Supplementary Material *for*
CNC-Net: Self-Supervised Learning for CNC Machining Operations

Mohsen Yavartanoo[1*] Sangmin Hong[2*] Reyhaneh Neshatavar[1*] Kyoung Mu Lee[1,2]
[1]Dept. of ECE & ASRI, [2]IPAI, Seoul National University, Seoul, Korea
{myavartanoo,mchiash2,reyhanehneshat,kyoungmu}@snu.ac.kr

S1. Inferior performance with CD and NC

In Sec. 4.3.1 of the main manuscript, we delve into the comparative analysis of the performance of our CNC-Net method against state-of-the-art (SOTA) techniques [23, 29, 30, 37]. In this comparison, our method exhibits inferior results in surface-based metrics, including CD [25] and NC [4], while excelling in volume-based metrics when compared with such SOTA methods. This discrepancy arises from the fundamental difference in approach: prior methods construct shapes by combining several smooth primitives, whereas our approach approximates shapes by carving a cube. To visually illustrate this distinction, we offer a closer examination of the surface in the CAD models produced by our CNC-Net method versus those generated by SOTA methods [23, 29, 37] in Fig. S1. The comparisons demonstrate that the results of our CNC-Net method are superior to those of these methods in forming the target shape to these methods [23, 29, 37] with the IoU metric that is directly reflected in the visual appearance. However, the discrepancy in performance, as indicated by CD and NC metrics, is attributed to the non-smooth surfaces resulting from our approach of carving a cube to approximate shapes.

S2. Milling paths

In Sec. 4.4.1 of the main manuscript, we present visualizations of the milling paths for some randomly selected samples from the ABC [16] dataset. In Fig. S2, we further provide the generated paths of the first four subsequent operations for more samples from the ABC [16] dataset. The results demonstrate the capacity of our method to intelligently learn paths in a self-supervised manner, eliminating the necessity for ground truth paths.

S3. Zero-shot Vs. fine-tuning

In Sec. 4.4.2 of the main manuscript, we conduct a quantitative comparison between results achieved through fine-tuning and those obtained via zero-shot learning. To further analyze the differences between the results obtained by these two training strategies, we present a visual comparison of our outcomes on the ABC [16] dataset, as illustrated in Fig. S3. The qualitative results indicate that our zero-shot learning method can effectively reproduce the general shapes of target objects. However, leveraging a pre-trained model and fine-tuning it for each sample notably enhances the reproduction of finer details. The zero-shot capability significantly enhances the practicality of our method, particularly when tackling target objects that differ substantially from shapes in the training dataset.

S4. Effect of operations

In Sec. 4.4.4 of the main manuscript, we quantitatively demonstrate the impact of each designed operation on achieving enhanced performance. Furthermore, we visually represent their effects using some samples from the ABC [16] dataset as shown in Fig. S4. Our qualitative results also underscore the necessity of the milling and rotation operations in forming the general shape of the target object. On the contrary, the drilling operation significantly improves precision in production. Specifically, in the case of Wo $\mathcal{O}^{\mathcal{M}}$ as depicted in Fig. S4, the method attempts to generate the shape by creating a hole, resulting in a shape that remains notably distant from the intended target shape. On the other hand, the method Wo $\mathcal{R}ot$ either removes some parts of the target object (first row) or leaves all parts intact (second row). Furthermore, while the method Wo $\mathcal{O}^{\mathcal{D}}$ reproduces shapes resembling the targe shapes, it still fails to capture finer details.

S5. More qualitative results

We expand our comparative analysis by offering additional visual comparisons between our CNC-Net method and SOTA 3D CAD reconstruction approaches [29, 30, 37, 23], assessed across both the ABC [16] and the ShapeNet [3] datasets. We illustrate these comparisons for the ABC and the ShapeNet datasets in Fig. S5 and Fig. S6, respectively.

[*]equal contribution

S6. Milling path efficiency and visualization

In Sec. 4.4.4 of the main manuscript, we present visualizations of milling paths for selected samples from the ABC [16] dataset. Fig. S7 offers a comparative analysis between our method and MeshCAM, a widely adopted CAM software for generating tool paths for 3-axis CNC machines. The comparison reveals that MeshCAM generates very dense paths, carving each layer from top down. In contrast, our method, designed for 5-axis CNC machines, affords greater flexibility in the positioning of carving actions. This flexibility translates into a significant efficiency gain in terms of the length of the tool path. We further quantify this result by presenting the total path lengths for each sample. On average the tool path length produced by our method is 13.01 times shorter than that generated by MeshCAM software. These results underscore the superiority of our approach in crafting considerably shorter, and thus more efficient, paths.

