Snowflakenet: Point cloud completion by snowflake point deconvolution with skip-transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 5499–5509, 2021.

Yifan Xie, Jihua Zhu, Shiqi Li, Naiwen Hu, and Pengcheng Shi. Hecpg: Hyperbolic embedding and confident patch-guided network for point cloud matching. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–12, 2024.

Hongyu Yan, Zijun Li, Kunming Luo, Li Lu, and Ping Tan. Symmcompletion: High-fidelity and high-consistency point cloud completion with symmetry guidance. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 9094–9102, 2025.

Hao Yang, Qianyu Zhou, Haijia Sun, Xiangtai Li, Fengqi Liu, Xuequan Lu, Lizhuang Ma, and Shuicheng Yan. Pointdgmamba: Domain generalization of point cloud classification via generalized state space model. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 9193–9201, 2025.

Menglin Yang, Harshit Verma, Delvin Ce Zhang, Jiahong Liu, Irwin King, and Rex Ying. Hypformer: Exploring efficient transformer fully in hyperbolic space. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3770–3781, 2024a.

Yunseo Yang, Jihun Kim, and Kuk-Jin Yoon. Syn-to-real domain adaptation for point cloud completion via part-based approach. In *European Conference on Computer Vision*, pp. 179–197, 2024b.

Zetong Yang, Li Chen, Yanan Sun, and Hongyang Li. Visual point cloud forecasting enables scalable autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 14673–14684, 2024c.

Xumin Yu, Yongming Rao, Ziyi Wang, Zuyan Liu, Jiwen Lu, and Jie Zhou. Pointr: Diverse point cloud completion with geometry-aware transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12498–12507, 2021.

Wentao Yuan, Tejas Khot, David Held, Christoph Mertz, and Martial Hebert. Pcn: Point completion network. In *International Conference on 3D Vision*, pp. 728–737, 2018.

Junzhe Zhang, Xinyi Chen, Zhongang Cai, Liang Pan, Haiyu Zhao, Shuai Yi, Chai Kiat Yeo, Bo Dai, and Chen Change Loy. Unsupervised 3d shape completion through gan inversion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1768–1777, 2021.

Lirui Zhang, Zhengkai Zhao, Zhi Zuo, Pan Gao, and Jie Qin. Simba: Towards high-fidelity and geometrically-consistent point cloud completion via transformation diffusion. *arXiv preprint arXiv:2511.16161*, 2025.

Haoran Zhou, Yun Cao, Wenqing Chu, Junwei Zhu, Tong Lu, Ying Tai, and Chengjie Wang. Seedformer: Patch seeds based point cloud completion with upsample transformer. In *European Conference on Computer Vision*, pp. 416–432, 2022.

A Supplementary Overview

This supplementary material provides additional implementation and optimization details, an expanded description of the proposed HDFA module, and further experimental analyses for Struct-DAMamba. Its organization is summarized as follows:

- **Section B: Additional Implementation and Optimization Details**
  - **Section B.1:** Experimental settings and training configurations.
  - **Section B.2:** Detailed reconstruction objectives, including Chamfer Distance and Hyperbolic Chamfer Distance.
  - **Section B.3:** Distribution-level structural alignment based on the Maximum Mean Discrepancy objective.
- **Section C: Detailed Architecture of HDFA**
  - **Section C.1:** Lorentz-model preliminaries and hyperbolic feature projection.
  - **Section C.2:** Detailed formulation of the hyperbolic attention mechanism.
  - **Section C.3:** Intra-domain structural integration and inter-domain structural aggregation.
- **Section D: Additional Experimental Results and Analyses**
  - **Section D.1:** Additional ablation studies, inference analysis, and computational-efficiency comparisons.
  - **Section D.2:** Additional qualitative evaluations on KITTI, MatterPort3D, ScanNet, and ModelNet.

B Additional Implementation and Optimization Details

B.1 Experimental Settings

All experiments are implemented in PyTorch and conducted on a single NVIDIA RTX 4090 GPU with 24 GB memory. The input point clouds are uniformly downsampled to 2,048 points, and the batch size is set to 32. For the Structure-aware Geometric Prior module, the number of nearest neighbors is set to $K = 8$. In GSSA, the bottleneck ratio of the depthwise one-dimensional convolution is set to $r = 8$. For HDFA, the curvature parameter κ of the Lorentz manifold is learnable and jointly optimized with the remaining network parameters. We optimize the network using AdamW with an initial learning rate of 1×10^{-3} and a weight decay of 5×10^{-2}. A cosine learning-rate scheduler is adopted for a total of 100 training epochs. The weight of the alignment loss is set to $\lambda = 0.001$. Unless otherwise specified, all quantitative results reported in the main paper and supplementary material are averaged over three independent runs with different random seeds.

