

DAPointMamba: Domain Adaptive Point Mamba for Point Cloud Completion

Yinghui Li^{1*}, Qianyu Zhou^{2*}, Di Shao³, Hao Yang⁴, Ye Zhu^{1†}, Richard Dazeley¹,
Xuequan Lu^{5†}

¹School of Information Technology, Deakin University, Australia

²College of Computer Science and Technology, Jilin University, Jilin, China

³Department of New Networks, PengCheng Laboratory, Shenzhen, China

⁴School of Computer Science, Shanghai Jiao Tong University, Shanghai, China

⁵Department of Computer Science and Software Engineering, The University of Western Australia, Australia
s218161821@deakin.edu.au, zhouqianyu@jlu.edu.cn, shaod@pcl.ac.cn, yanghao.cs@sjtu.edu.cn, ye.zhu@deakin.edu.au,
richard.dazeley@deakin.edu.au, bruce.lu@uwa.edu.au

Abstract

Domain adaptive point cloud completion (DA PCC) aims to narrow the geometric and semantic discrepancies between the labeled source and unlabeled target domains. Existing methods either suffer from limited receptive fields or quadratic complexity due to using CNNs or vision Transformers. In this paper, we present the first work that studies the adaptability of State Space Models (SSMs) in DA PCC and find that directly applying SSMs to DA PCC will encounter several challenges: directly serializing 3D point clouds into 1D sequences often disrupts the spatial topology and local geometric features of the target domain. Besides, the overlook of designs in the learning domain-agnostic representations hinders the adaptation performance. To address these issues, we propose a novel framework, DAPointMamba for DA PCC, that exhibits strong adaptability across domains and has the advantages of global receptive fields and efficient linear complexity. It has three novel modules. In particular, Cross-Domain Patch-Level Scanning introduces patch-level geometric correspondences, enabling effective local alignment. Cross-Domain Spatial SSM Alignment further strengthens spatial consistency by modulating patch features based on cross-domain similarity, effectively mitigating fine-grained structural discrepancies. Cross-Domain Channel SSM Alignment actively addresses global semantic gaps by interleaving and aligning feature channels. Extensive experiments on both synthetic and real-world benchmarks demonstrate that our DAPointMamba outperforms state-of-the-art methods with less computational complexity and inference latency.

Code — <https://github.com/Yinghui-Li-New/DAMamba>

Introduction

3D point cloud completion (PCC) is a fundamental task in 3D vision, essential for various real-world applications such as autonomous driving, robotics, and virtual reality. Its objective is to recover complete shapes from partial input point clouds accurately. While recent supervised methods (Tesema et al. 2023; Fei et al. 2022; Yu et al. 2021;

*These authors contributed equally.

†Corresponding Authors.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Wang et al. 2024a) have achieved great progress, their performance often drops significantly when directly applied to unseen datasets due to substantial distribution shifts induced by different sensors, shapes, and *etc.*, across domains.

To address this issue, unsupervised domain adaptation (UDA) techniques are introduced to PCC to narrow the domain shifts between the labeled source domain and the unlabeled target domain. The existing UDA PCC methods can be mainly categorized into several types: adversarial learning (Chen, Chen, and Mitra 2019; Zhang et al. 2021; Yang, Kim, and Yoon 2024), data reconstruction (Wen et al. 2021; Liu et al. 2024a), disentangled learning (Gong et al. 2022), and self-supervised learning (Hong et al. 2023; Yang, Kim, and Yoon 2024). Nevertheless, most of them are built upon convolution-based backbones (Figure 1(a)) and suffer from limited receptive fields and struggle to model global geometric structures during the learning of domain-invariant features for adaptation. Recently, DAPoinTr (Li et al. 2025b) introduces the Transformer architecture (Yu et al. 2021) in UDA PCC to align the sequence-wise features at the global level and the local level. However, the quadratic complexity of the attention mechanism (Figure 1(b)) drastically leads to unacceptable computational inefficiency and memory overhead, especially for long sequence patches.

Recently, Mamba, one of emerging State Space Models (SSMs), is capable of modeling long-range dependencies with a global receptive field (Gu and Dao 2023; Liu et al. 2024b). Moreover, Mamba stands out due to its remarkable efficiency, achieving linear computational complexity compared to Transformer-based methods. Recent studies (Liang et al. 2024; Han et al. 2024) have validated Mamba’s efficacy in point cloud tasks, demonstrating strong capabilities for efficient sequence modeling. Despite its gratifying progress, it is non-trivial to directly apply existing Mamba-based approaches to UDA PCC due to the lack of consideration of domain shifts and tailored designs. Therefore, how to improve the transferability of Mamba-based models in UDA PCC still remains a very critical and open problem. In this paper, our goal is to enhance the transferability of Mamba-like models across domains in UDA point cloud completion. Our motivations mainly lie in two aspects. Firstly, directly