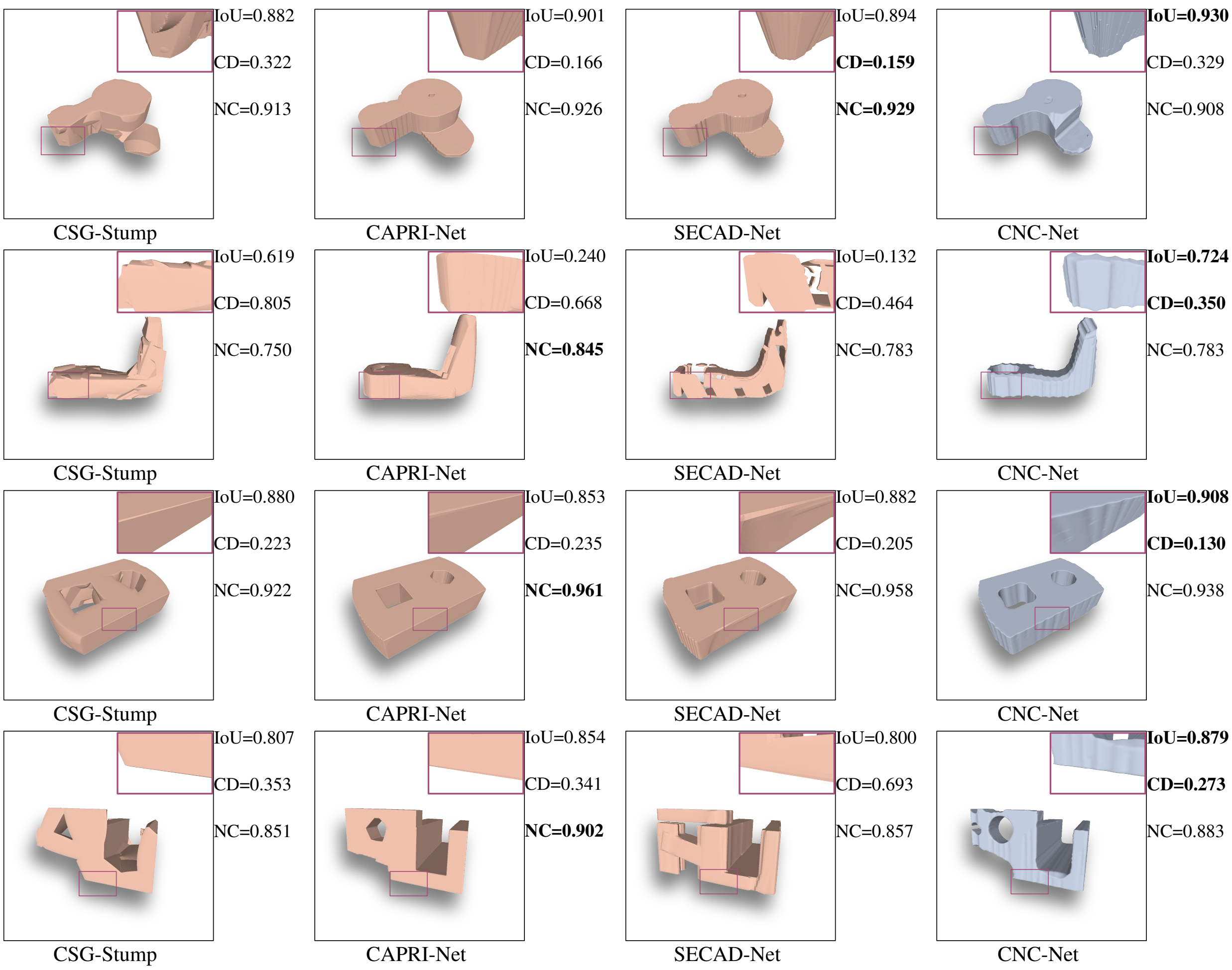

Figure S1. **Comparison on the surface of the reconstructed shapes.**

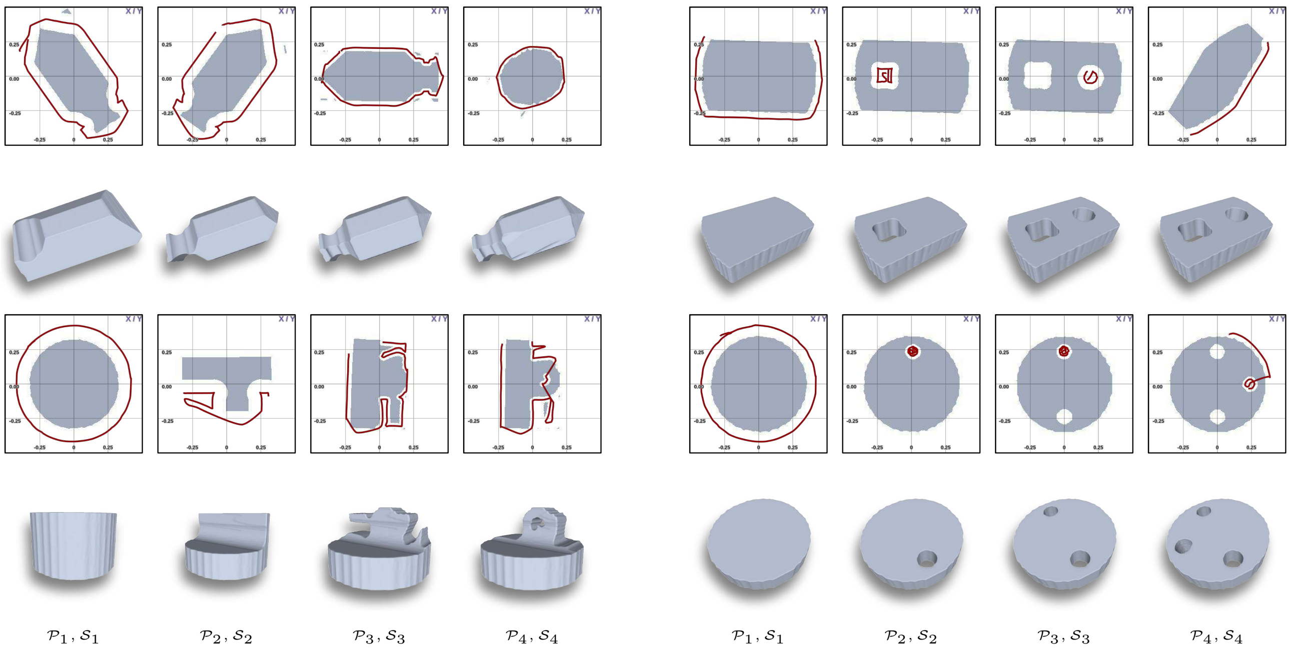

Figure S2. **Ablation study for milling paths.** The odd-numbered rows display the