B.2 Reconstruction Objectives

To further clarify the optimization objective, we provide the detailed formulations of the Chamfer Distance (CD) and Hyperbolic Chamfer Distance (HCD) used for point cloud reconstruction.

$$\begin{aligned}\mathcal{L}_{\mathrm{CD}}(X,Y) = &\frac{1}{|X|}\sum_{x\in X}\min_{y\in Y}\|x-y\|_2^2 \\ &+\frac{1}{|Y|}\sum_{y\in Y}\min_{x\in X}\|x-y\|_2^2, \\ \mathcal{L}_{\mathrm{HCD}}(X,Y) = &\ \operatorname{arcosh}\left(1+\mathcal{L}_{\mathrm{CD}}(X,Y)\right).\end{aligned} \tag{13}$$

where X and Y denote the predicted and ground-truth point sets, respectively, and $\|x-y\|_2^2$ denotes the squared Euclidean distance between points x and y.

B.3 Distribution-Level Structural Alignment

To further clarify the optimization of GSSA, we provide the detailed formulation of the MMD-based alignment objective used to match compact structural representations across domains. Since source and target samples are unpaired in UDA-PCC, their visible regions and local semantics are not necessarily consistent. Enforcing direct point-wise or patch-wise correspondence may therefore introduce misleading matches and negative transfer. To avoid such hard local alignment, we align the compact structural bases Z_s and Z_t at the distribution level using a Maximum Mean Discrepancy (MMD) loss. In this way, GSSA encourages global structural consistency across domains while reducing the risk of semantic mismatches. The MMD loss is formulated as:

$$\begin{aligned}\mathcal{L}_{\text{align}} = &\frac{1}{n_s^2}\sum_{i=1}^{n_s}\sum_{j=1}^{n_s}\kappa(z_i^s, z_j^s) + \frac{1}{n_t^2}\sum_{i=1}^{n_t}\sum_{j=1}^{n_t}\kappa(z_i^t, z_j^t) \\ &- \frac{2}{n_s n_t}\sum_{i=1}^{n_s}\sum_{j=1}^{n_t}\kappa(z_i^s, z_j^t).\end{aligned} \tag{14}$$

Here, n_s and n_t denote the numbers of source- and target-domain structural representations used for MMD computation, respectively. Moreover, z_i^s and z_j^t denote the i-th source-domain and j-th target-domain structural representations, respectively, and $\kappa(\cdot,\cdot)$ denotes the kernel function. In our implementation, we employ a Gaussian radial basis function (RBF) kernel with multiple bandwidths.

C Detailed Architecture of HDFA

In this section, we provide the implementation details of the proposed HDFA. We first introduce the necessary preliminaries of hyperbolic geometry under the Lorentz model. Then, we describe the core component of HDFA, namely hyperbolic attention. Finally, we instantiate this attention mechanism into two structural modeling modules: Inter-domain Structural Aggregation and Intra-domain Structural Integration.

C.1 Lorentz Model and Hyperbolic Projection

Point cloud representations usually contain hierarchical and non-Euclidean structural dependencies. To better preserve such structures, HDFA projects Euclidean features into a hyperbolic space. Specifically, we adopt the Lorentz model due to its numerical stability and well-defined geometric operations(Chen et al., 2022). This model is an n-dimensional Riemannian manifold $\mathbb{L}_\kappa^n = (\mathcal{L}^n, g_{\mathbf{x}}^\kappa)$ with constant negative curvature $\kappa < 0$. The manifold $\mathcal{L}_\kappa^n$ corresponds to the upper sheet of a hyperboloid in $(n+1)$-dimensional Minkowski space:

$$\mathcal{L}^n = \{\mathbf{x} \in \mathbb{R}^{n+1} : \langle \mathbf{x}, \mathbf{x}\rangle_{\mathcal{L}} = 1/\kappa,\ \mathbf{x}_0 > 0\}, \tag{15}$$

where $\langle\cdot,\cdot\rangle\mathcal{L}$ denotes the Lorentzian inner product:

$$\langle \mathbf{x}, \mathbf{y}\rangle_{\mathcal{L}} = -x_0 y_0 + \sum_{i=1}^{D} x_i y_i. \tag{16}$$

Here, x_0 is the temporal component, while the remaining components correspond to the spatial coordinates.