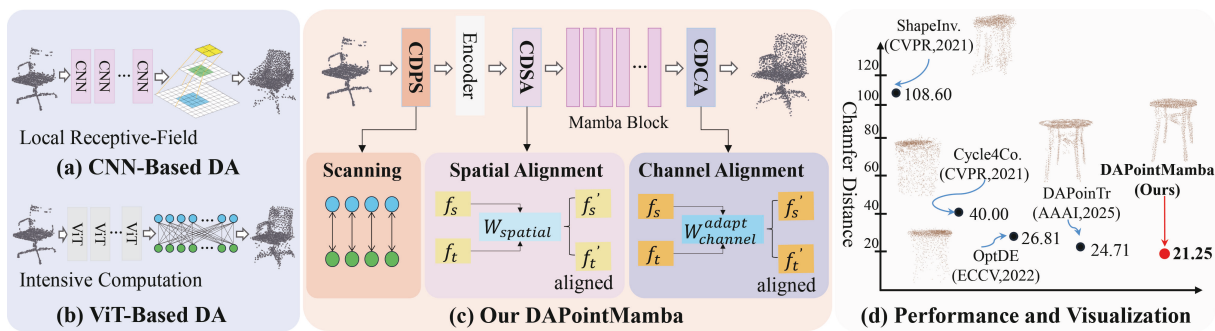


Figure 1: **Left:** Previous domain adaptive point cloud completion works suffer from limited receptive fields (a) or quadratic complexity (b) due to using CNNs or vision Transformers. **Middle:** In contrast, we propose a novel framework, DAPointMamba (c), that consists of Cross-Domain Patch-Level Scanning (CDPS), Cross-Domain Spatial SSM Alignment (CDSA), and Cross-Domain Channel SSM Alignment (CDCA). Our model exhibits strong adaptability across domains and has the advantages of global receptive fields and efficient linear complexity. **Right:** Visualization of the Table category from widely-used 3D-FUTURE benchmark, and our DAPointMamba demonstrates superior domain adaptability (lower Chamfer Distance is better) against state-of-the-art methods.

serializing sparse, unstructured 3D point clouds into 1D sequences often disrupts the spatial topology and local geometric features of point clouds, undermining the completion performance in the target domain. For instance, a rule-based scanning strategy may miss critical geometric local details under specific viewpoints, and spatial relationships between points may not be preserved after serialization. Secondly, current Point Mamba architectures lack tailored designs for learning domain-invariant features, making them sensitive to domain shifts and consequently resulting in poor completion results in target domains.

In this paper, we propose DAPointMamba, the first Mamba-based UDA framework (Figure 1(c)) explicitly tailored for domain adaptive point cloud completion. Our DAPointMamba has the advantages of strong transferability toward the target domain, linear computational complexity, and a global receptive field in completion point clouds. Our DAPointMamba comprises three key components. Firstly, Cross-Domain Patch-Level Scanning (CDPS) is introduced to explicitly ensure spatial correspondence across domains by partitioning incomplete point clouds into aligned local patches using shared coordinate normalization and Z-order serialization. Secondly, Cross-Domain Spatial SSM Alignment (CDSA) is presented to explicitly address fine-grained spatial misalignment at the patch level through similarity-guided modulation, reinforcing local feature alignment and effectively mitigating structural discrepancies. Finally, Cross-Domain Channel SSM Alignment (CDCA) is designed to tackle global semantic misalignment by adaptively mixing and modulating feature channels from both source and target domains. Extensive experiments are conducted on diverse benchmarks, including real-world scans (KITTI, ScanNet, and MatterPort3D) and synthetic datasets (3D-FUTURE and ModelNet), showing the effectiveness of DAPointMamba.

In summary, our contributions are three-fold:

- We propose DAPointMamba, a novel Mamba-based framework, for UDA point cloud completion. Our frame-

work exhibits strong adaptability across domains and advantages of global receptive fields and linear complexity.

- We design Cross-Domain Patch-Level Scanning to ensure patch-level geometric correspondences, and Cross-Domain Spatial SSM Alignment and Cross-Domain Channel SSM Alignment to enforce fine-grained spatial alignment and global semantic alignment.
- Extensive experiments with visualization on both synthetic and real-world datasets consistently confirm the superiority of our approach to state-of-the-art methods in terms of performance and computational efficiency.

Related Work

Domain Adaptive Point Cloud Completion

Point cloud completion (PCC) aims to generate complete shapes from partial input. Early CNN-based methods (Yuan et al. 2018; Groueix et al. 2018; Tesema et al. 2023; Fei et al. 2022) used folding decoders and convolutional architectures to generate coarse shapes. Later advancements (Tchapmi et al. 2019; Xiang et al. 2021) refined the output resolution and structural detail with hierarchical or progressive generation strategies. Transformer-based models (Yu et al. 2021; Zhou et al. 2022) improve global feature modeling but suffer from a high computational cost. Unfortunately, PCC models often degrade when transferred across different domains. UDA has thus been explored to bridge this gap via adversarial learning (Chen, Chen, and Mitra 2019; Zhang et al. 2021), data reconstruction (Wen et al. 2021; Liu et al. 2024a), disentangled learning (Gong et al. 2022), and self-supervised learning (Hong et al. 2023; Yang, Kim, and Yoon 2024). Nonetheless, existing methods either suffer from limited receptive fields or quadratic complexity due to using CNNs or Transformers. Thus, it is an urgent need to investigate UDA PCC with global modeling, linear complexity, and high transferability toward the target domain.

