HiDiGen: Hierarchical Diffusion for B-Rep Generation with Explicit Topological Constraints

Supplementary Material

This supplementary material provides additional details and results to support the claims in the main paper. We first elaborate on the evaluation metrics used to assess geometric fidelity, topological validity, and generation diversity. We then present representative failure cases that highlight current limitations in jointly modeling topology and geometry , followed by extended qualitative results on the DeepCAD and ABC datasets.

1. Details of Evaluation Metrics

In this section, we present more details about evaluation metric used in this paper.

- **Minimum Matching Distance (MMD-CD).** This metric measures the average Chamfer Distance from each reference point cloud to its nearest neighbor in the generated set, reflecting the quality of the best-matching samples.
- **Coverage (COV-CD).** COV-CD quantifies the diversity of the generated point clouds by computing the fraction of reference point clouds whose nearest neighbor in the generated set is unique. Higher coverage suggests that the model captures a broader range of the underlying data distribution.
- **Jensen–Shannon Divergence (JSD).** JSD evaluates the similarity between the occupancy distributions of real and generated point clouds on a discretized 3D grid. It provides a symmetric measure of distributional discrepancy, where a lower JSD implies better alignment between the synthetic and ground-truth data manifolds.
- **Valid.** We report the percentage of generated B-reps that form watertight, non-manifold-free solids. A model is considered valid only if it satisfies fundamental topological constraints (e.g., each edge belongs to exactly two distinct faces and connects two distinct vertices) and passes geometric validation via a standard CAD kernel such as OpenCascade.
- **Novel.** Novelty measures the fraction of generated models that do not appear in the training set. We employ a hash-based identity check on the full B-rep structure.
- **Unique.** Uniqueness quantifies intra-sample diversity by computing the proportion of generated models that occur only once in the output set. Using the B-rep hashing scheme, this metric reflects the model's ability to avoid repetitive or mode-collapsed outputs.
- **Average Number of Boundary Loops.** This metric counts the total number of topological boundary loops (i.e., `WIRE`s) across all faces of a CAD model and averages it over the dataset. Each loop corresponds to a connected boundary component of a face, including outer contours and inner holes. A higher value indicates more complex face topologies, often associated with intricate engineering features such as cutouts or fillets.
- **Average Mean Curvature.** We compute the mean curvature $(k_1 + k_2)/2$ at uniformly sampled points on each face and average it across all faces and shapes. This scalar quantity captures the intrinsic smoothness and bending of surfaces: planar regions contribute zero curvature, while cylindrical or spherical regions yield non-zero values. The metric serves as a proxy for geometric complexity, with higher average curvature reflecting richer shape variation.

2. Failure Analysis

Due to the intricate interplay between topological structure and geometric attributes in B-rep CAD generation, our approach employs four independent diffusion models and two autoregressive models. Although these models can leverage prior information during training, they may generate out-of-distribution predictions at inference time. This leads to two primary failure modes: (1) topological errors that propagate to produce geometrically invalid faces, and (2) inaccurate geometry that induces spurious topological connections. We illustrate representative failure cases in Fig. 1.

3. Loss Functions

Our framework is trained with three task-specific objectives corresponding to edge-face adjacency prediction, autoregressive edge-vertex sequence generation, and face geometry synthesis.

The edge-face module is trained as a variational autoencoder. Its loss combines a reconstruction term and a KL regularization term:

$$\mathcal{L}_{\text{ef}} = \underbrace{\mathbb{E}\big[-\log p(\mathbf{EF}_i^{seq} \mid \mathbf{z})\big]}_{\text{reconstruction}} + \underbrace{D_{\text{KL}}\big(q(\mathbf{z} \mid \mathbf{EF}_i^{seq}) \,\|\, \mathcal{N}(0, \mathbf{I})\big)}_{\text{regularization}}, \tag{1}$$

where$\mathbf{EF}_i^{seq}$ denotes the ground-truth upper-triangular edge–face adjacency matrix, and $\mathbf{z}$ is the stochastic latent code. The reconstruction loss is implemented as cross-entropy over discrete adjacency entries, and the KL term admits the closed-form expression $-\frac{1}{2}\mathbb{E}[1 + \log\sigma^2 - \mu^2 - \sigma^2]$.

The edge-vertex generator is trained autoregressively via maximum likelihood. Given the ground-truth vertex sequence $\mathbf{v}_{1:T}$, the model predicts a categorical distribution over the next token at each step. The loss is the negative log-likelihood of the sequence, masked to exclude padding and special tokens, and normalized by the total number of valid prediction positions:

$$\mathcal{L}_{\text{ev}} = -\frac{1}{\sum_{i,t} m_{i,t}} \sum_{i=1}^{B} \sum_{t=1}^{T-1} m_{i,t} \log p(v_{i,t+1} \mid v_{i,1:t}), \quad (2)$$

where $m_{i,t}$ indicates whether position t in batch i is a valid prediction target.

For geometry synthesis, we adopt a diffusion-based approach trained to predict noise added to face coordinates. The loss is the mean squared error between predicted and ground-truth noise, restricted to valid (non-padded) attributes:

$$\mathcal{L}_{\text{geo}} = \frac{1}{|\mathcal{F}_{\text{valid}}|} \sum_{f \in \mathcal{F}_{\text{valid}}} \left\| \epsilon_\theta(f) - \epsilon \right\|_2^2, \quad (3)$$

where ϵ is the sampled noise, $\epsilon_\theta(f)$ is the model's prediction, and $\mathcal{F}_{\text{valid}}$ denotes the set of faces present in the input attributes.

All three losses are optimized jointly using AdamW with a learning rate of 5×10^{-4} and weight decay of 1×10^{-6}.

4. Additional Qualitative Results

Fig. 2 presents additional qualitative results of our method on the DeepCAD and ABC datasets.

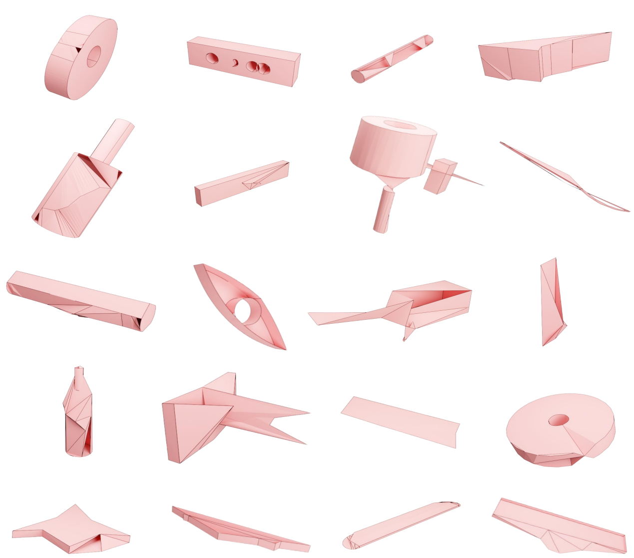

Figure 1. Failure cases illustrating topological and geometric inconsistencies in generated B-reps.

Figure 2. Additional qualitative results of HiDiGen on the DeepCAD datasets.

Figure 3. Additional qualitative results of HiDiGen on the ABC datasets.