At any point $x \in \mathbb{L}_k^n$, the tangent space$\mathcal{T}_x\mathbb{L}_\kappa^n$ represents a local linear approximation to the hyperboloid and is defined by:

$$\mathcal{T}_x\mathbb{L}_\kappa^n = \{\mathbf{y} \in \mathbb{R}^{n+1} : \langle \mathbf{y}, \mathbf{x}\rangle_{\mathcal{L}} = 0\}. \tag{17}$$

This tangent space is essentially a Euclidean subspace embedded in $\mathbb{R}^{n+1}$ and is crucial for performing optimization and mapping operations in hyperbolic neural networks(Yang et al., 2024a).

Operations such as neural network training in hyperbolic spaces require seamless transitions between hyperbolic manifolds and their tangent spaces. This transition is achieved via exponential and

logarithmic maps, essential mathematical tools defined in hyperbolic geometry. The exponential map exp_x^κ transfers a vector from the tangent space at point x to the hyperbolic manifold, mathematically defined as:

$$\exp_{\mathbf{x}}^{\kappa}(\mathbf{v}) = \cosh(\alpha)\mathbf{x} + \sinh(\alpha)\frac{\mathbf{v}}{\alpha}, \quad \alpha = \sqrt{-\kappa\langle \mathbf{v}, \mathbf{v}\rangle_{\mathcal{L}}}. \tag{18}$$

Conversely, the logarithmic map log_x^κ projects points from the manifold back onto the tangent space:

$$\log_x^{\kappa}(\mathbf{y}) = \frac{\operatorname{arcosh}(\beta)}{\sqrt{\beta^2 - 1}}(\mathbf{y} - \beta\mathbf{x}), \quad \beta = \kappa\langle \mathbf{x}, \mathbf{y}\rangle_{\mathcal{L}}. \tag{19}$$

These mappings enable efficient geometric manipulations necessary for representation learning and model optimization within hyperbolic spaces, particularly in tasks involving hierarchical data structures such 3D point clouds.

To better model hierarchical and complex structural relationships, we project the extracted Euclidean features onto a hyperbolic manifold. To map Euclidean features onto hyperbolic manifold $\mathbb{L}_\kappa^D$ we first associate each Euclidean feature vector $x_t \in \mathbb{R}^D$ with a vector on the tangent space $\mathcal{T}_0\mathbb{L}_\kappa^D$ at the manifold origin $\mathbf{o} = (\sqrt{-1/\kappa}, 0, \ldots, 0)$:

$$\mathbf{x}_t^{(0)} = [0, \mathbf{x}_t] \in \mathcal{T}_0\mathbb{L}_\kappa^D. \tag{20}$$

Then, we use the exponential map: $\mathbf{x}_t^{\perp} = \exp_{\mathbf{o}}^{\kappa}\left(\mathbf{x}_t^{(0)}\right)$ at the origin 0 to project $x_t^{(0)}$ onto the hyperbolic manifold $\mathbb{L}_\kappa^D$. Consequently, the Euclidean features from the previous stage are transformed into a hyperbolic representation $\mathbf{X}^{\mathbb{L}} = [\mathbf{x}_1^{\mathbb{L}}, \mathbf{x}_2^{\mathbb{L}}, \ldots, \mathbf{x}_T^{\mathbb{L}}] \in \mathbb{L}_\kappa^{B\times T\times(D+1)}$.

C.2 Hyperbolic Attention

The core mechanism of HDFA is a hyperbolic attention operation defined on the Lorentz manifold. Different from standard Euclidean attention, HDFA performs feature projection, similarity estimation, aggregation, and refinement while preserving the manifold constraint.

Given hyperbolic input features, we first use a curvature-aware hyperbolic transformation with curvatures, denoted as HTC(Yang et al., 2024a), to obtain query, key, and value embeddings.Specifically, given a Lorentzian representation $\mathbf{X}^{\mathbb{L}} \in \mathbb{L}_\kappa^{B\times T\times(D+1)}$, HTC is formulated as:

$$\text{HTC}(\mathbf{x}_t^{\mathbb{L}}; \mathbf{W}_h, \kappa_1, \kappa_2) = \left(\sqrt{\frac{\kappa_1}{\kappa_2}\|\mathbf{x}_t^{\mathbb{L}}\mathbf{W}_h\|_2^2} - \frac{1}{\kappa_2}, \sqrt{\frac{\kappa_1}{\kappa_2}}\mathbf{x}_t^{\mathbb{L}}\mathbf{W}_h\right), \tag{21}$$

where $\mathbf{W}_h$ are learnable parameters per attention head, and κ_1, κ_2 are the curvatures before and after transformation.