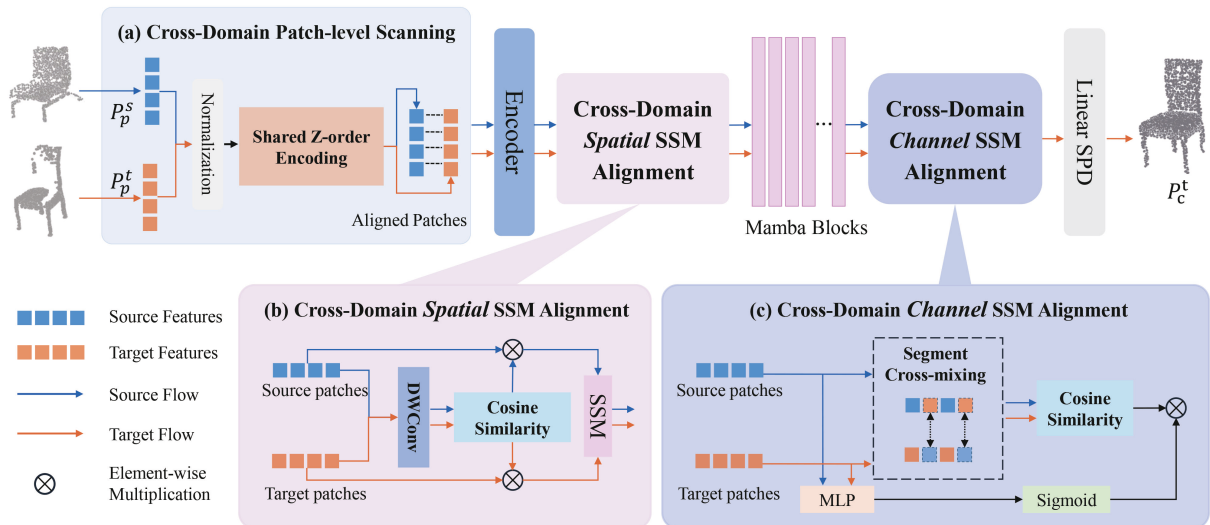


Figure 2: The framework of DAPointMamba for cross-domain point cloud completion, including three key components: (a) Cross-Domain Patch-Level Scanning is designed to close domain shifts by creating spatial correspondence at the patch level. (b) Cross-Domain Spatial SSM Alignment is proposed to solve fine-grained spatial discrepancies across domains. (c) Cross-Domain Channel SSM Alignment is presented to address the semantic structure of feature channels through cross-domain channel mixing and similarity-based modulation.

State Space Models (SSMs)

Mamba, as well as State space models (SSMs) (Gu and Dao 2023), have recently garnered increasing attention due to their strong capabilities in modeling long-range dependencies with linear computational complexity. VMamba and Vision Mamba (Liu et al. 2024b; Zhu et al. 2024) introduce Mamba architecture in vision tasks. Recently, a series of works has been inspired to study Mamba in point cloud analysis, where the representative works are PointMamba, Point Cloud Mamba, Mamba3D, *etc* (Liang et al. 2024; Zhang et al. 2025; Han et al. 2024; Bahri et al. 2025; Wang et al. 2024b; Li et al. 2025a). In the completion task, 3DMambaComplete (Li, Yang, and Fei 2024) incorporates Mamba blocks to enhance local details and global structure, while it also incorporates cross-attention mechanisms and is not a pure Mamba design. To the best of our knowledge, this is the first work that studies Mamba’s adaptability in UDA PCC.

Methodology

In this section, we present DAPointMamba, a novel Mamba-based architecture specifically designed for domain-adaptive point cloud completion. DAPointMamba demonstrates superior adaptability to unseen domains while maintaining the advantages of global receptive fields and efficient linear complexity. As illustrated in Figure 2, our framework comprises three key components: Cross-Domain Patch-Level Scanning (CDPS), Cross-Domain Spatial SSM Alignment (CDSA), and Cross-Domain Channel SSM Alignment (CDCA). Specifically, CDPS is designed to spatially align point patches, mitigating domain discrepancies between source and target domains during the serialization process. Furthermore, CDSA explicitly models and strengthens spatial correspondences across domain-specific features at the

patch level, effectively aligning spatial information. Complementarily, CDCA addresses global semantic misalignment by adaptively mixing and modulating feature channels across domains, enabling robust and consistent alignment of both spatial and semantic representations between source and target domains.

Cross-Domain Patch-Level Scanning

We have noticed that cross-domain point cloud completion faces severe geometric inconsistencies between source and target domains, where direct serialization of sparse, unstructured 3D point clouds into 1D sequences often disrupts spatial topology and undermines local geometric features. Recent works (Wei, Gu, and Sun 2022, 2024) have demonstrated that local geometric structures and patch-level context are crucial for both capturing fine-grained details and achieving robust adaptation. However, most existing methods (Liang et al. 2024; Wu et al. 2024) rely on global or static grouping strategies, which ignore domain-specific variations and fail to guarantee spatial correspondence across domains during serialization.

To address this, we propose a cross-domain patch-level scanning strategy with shared Z-order curve serialization named *Cross-Domain Patch-Level Scanning* (CDPS). Unlike heuristic global scanning (Liang et al. 2024; Wu et al. 2024), CDPS ensures that local patches in the source domain and target domain are spatially aligned, thereby facilitating more effective feature adaptation across domains. Specifically, given a sample $X_s, X_t \in \mathbb{R}^{B \times N \times 3}$, from source and target domains, respectively, we first normalize and compute the shared minimum coordinate across both domains:

$$C_{min} = \min(\min(X_s, \dim = 1), \min(X_t, \dim = 1)) \quad (1)$$

Each normalized point cloud is then discretized into a shared grid space and serialized into a 1D sequence using a consistent Z-order curve encoding, which maps 3D coordinates into a 1D sequence while maintaining spatial proximity:

$$G_s = [(X_s - C_{min}) * scale], \quad G_t = [(X_t - C_{min}) * scale] \quad (2)$$

$$Z_s = \text{Z-order}(G_s), Z_t = \text{Z-order}(G_t) \quad (3)$$

We sort the points from both domains according to their Z-order values and partition the sequence into G patches, each containing K points:

$$\begin{aligned} X_s &= \{P_s^{(1)}, P_s^{(2)}, \dots, P_s^{(G)}\}, \\ X_t &= \{P_t^{(1)}, P_t^{(2)}, \dots, P_t^{(G)}\} \end{aligned} \quad (4)$$

where $P_s^{(g)}, P_t^{(g)} \in \mathbb{R}^{K \times 3}$ denote the g -th patch in the source and target domains, respectively. Due to the unified normalization and Z-order serialization, the g -th patch in both domains corresponds to the same region, enabling precise patch-level alignment.

This patch-level alignment guarantees that each pair of patches consistently covers an identical spatial region across domains, thus providing a reliable basis for patch-level feature alignment and cross-domain adaptation. CDPS explicitly preserves spatial consistency during serialization, fundamentally mitigating the mismatch issues inherent in traditional global or purely sequential approaches.

Cross-Domain Spatial SSM Alignment

While the CDPS module effectively aligns feature distributions at the patch level, residual fine-grained discrepancies within corresponding local patches persist, limiting the precise alignment of local geometric details and ultimately hindering the domain adaptability to unseen domains.

