1. Details of Evaluation Metrics

In the main paper, we used symmetric Chamfer Distance (CD), Normal Consistency (NC), Edge Chamfer Distance (ECD) as evaluation metrics. Specifically, CD measures the squared distances between nearest neighbors of two point clouds:

\[ CD = \mathbb{E}_{x \in P} \min_{y \in G} \|x - y\|_2^2 + \mathbb{E}_{y \in G} \min_{x \in P} \|x - y\|_2^2, \]  

where P and G are the point clouds uniformly sampled from the prediction mesh and ground truth mesh.

ECD is calculated in the same way as CD, except that the targets are points that lie on the sharp edges of the shape, forming subsets \( P_E \) and \( G_E \) of uniformly sampled points:

\[ ECD = \mathbb{E}_{x \in P_E} \min_{y \in G_E} \|x - y\|_2^2 + \mathbb{E}_{y \in G_E} \min_{x \in P_E} \|x - y\|_2^2, \]  

The close-to-edge points for calculating ECD are obtained from 12,000 surface-sampled points. We adopt the settings of distance and sharpness threshold from [1]. Also utilizing the nearest neighbors, NC calculates the dot product of their normal vectors:

\[ NC = \left( \mathbb{E}_{x \in P} [n_x \cdot n_{x_y}] + \mathbb{E}_{y \in G} [n_y \cdot n_{x_y}] \right) / 2, \]  

where \( n \) represents the normal vector, \( x_y \) stands for the neighbor points of point cloud \( x \) in point cloud \( y \), and vice versa. We use 8,192 surface-sampled points to compute CD and NC. For easy viewing, we scale CD by 1,000 and ECD by 100.

2. Discussion of primitives and assemblies of different methods

For primitive shapes, there is a trade-off between expressiveness and ease of editing. In addition to the methods we compared in the main paper, BSP-Net [1] and CAPRI-Net [6] are two other methods that use primitive shapes as well as CSG operations for shape reconstruction. As shown in Table 1, the primitive shapes (convexes in the original papers) of BSP-Net and CAPRI-Net are intersections of linear partitions or quadratic partitions, which cannot be edited directly in 2D. UCSG-Net and CSG-Stump adopt fixed shapes, such as boxes and spheres, as primitives. The outputs of our method, point2Cyl and ExtrudeNet contain sketches, making the cylinder primitives editable in 2D. As shown in Table 1, the property of generating primitives that can be edited in 2D makes the methods in the last four rows within a fair comparison, so we report qualitative and quantitative comparisons of such methods in the main text.

Table 1. Comparison of different shape representations for reconstruction methods based on primitives.

<table>
<thead>
<tr>
<th>Method</th>
<th>Primitives</th>
<th>2D editable</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSP-Net [1]</td>
<td>Intersection of linear partitions</td>
<td>✗</td>
</tr>
<tr>
<td>CAPRI-Net [6]</td>
<td>Intersection of quadratic partitions</td>
<td>✗</td>
</tr>
<tr>
<td>UCSG-Net [2]</td>
<td>Box, Sphere</td>
<td>✓</td>
</tr>
<tr>
<td>CSG-Stump [3]</td>
<td>Box, Sphere, Cylinder, Cone</td>
<td>✓</td>
</tr>
<tr>
<td>Point2Cyl [5]</td>
<td>Extrusion cylinder</td>
<td>✓</td>
</tr>
<tr>
<td>Ours</td>
<td>Extrusion cylinder</td>
<td>✓</td>
</tr>
</tbody>
</table>

3. More comparison results

We show more reconstruction results of each method on the ABC and Fusion 360 datasets in Fig. 2 and Fig. 3, respectively. The modeling algorithm of the output meshes and the display order of the results are the same as in the main paper. To highlight the differences between SECAD-Net and ExtrudeNet [4], we further show the segmentation of the reconstruction results of the two methods in Fig. 1. Each color in the figure represents a primitive before the final union operation.
4. Comparison results without fine-tuning

We compare the reconstruction performance of each method before fine-tuning on the ABC and Fusion datasets. Using the same evaluation metrics as the main paper, we list the quantitative results in Table 2. The results show that our method still achieves the best results in five out of six evaluations.

Table 2. Quantitative comparison between different methods without fine-tuning on ABC/Fusion 360 datasets.

<table>
<thead>
<tr>
<th>Methods</th>
<th>CD↓ [m]</th>
<th>ECD↓ [m]</th>
<th>NC↑ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCSG-Net [2]</td>
<td>3.14/4.45</td>
<td>12.0/17.0</td>
<td>70.48/66.63</td>
</tr>
<tr>
<td>CSG-Stump [3]</td>
<td>3.34/5.22</td>
<td>3.14/6.88</td>
<td>70.70/66.28</td>
</tr>
<tr>
<td>ExtrudeNet [4]</td>
<td>3.07/5.61</td>
<td>2.99/6.92</td>
<td>71.35/68.15</td>
</tr>
<tr>
<td>Ours</td>
<td>2.94/4.20</td>
<td>3.00/5.53</td>
<td>71.98/68.68</td>
</tr>
</tbody>
</table>

5. More ablation comparisons of sketch representations

In order to demonstrate the advantages of our sketch head network (SK-head) over other classical shape representations, we further show the profiles predicted by different sketch representations in the ablation study in Fig. 6. It can be seen that our SK-head can predict smoother and more complete profiles compared to BSP and box primitives.

6. More visual results of sketch interpolations

To further explore the interpretability of the proposed network, we show the 3D shapes and profiles corresponding to the linearly interpolated embedding codes in Fig. 4 and Fig. 5. The new sketches obtained by interpolation can be directly sent to CAD software for editing and re-creation.

7. CSG assembly using cylinder primitives

In our final CAD result, the smallest unit is the cylinder primitives, which can be assembled into the final shape using CSG operations. The assembly process forms a CSG tree-like structure, and we show some assembly visualization results in Fig. 7.

References


Figure 2. Visual comparison of our approach with UCSG-Net [2], CSG-Stump [3], and ExtrudeNet [4] on the ABC dataset. Our reconstruction results are generated using both marching cubes (MC) and our proposed sketch-extrude operations (SE).
Figure 3. Visual comparison between reconstruction results on Fusion 360 dataset.
Figure 4. The leftmost and rightmost columns are the prediction results of SECAD-Net on the ABC dataset. The middle results are obtained by interpolating the latent embeddings. For each case, the top row shows the 3D shape, and each remaining row represents the sketches learned by one sketch head network.
Figure 5. The leftmost and rightmost columns are the prediction results of SECAD-Net on the ABC dataset. The middle results are obtained by interpolating the latent embeddings. For each case, the top row shows the 3D shape, and each remaining row represents the sketches learned by one sketch head network.
Figure 6. Visual comparison results for ablation study on sketch representation. Each case contains 3D shapes modeled with different sketch representations and the corresponding binary profile images.
Figure 7. Visualization of assembling cylinder primitives into complete shapes with Boolean operations. Therefore, the final output CAD models are represented as a CSG tree-like structure.