For each attention head h, the hyperbolic query, key, and value features are computed as:

$$\begin{aligned} \mathbf{Q}_h &= \text{HTC}(\mathbf{X}'; \mathbf{W}_h^Q), \\ \mathbf{K}_h &= \text{HTC}(\mathbf{X}'; \mathbf{W}_h^K), \\ \mathbf{V}_h &= \text{HTC}(\mathbf{X}'; \mathbf{W}_h^V), \end{aligned} \tag{22}$$

where $\mathbf{Q}_h, \mathbf{K}_h, \mathbf{V}_h \in \mathbb{L}_\kappa^{B\times T\times D_h}$, $D_h = (D+1)/H$, and H is the number of heads. Attention weights are computed via negative Lorentzian distances using the inner product followed by softmax:

$$\mathbf{A}_h = \text{softmax}\left(\frac{2 + 2\langle \mathbf{Q}_h, \mathbf{K}_h\rangle_{\mathcal{L}}}{\sqrt{D_h}} + b_h\right), \tag{23}$$

where b_h is a learnable scalar bias. The weighted values are aggregated via the Lorentzian midpoint:

$$\mathbf{A}_h \odot^{\kappa} \mathbf{V}_h := \frac{\mathbf{A}_h\mathbf{V}_h}{\sqrt{\kappa\|\mathbf{A}_h\mathbf{V}_h\|_{\mathcal{L}}}}, \tag{24}$$

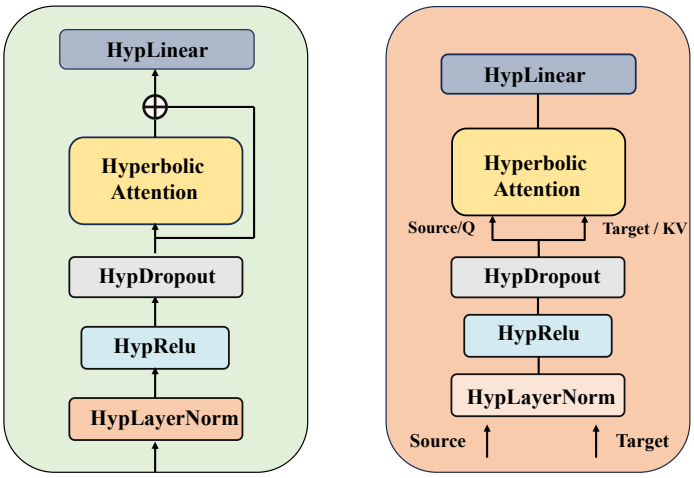


Figure 6: Detailed architectures of intra-domain structural integration and inter-domain structural aggregation.

denoted as $\mathbf{O}_h$. Finally, concatenating all heads yields

$$\mathbf{O} = [\mathbf{O}_1 \| \mathbf{O}_2 \| \dots \| \mathbf{O}_H] \in \mathbb{L}_\kappa^{B\times T\times (D+1)}. \tag{25}$$

In addition to the above multi-head attention operation, we further introduce hyperbolic counterparts of standard components including layer normalization, activation, and dropout, so as to enhance feature stability and representation capacity. We integrate these operations into the HRC module (Yang et al., 2024a), which performs curvature-aware refinement directly within the Lorentzian manifold. Specifically, the refining function $f_r(\cdot)$ encapsulates hyperbolic layer normalization $f_{\text{LayerNorm}}(\cdot)$, hyperbolic activation $f_\sigma(\cdot)$, and hyperbolic dropout $f_{\text{Dropout}}(\cdot)$, as formally defined as follows:

$$\begin{aligned} \text{HypLayerNorm}(\mathbf{O}) &= \text{HRC}(\mathbf{O}, f_{\text{LayerNorm}}), \\ \text{HypReLU}(\mathbf{O}) &= \text{HRC}(\mathbf{O}, f_\sigma), \\ \text{HypDropout}(\mathbf{O}) &= \text{HRC}(\mathbf{O}, f_{\text{Dropout}}). \end{aligned} \tag{26}$$