To solve this problem, we introduce the *Cross-Domain Spatial SSM Alignment* (CDSA) module. The CDSA effectively reinforces the alignment of spatially corresponding local patches while preserving structural diversity in regions with significant geometric variation. Specifically, given patch-level features $X_s, X_t \in \mathbb{R}^{B \times D \times G}$, where B is the batch size, D is the feature dimension, and G is the number of patches, the Spatial SSM module operates as follows:

$$\mathcal{D}_s = \text{DWConv}(X_s), \quad \mathcal{D}_t = \text{DWConv}(X_t) \quad (5)$$

where $\text{DWConv}(\cdot)$ denotes a depthwise 1D convolution applied along the patch dimension. For each patch g in the batch, we compute the cosine similarity between the convolved features to obtain the space similarity weight $\mathcal{W}_{\text{spatial}}$, which can be represented as $\mathcal{W}_{\text{spatial}} = \cos(\mathcal{D}_s, \mathcal{D}_t)$. Finally, we modulate the patch features accordingly:

$$\tilde{X}_s = X_s \odot \mathcal{W}_{\text{spatial}}, \quad \tilde{X}_t = X_t \odot \mathcal{W}_{\text{spatial}} \quad (6)$$

where \tilde{X}_s and \tilde{X}_t denote the patch features of the source and target domains, respectively. To explicitly encourage local feature consistency across domains, we further compute a mean squared error (MSE) loss between \tilde{X}_s and \tilde{X}_t :

$$\mathcal{L}_{sp} = \frac{1}{BDG} \sum_{b=1}^B \sum_{d=1}^D \sum_{g=1}^G \left(\tilde{X}_{s,b,d,g} - \tilde{X}_{t,b,d,g} \right)^2 \quad (7)$$

This design significantly improves the model’s capacity to capture fine-grained local details by effectively modulating patch features based on cross-domain spatial similarities, enhancing its ability to generalize to unseen domains.

Cross-Domain Channel SSM Alignment

Despite effective local alignments from CDSA, global semantic inconsistencies across domains remain a critical challenge due to significant domain gaps. Meanwhile, Mamba-based architectures have demonstrated impressive efficiency and scalability for long-sequence modeling with linear computational complexity. However, recent studies (Liang et al. 2024; Han et al. 2024) indicate that Mamba’s inherently one-dimensional and causal formulation limits its capability to capture complex global interactions across feature channels, particularly when dealing with unordered and high-dimensional point cloud data. Consequently, Mamba-based backbones often exhibit insufficient global feature interaction and suboptimal performance in addressing cross-domain semantic discrepancies, especially under severe distribution shifts between source and target domains.

To mitigate these limitations, we propose the *Cross-Domain Channel SSM Alignment* (CDCA), which explicitly enhances global feature consistency at the channel level across domains. As illustrated in Figure 2 (c), we first compute the global feature for each domain:

$$g_s = \frac{1}{G} \sum_{g=1}^G X_s \in \mathbb{R}^{B \times D}, \quad g_t = \frac{1}{G} \sum_{g=1}^G X_t \in \mathbb{R}^{B \times D} \quad (8)$$

Then, the global features are concatenated and passed through a learnable alignment strength estimator:

$$\alpha = \text{Sigmoid}(\text{MLP}([g_s, g_t])) \in \mathbb{R}^{B \times 1} \quad (9)$$

The feature channels are first divided into S non-overlapping segments for both the source and target domain representations, X_s and X_t , respectively.

To reduce the domain gap and encourage information exchange, we construct mixed channel representations $X_{s,mix}$ and $X_{t,mix}$ by interleaving segments from both domains:

$$\begin{aligned} X_{s,mix} &= [X_s^{(1)}, X_t^{(2)}, X_s^{(3)}, X_t^{(4)}, \dots], \\ X_{t,mix} &= [X_t^{(1)}, X_s^{(2)}, X_t^{(3)}, X_s^{(4)}, \dots] \end{aligned} \quad (10)$$

We then compute the cosine similarity between the mixed representations to derive a channel-wise alignment weight, which can be represented as:

$$\mathcal{W}_{\text{channel}} = \cos(X_{s,mix}, X_{t,mix}) \in \mathbb{R}^{B \times G} \quad (11)$$

And the similarity weights $\mathcal{W}_{\text{channel}}$ are modulated by the estimated alignment strength α to generate adaptive similarity $\tilde{\mathcal{W}}_{\text{channel}}$. Finally, the original features are modulated by the adaptive similarity weights $\tilde{\mathcal{W}}_{\text{channel}}$:

$$\tilde{F}_s = X_s \odot \tilde{\mathcal{W}}_{\text{channel}}, \quad \tilde{F}_t = X_t \odot \tilde{\mathcal{W}}_{\text{channel}} \quad (12)$$

To explicitly regularize global feature alignment, we minimize the mean squared error between the modulated \tilde{F}_s and

Methods	Avg	Cabinet	Chair	Lamp	Sofa	Table
Pcl2Pcl	92.83	57.23	43.91	157.86	63.23	141.92
ShapeInv.	53.21	38.54	26.30	48.57	44.02	108.60
Cycle4Comp	45.39	32.62	34.08	77.19	43.05	40.00
ACL-SPC	35.97	70.12	23.87	31.75	28.74	25.38
OptDE	28.99	28.37	21.87	29.92	37.98	26.81
DAPoinTr	22.35	18.46	17.60	27.91	23.08	24.71
DAPointMamba	20.40	19.35	16.21	22.81	22.38	21.25

Table 1: Cross-domain completion results on 3D-FUTURE. Chamfer Distance (CD) \downarrow is used as the metric, and the scale factor is 10^4 . Lower is better.

\tilde{F}_t from source and target features, respectively:

$$\mathcal{L}_{ch} = \frac{1}{BDG} \sum_{b=1}^B \sum_{d=1}^D \sum_{g=1}^G \left(\tilde{F}_{s,b,d,g} - \tilde{F}_{t,b,d,g} \right)^2 \quad (13)$$

By explicitly modeling and aligning global channel-wise information, the Channel SSM compensates for the inherent limitations of the Mamba backbone in global context aggregation. This design ensures that high-level semantic structures remain consistent across domains, which is crucial for robust domain adaptation point cloud completion. Therefore, the total loss function can be formulated as:

$$\mathcal{L}_{total} = \mathcal{L}_{CD} + \lambda \mathcal{L}_{sp} + \beta \mathcal{L}_{ch} \quad (14)$$

Experiments

Experimental Settings

Datasets. Following previous standard protocol in UDA PCC (Chen, Chen, and Mitra 2019; Li et al. 2025b), we evaluate the effectiveness of our proposed DAPointMamba and compare it with state-of-the-art models on three commonly used datasets. We use 3D data from CRN (Wang, Ang Jr, and Lee 2020) as our source domain, and the datasets, including Real-World Scans, 3D-FUTURE (Fu et al. 2021), and ModelNet (Wu et al. 2015) as our target domains. For the CRN dataset, we use 26,863 samples drawn from categories shared between CRN and the other datasets for DA. For the real-world scan datasets, we evaluate our model on ScanNet (Dai et al. 2017), MatterPort3D (Chang et al. 2017), and KITTI (Geiger, Lenz, and Urtasun 2012). We select shared categories between source and target domains, and all input scans are uniformly downsampled to 2048 points for unpaired training and inference.

Implementation Details. All experiments are conducted on an RTX 4090 with 64GB RAM. We use refinement modules of PointMamba (Liang et al. 2024) as the backbone of our UDA PCC network. For the training, we employ an initial learning rate of 1×10^{-3} and a weight decay of 5×10^{-2} . The batch size is set to 32. To balance losses, weights of λ and β are set to 0.1 and 0.1, respectively. Following prior works, we adopt the Unidirectional Chamfer Distance (UCD), Unidirectional Hausdorff Distance (UHD), and Chamfer Distance (CD) as evaluation metrics.

Methods	Avg	Plane	Car	Chair	Lamp	Sofa	Table
Pcl2Pcl	68.14	18.53	17.54	43.58	126.80	38.78	163.62
ShapeInv.	41.61	3.78	15.66	22.25	60.42	22.25	125.31
Cycle4Comp	28.65	5.77	11.85	26.67	83.34	22.82	21.47
ACL-SPC	34.89	5.75	11.73	43.08	106.29	25.62	16.89
OptDE	15.94	2.18	9.80	14.71	39.74	19.43	9.75
DAPoinTr	13.79	2.38	8.04	13.83	33.26	12.72	12.51
DAPointMamba	13.11	2.30	7.58	13.15	32.04	12.48	11.08

Table 2: Cross-domain completion results on ModelNet. Chamfer Distance (CD) \downarrow is used as the metric, and the scale factor is 10^4 . Lower is better.

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	0.95/2.8	0.91/2.6	1.1/3.0	1.0/3.0	0.40/2.1

Table 3: Cross-domain completion results on real-world scans. [UCD \downarrow /UHD \downarrow] are used as the metrics, and the scale factor is 10^4 and 10^2 for UCD (Unidirectional Chamfer Distance) and UHD (Unidirectional Hausdorff Distance), respectively. A lower value of UCD or UHD is better.

Main Results

In this section, we conduct comprehensive experimental comparisons on widely used benchmarks spanning both real-world and synthetic domains. Specifically, we evaluate our method on real-world scans from KITTI, ScanNet, and MatterPort3D, as well as synthetic datasets including 3D-FUTURE and ModelNet, to thoroughly assess its performance across diverse scenarios.

Comparison Methods. We compare our method with state-of-the-art approaches in UDA PCC. These include Pcl2Pcl (Chen, Chen, and Mitra 2019), ShapeInversion (Zhang et al. 2021), Cycle4Completion (Zhang et al. 2021), OptDE (Gong et al. 2022), ACL-SPC (Hong et al. 2023) and DAPoinTr (Li et al. 2025b).

Results on 3D-FUTURE benchmark. As shown in Table 1, our proposed method DAPointMamba achieves consistent and significant improvements over the state-of-the-art UDA PCC methods on the 3D-FUTURE benchmark. Compared with the second-best method, DAPoinTr, our approach achieves an average improvement of 1.95 (from 22.35 to **20.40**) in Chamfer Distance (CD) across all categories. Notably, our model brings substantial improvements in challenging categories such as lamp and table, with remarkable margins (see the 5th and 7th columns in Table 1). Although the result on the cabinet is slightly lower than DAPoinTr, the overall performance proves that our model can better align feature distributions between source and target domains, while effectively leveraging the capacity of Mamba

Methods	CDPS	CDSA	CDCA	Cabinet	Chair	Lamp	Sofa	Table	Avg
Baseline				23.14	17.21	26.68	25.71	24.16	23.38
DAPointMamba	✓			20.06	17.02	25.37	23.35	22.85	21.73
	✓	✓		19.68	16.45	24.29	22.80	22.61	21.17
	✓	✓	✓	19.35	16.21	22.81	22.38	21.25	20.40

Table 4: Ablation studies on each proposed module of our DAPointMamba on 3D-FUTURE dataset. Chamfer Distance (CD) \downarrow is used as the metric, and the scale factor is 10^4 . Lower is better.

to capture both local and global geometric structures.

Results on ModelNet benchmark. As presented in Table 2, our DAPointMamba demonstrates superior average performance compared to the state-of-the-art methods. Specifically, DAPointMamba achieves the best performance across all categories compared with the current best model, DAPoinTr. This superiority can be attributed to our innovative cross-domain alignment designs, which explicitly address local geometric inconsistencies and global semantic variations. Our innovative designs can effectively mitigate domain shifts and narrow domain gaps at the patch and global level. Such consistent improvements highlight the robustness and domain adaptability of our proposed architecture when tackling complex UDA PCC tasks.