The output of HRC is a refined hyperbolic embedding $\hat{X} = \text{HRC}(\mathbf{O}; f_r, \kappa_1, \kappa_2) \in \mathbb{L}_\kappa^{B\times T\times (D+1)}$ suitable for downstream tasks, where κ_1, κ_2 represent the curvatures before and after the transformation.Specifically, the HRC module is defined as follows:

$$\text{HRC}(\mathbf{O}; f_r, \kappa_1, \kappa_2) = \left(\sqrt{\frac{\kappa_1}{\kappa_2}} \| f_r(\mathbf{O}) \|_2^2 - \frac{1}{\kappa_2}, \sqrt{\frac{\kappa_1}{\kappa_2}} f_r(\mathbf{O}) \right) \tag{27}$$

C.3 Intra- and Inter-Domain Structural Modeling

Overall design. Based on the hyperbolic attention mechanism introduced above, HDFA models structural dependencies at both the intra-domain and inter-domain levels. Given the source- and target-domain Euclidean features, HDFA first projects them onto the Lorentz manifold, where hierarchical geometric relationships can be represented more effectively than in a flat Euclidean space. It then performs intra-domain structural integration independently for each domain, followed by inter-domain structural aggregation between the resulting domain-specific representations. These two operations play complementary roles: the former improves structural consistency within each individual point cloud, while the latter transfers shared structural cues across domains and suppresses domain-specific feature biases.

Intra-domain structural integration. Let the hyperbolic representation of domain $d \in \{s, t\}$ be denoted as $\mathbf{H}_d \in \mathbb{L}_\kappa^{B\times T\times (D+1)}$. HDFA first applies hyperbolic self-attention to the source and target

domains separately. Specifically, the query, key, and value are all generated from the same domain representation, such that $\bar{\mathbf{H}}_s = \mathrm{ISI}(\mathbf{H}_s)$ and $\bar{\mathbf{H}}_t = \mathrm{ISI}(\mathbf{H}_t)$. This self-aggregation allows each patch to interact with other patches within the same point cloud and capture their contextual and hierarchical relationships. As a result, local patch features are enhanced with domain-wise structural context before cross-domain interaction, providing more coherent and discriminative structural representations for the subsequent aggregation stage.

Inter-domain structural aggregation. After obtaining the domain-specific hyperbolic representations, HDFA further aggregates target-domain structural information into the source-domain feature space. In contrast to intra-domain integration, the source-domain representation is used to generate the queries, whereas the target-domain representation provides the keys and values, which can be written as $\mathbf{H}_{t\to s} = \mathrm{ISA}(\bar{\mathbf{H}}_s, \bar{\mathbf{H}}_t)$. This target-to-source interaction encourages the supervised source branch to incorporate structural patterns observed in the unlabeled target domain, thereby reducing its dependence on source-specific geometry and promoting more transferable representations. The aggregated hyperbolic features are subsequently mapped back to the Euclidean tangent space and residually fused with the original features, which preserves fine-grained geometric information while introducing cross-domain structural cues. During inference, when source samples are unavailable, the same aggregation mechanism operates using target-domain features as the query, key, and value, allowing the model to retain its learned structure-aware aggregation behavior.

D Additional Experimental Results and Analyses

D.1 Further Analysis of Ablation Results

Table 4: HDFA interactions (left; S/T denotes source/target features) and RTX 4090 efficiency (right; parameters in M, FLOPs in G, and batch-one time in ms for 2,048 points).

HDFA interaction ablation (CD↓)					*Computational efficiency*			
Method	Train	Test	3D-FUTURE	ModelNet	Model	Params	FLOPs	Time
w/o inference self	$S \leftarrow T$	None	19.34	12.82	DAPoinTr	36.77	18.92	28.69
Self-only	$S \leftarrow S$	$T \leftarrow T$	20.32	13.04	DAPointMamba	9.77	5.20	5.64
Full model	$S \leftarrow T$	$T \leftarrow T$	**19.15**	**12.61**	StructDAMamba	10.17	5.23	5.93

Inference and efficiency. Table 4 separates the effects of cross-domain aggregation during training and target self-interaction at inference. Removing inference self-interaction raises average CD from 19.15 to 19.34 on 3D-FUTURE and from 12.61 to 12.82 on ModelNet. Replacing cross-domain training interaction with source self-interaction further raises CD to 20.32 and 13.04, respectively, despite retaining target self-interaction at test time. Thus, the inference operation helps, but does not replace the cross-domain training signal. The full model uses 10.17M parameters and 5.23G FLOPs, compared with 36.77M and 18.92G for DAPoinTr, and runs in 5.93 rather than 28.69 ms. These differences correspond to approximately 72% fewer parameters and FLOPs and 79% lower latency than DAPoinTr. Relative to DAPointMamba, it adds 0.40M parameters, 0.03G FLOPs, and 0.29 ms, while reducing average CD on both datasets (Table 1).