Results on Real-World Scans benchmark. Table 3 presents quantitative results on three widely used real-world scan datasets: ScanNet, MatterPort3D, and KITTI. We evaluate performance using both UCD and UHD, which respectively measure overall geometric fidelity and worst-case point-wise deviation. As shown in Table 3, DAPointMamba consistently outperforms all prior methods across all datasets in terms of the UCD metric, indicating our approach can preserve global shape structures and align feature distributions across domains. While our UHD performance is slightly inferior to previous methods, we consider that this is due to the nature of the UHD metric, which emphasizes the maximum point-wise error and is highly sensitive to isolated boundary outliers. Since DAPointMamba is designed to optimize holistic shape prediction through consistent local-global alignment, rather than minimizing extreme outliers, it is naturally more aligned with the UCD objective.

Ablation Studies

In this section, we analyze the effects of the designed components, including the Cross-Domain Patch-Level Scanning (CDPS), Cross-Domain Spatial SSM Alignment (CDSA), and Cross-Domain Channel SSM Alignment (CDCA). As shown in Table 4, we conduct all ablation experiments on the 3D-FUTURE dataset, using the CRN dataset as the source domain and 3D-FUTURE as the target domain.

Effectiveness of each component. We begin with a **Baseline** that independently serializes source and target point clouds using the Z-order curve without considering cross-domain spatial correspondence. As shown in Table 4, this naive strategy yields suboptimal performance, with an average Chamfer Distance (CD) of 23.38, due to the misalignment of spatial structures and the inability to learn domain-invariant features. Introducing the **CDPS** leads to a sub-

Model	Params(M)	FLOPs(G)	Time(ms)
DAPoinTr	36.904	24.912	23.774
DAPointMamba	9.571	5.192	3.820

Table 5: Computational efficiency of the current state-of-the-art approach DAPoinTr and our DAPointMamba.

stantial improvement, reducing the average CD to 21.73. This module establishes spatial correspondence by aligning patches within a shared coordinate space, effectively mitigating local geometric discrepancies. Notably, the cabinet and sofa categories exhibit marked gains, with CD values dropping from 23.14 to 20.06 and from 25.71 to 23.35, respectively. Building on this, we incorporate the **CDSA**, which explicitly models patch-level similarities between source and target domains. This further reduces the average CD to 21.17. The improvement is particularly evident in geometrically complex categories such as lamp (from 25.37 to 24.29) and chair (from 17.02 to 16.45), highlighting CDSA’s ability to handle fine-grained local misalignments. Finally, by adding **CDCA**, we achieve an average CD of 20.40—the best overall performance. CDCA addresses residual domain shifts in the global semantic space by adaptively aligning feature channels. Its effectiveness is especially pronounced in high-variance categories such as lamp and table, where CD drops from 24.29 to 22.81 and from 22.61 to 21.25, respectively. These results demonstrate that all individual components are complementary, and their integration leads to significant improvements in UDA point cloud completion.

Visualization and Analysis

Computational Efficiency. To assess the computational efficiency of our model, we conduct experiments on 3D-FUTURE and compare it with the state-of-the-art DAPoinTr from different perspectives, including parameters, FLOPs, and inference time. For fair comparison, all evaluations are conducted on an RTX 4090 GPU with 64GB RAM. We fix the batch size to 1 and the number of points to 2048, and the inference time is averaged over 100 experiments to ensure stability. As shown in Table 5, our model exhibits significantly lower computational cost compared to the current best model, DAPoinTr. Our model greatly reduces parameters (9.571M vs. 36.904M), FLOPs (5.192G vs. 24.912G), and inference time (3.820 ms vs. 23.774 ms). These results demonstrate that DAPointMamba achieves superior performance with over **80%** lower computational cost and latency.

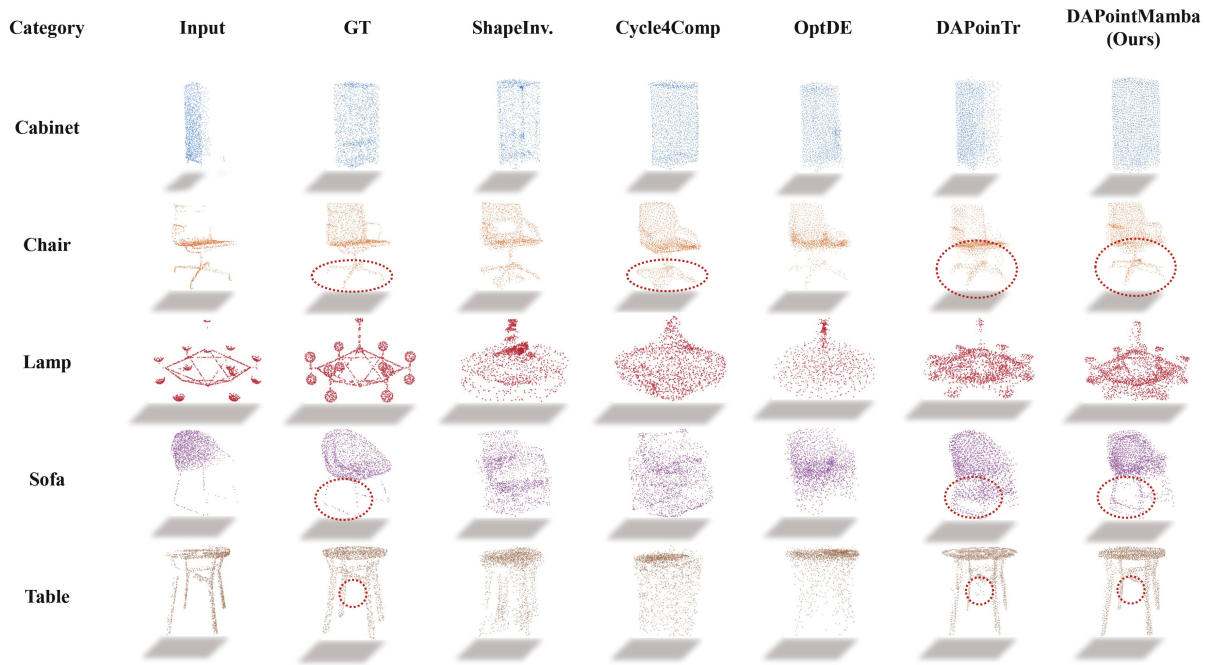


Figure 3: Visualization comparisons between ours and state-of-the-art PCC methods on 3D-FUTURE dataset.