Table 5: HDFA component and GSSA setting ablations (average CD↓, scale 10^4). Lower is better.

Dataset	HDFA components			GSSA bottleneck			SGP neighborhood		
	Euclidean	w/o ISI	Full	$r = 4$	$r = 8$	$r = 16$	$k = 4$	$k = 8$	$k = 16$
3D-FUTURE	21.03	20.56	**19.15**	20.08	**19.15**	19.72	19.78	**19.15**	19.46
ModelNet	14.27	12.98	**12.61**	12.82	**12.61**	13.12	13.04	**12.61**	12.96

Further module ablations. Table 5 isolates HDFA's geometry and intra-domain integration, and GSSA's bottleneck ratio r and neighborhood size k. Euclidean HDFA yields CD values of 21.03 and 14.27 on 3D-FUTURE and ModelNet, while removing intra-domain integration gives 20.56 and 12.98; both exceed the full model's 19.15 and 12.61. Relative to the full model, Euclidean aggregation incurs larger penalties than removing ISI on both datasets: 1.88 versus 1.41 on 3D-FUTURE, and 1.66 versus 0.37 on ModelNet. This indicates a stronger contribution from hyperbolic geometry in these ablations, with an additional benefit from intra-domain integration. For GSSA, changing r from 8 to either 4 or 16 increases CD on both datasets. The same pattern holds when changing k from 8 to either 4 or 16. On 3D-FUTURE, the smaller settings cause larger penalties than the larger ones: $r = 4$ adds 0.93 versus 0.57 for $r = 16$, while $k = 4$ adds 0.63 versus 0.31 for $k = 16$. These tests support the selected settings within the evaluated ranges.

D.2 Additional Qualitative Results

Visualization Results on KITTI, MatterPort3D, and ScanNet. Figure 7 and Figure 8 show the qualitative completion results on real-world scan benchmarks, including indoor chair/table samples from ScanNet (Dai et al., 2017) and MatterPort3D (Chang et al., 2017), and outdoor car instances from KITTI (Geiger et al., 2012). Compared with synthetic data, these real scans present tougher challenges: inherent sensor noise, irregular occlusion, non-uniform point density, and large domain shifts from the source domain. Existing state-of-the-art methods (Li et al., 2025c;b; Wen et al., 2021; Gong et al., 2022) lack a structure-aware alignment strategy; their isolated point- or patch-level matching paradigm ignores the intrinsic topological dependencies of point clouds, leading to structural distortion, detail loss, local mismatches, and structural drift under real-world domain shifts. In contrast, StructDAMamba consistently generates high-fidelity complete point clouds. Benefiting from GSSA and HDFA, our method mitigates noise-induced misalignment, preserves holistic object structures, and reconstructs fine-grained geometric details. **Visualization Results on ModelNet.** Figure 9 presents qualitative results on ModelNet. Under severe domain shifts, baseline methods such as OptDE struggle to preserve global geometric integrity, resulting in structural distortion and fragmented geometries. Although DAPoinTr and DAPointMamba alleviate these issues to some extent, they still suffer from local feature mismatches and topological misalignments. By contrast, StructDAMamba yields complete, high-fidelity point clouds with precise topology and fine-grained details. These visual comparisons demonstrate that our structure-aware design effectively mitigates cross-domain misalignment and structural drift.