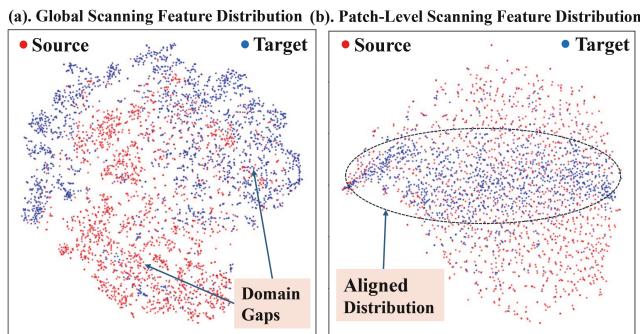


Figure 4: Visualization of the feature distribution of Cross-Domain Patch-Level Scanning (CDPS) compared with the global grouping scan strategy.

Visualization results on 3D-FUTURE. Figure 3 presents visualizations of point cloud completion results on the 3D-FUTURE dataset. Compared to prior state-of-the-art methods, our DAPointMamba consistently predicts more accurate and structurally complete outputs, with a clear advantage in preserving fine-grained local details and global shape. This improvement is particularly evident in the cases of chairs, sofas, and tables, as highlighted in the third, fourth, and last rows of the figure. These visual results further demonstrate our robustness in delivering better reconstructions across various cross-domain scenarios.

TSNE Visualization of Global and Patch-Level Scanning. The TSNE visualization (Figure 4) compares the effectiveness of our proposed Cross-Domain Patch-Level Scanning (CDPS) strategy with the conventional global scan-

ning approach in terms of feature distribution alignment between source and target domains. As illustrated in Figure 4(a), the global scanning method leads to a distinct separation between source (blue points) and target domain (red points), indicating apparent domain gaps and inadequate feature alignment. In contrast, CDPS, as demonstrated in Figure 4(b), acquires superior representations, exhibiting significantly improved overlap between source and target domain features. Specifically, the alignment achieved through our CDPS module ensures that features from different domains occupy more consistent and coherent regions in the embedding space. This closer alignment underscores CDPS’s effectiveness in explicitly maintaining spatial correspondences and mitigating local geometric discrepancies at the patch level, thereby facilitating more robust and generalizable cross-domain adaptation.

Conclusion

In this paper, we present DAPointMamba, a novel and pioneering framework aimed at enhancing the transferability of Mamba blocks for point cloud completion across various domains. Through the introduction of several new modules, including Cross-Domain Patch-level Scanning, Cross-Domain Spatial SSM Alignment, and Cross-Domain Channel SSM Alignment, our DAPointMamba effectively addresses both local and global domain discrepancies, resulting in strong adaptability across domains. Furthermore, compared to the Transformer-based architecture, our DAPointMamba achieves efficient long-sequence modeling with linear complexity. Extensive experiments on synthetic and real-world datasets demonstrate superior performance to state-of-the-art methods, particularly in complex object categories.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 62502178).

References

- Bahri, A.; Yazdanpanah, M.; Noori, M.; Dastani, S.; Cheraghalikhani, M.; Hakim, G. A. V.; Osowiechi, D.; Beizaee, F.; Ben Ayed, I.; and Desrosiers, C. 2025. Spectral informed mamba for robust point cloud processing. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 11799–11809.
- Chang, A.; Dai, A.; Funkhouser, T.; Halber, M.; Niessner, M.; Savva, M.; Song, S.; Zeng, A.; and Zhang, Y. 2017. Matterport3d: Learning from rgb-d data in indoor environments. *arXiv preprint arXiv:1709.06158*.
- Chen, X.; Chen, B.; and Mitra, N. J. 2019. Unpaired point cloud completion on real scans using adversarial training. *arXiv preprint arXiv:1904.00069*.
- Dai, A.; Chang, A. X.; Savva, M.; Halber, M.; Funkhouser, T.; and Nießner, M. 2017. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5828–5839.
- Fei, B.; Yang, W.; Chen, W.-M.; Li, Z.; Li, Y.; Ma, T.; Hu, X.; and Ma, L. 2022. Comprehensive review of deep learning-based 3d point cloud completion processing and analysis. *IEEE Transactions on Intelligent Transportation Systems*, 23(12): 22862–22883.
- Fu, H.; Jia, R.; Gao, L.; Gong, M.; Zhao, B.; Maybank, S.; and Tao, D. 2021. 3d-future: 3d furniture shape with texture. *International Journal of Computer Vision*, 129: 3313–3337.
- Geiger, A.; Lenz, P.; and Urtasun, R. 2012. Are we ready for autonomous driving? the kitti vision benchmark suite. In *IEEE Conference on Computer Vision and Pattern Recognition*, 3354–3361.
- Gong, J.; Liu, F.; Xu, J.; Wang, M.; Tan, X.; Zhang, Z.; Yi, R.; Song, H.; Xie, Y.; and Ma, L. 2022. Optimization over disentangled encoding: Unsupervised cross-domain point cloud completion via occlusion factor manipulation. In *European Conference on Computer Vision*, 517–533.
- Groueix, T.; Fisher, M.; Kim, V. G.; Russell, B. C.; and Aubry, M. 2018. A papier-mâché approach to learning 3d surface generation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 216–224.
- Gu, A.; and Dao, T. 2023. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*.
- Han, X.; Tang, Y.; Wang, Z.; and Li, X. 2024. Mamba3d: Enhancing local features for 3d point cloud analysis via state space model. In *Proceedings of the 32nd ACM International Conference on Multimedia*, 4995–5004.
- Hong, S.; Yavartanoo, M.; Neshatavar, R.; and Lee, K. M. 2023. 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*, 9435–9444.
- Li, D.; Zhao, J.; Chang, C.; Chen, Z.; and Du, J. 2025a. LGMamba: Large-Scale ALS Point Cloud Semantic Segmentation With Local and Global State-Space Model. *IEEE Geoscience and Remote Sensing Letters*, 22: 1–5.
- Li, Y.; Yang, W.; and Fei, B. 2024. 3dmambacomplete: Exploring structured state space model for point cloud completion. *arXiv preprint arXiv:2404.07106*.
- Li, Y.; Zhou, Q.; Gong, J.; Zhu, Y.; Dazeley, R.; Zhao, X.; and Lu, X. 2025b. Dapointr: Domain adaptive point transformer for point cloud completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 5066–5074.
- Liang, D.; Zhou, X.; Xu, W.; Zhu, X.; Zou, Z.; Ye, X.; Tan, X.; and Bai, X. 2024. Pointmamba: A simple state space model for point cloud analysis. *arXiv preprint arXiv:2402.10739*.
- Liu, F.; Gong, J.; Zhou, Q.; Lu, X.; Yi, R.; Xie, Y.; and Ma, L. 2024a. Cloudmix: Dual mixup consistency for unpaired point cloud completion. *IEEE Transactions on Visualization and Computer Graphics*, 31(4): 2182–2195.
- Liu, Y.; Tian, Y.; Zhao, Y.; Yu, H.; Xie, L.; Wang, Y.; Ye, Q.; Jiao, J.; and Liu, Y. 2024b. Vmamba: Visual state space model. *Advances in Neural Information Processing Systems*, 37: 103031–103063.
- Tchapmi, L. P.; Kosaraju, V.; Rezatofighi, H.; Reid, I.; and Savarese, S. 2019. Topnet: Structural point cloud decoder. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 383–392.
- Tesema, K. W.; Hill, L.; Jones, M. W.; Ahmad, M. I.; and Tam, G. K. 2023. Point cloud completion: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 30(10): 6880–6899.
- Wang, J.; Cui, Y.; Guo, D.; Li, J.; Liu, Q.; and Shen, C. 2024a. Pointattn: You only need attention for point cloud completion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 5472–5480.
- Wang, X.; Ang Jr, M. H.; and Lee, G. H. 2020. Cascaded refinement network for point cloud completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 790–799.
- Wang, Z.; Chen, Z.; Wu, Y.; Zhao, Z.; Zhou, L.; and Xu, D. 2024b. Pointramamba: A hybrid transformer-mamba framework for point cloud analysis. *arXiv preprint arXiv:2405.15463*.
- Wei, X.; Gu, X.; and Sun, J. 2022. Learning generalizable part-based feature representation for 3d point clouds. *Advances in Neural Information Processing Systems*, 35: 29305–29318.
- Wei, X.; Gu, X.; and Sun, J. 2024. Multi-Scale Part-Based Feature Representation for 3D Domain Generalization and Adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Wen, X.; Han, Z.; Cao, Y.-P.; Wan, P.; Zheng, W.; and Liu, Y.-S. 2021. 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*, 13080–13089.

Wu, X.; Jiang, L.; Wang, P.-S.; Liu, Z.; Liu, X.; Qiao, Y.; Ouyang, W.; He, T.; and Zhao, H. 2024. Point transformer v3: Simpler faster stronger. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4840–4851.

Wu, Z.; Song, S.; Khosla, A.; Yu, F.; Zhang, L.; Tang, X.; and Xiao, J. 2015. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1912–1920.

Xiang, P.; Wen, X.; Liu, Y.-S.; Cao, Y.-P.; Wan, P.; Zheng, W.; and Han, Z. 2021. Snowflakenet: Point cloud completion by snowflake point deconvolution with skip-transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 5499–5509.

Yang, Y.; Kim, J.; and Yoon, K.-J. 2024. Syn-to-real domain adaptation for point cloud completion via part-based approach. In *European Conference on Computer Vision*, 179–197.

Yu, X.; Rao, Y.; Wang, Z.; Liu, Z.; Lu, J.; and Zhou, J. 2021. PointR: Diverse point cloud completion with geometry-aware transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 12498–12507.

Yuan, W.; Khot, T.; Held, D.; Mertz, C.; and Hebert, M. 2018. Pcn: Point completion network. In *International Conference on 3D Vision*, 728–737.

Zhang, J.; Chen, X.; Cai, Z.; Pan, L.; Zhao, H.; Yi, S.; Yeo, C. K.; Dai, B.; and Loy, C. C. 2021. Unsupervised 3d shape completion through gan inversion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1768–1777.

Zhang, T.; Yuan, H.; Qi, L.; Zhang, J.; Zhou, Q.; Ji, S.; Yan, S.; and Li, X. 2025. Point cloud mamba: Point cloud learning via state space model. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 10121–10130.

Zhou, H.; Cao, Y.; Chu, W.; Zhu, J.; Lu, T.; Tai, Y.; and Wang, C. 2022. Seedformer: Patch seeds based point cloud completion with upsample transformer. In *European Conference on Computer Vision*, 416–432. Springer.

Zhu, L.; Liao, B.; Zhang, Q.; Wang, X.; Liu, W.; and Wang, X. 2024. Vision Mamba: Efficient Visual Representation Learning with Bidirectional State Space Model. In *International Conference on Machine Learning*, 62429–62442.