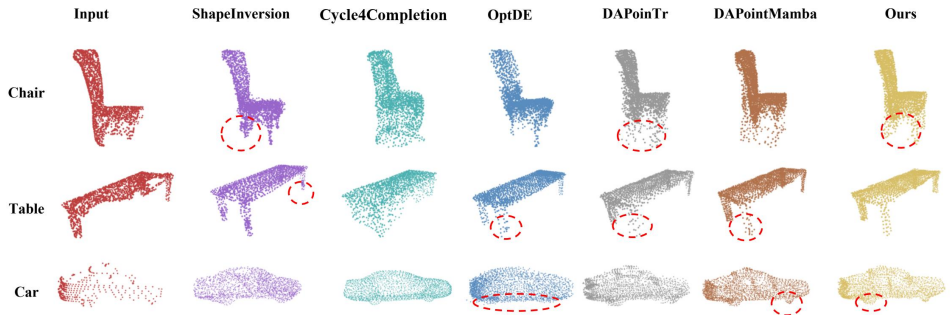


Figure 7: Visualization comparisons between ours and state-of-the-art PCC methods on KITTI and MatterPort3D.

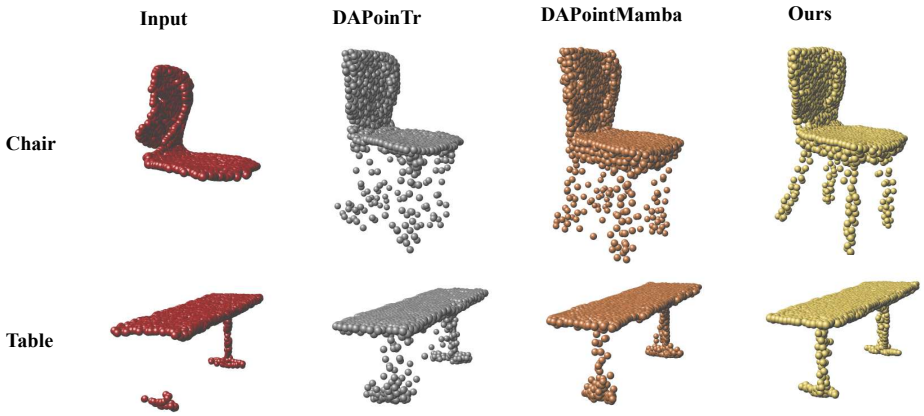


Figure 8: Visualization comparisons between ours and state-of-the-art PCC methods on ScanNet.

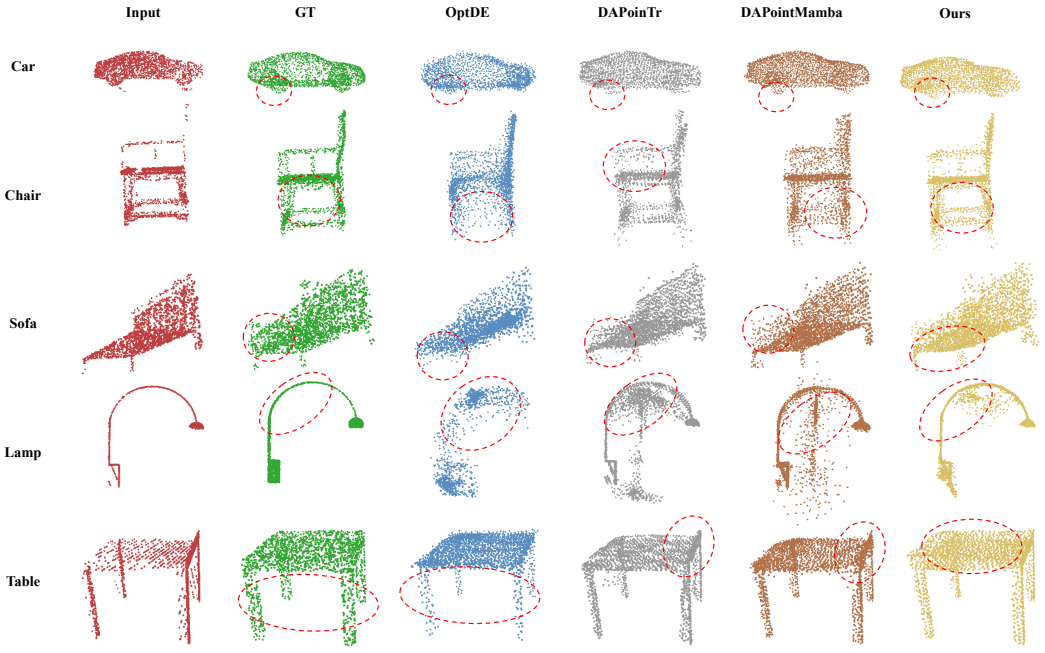


Figure 9: Visualization comparisons between ours and state-of-the-art PCC methods on ModelNet.