path $\mathcal{P}$ in steps $s = 1, \ldots, 4$ from the top view. The even-numbered rows depict the reproduced shapes in each step.

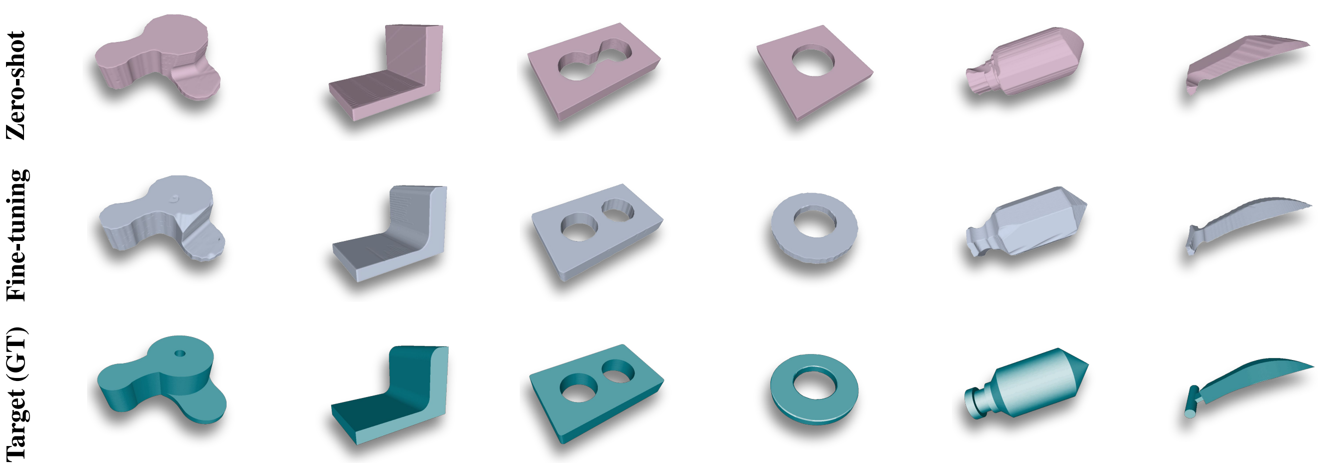

Figure S3. **Qualitative results for zero-shot Vs. fine-tuning**

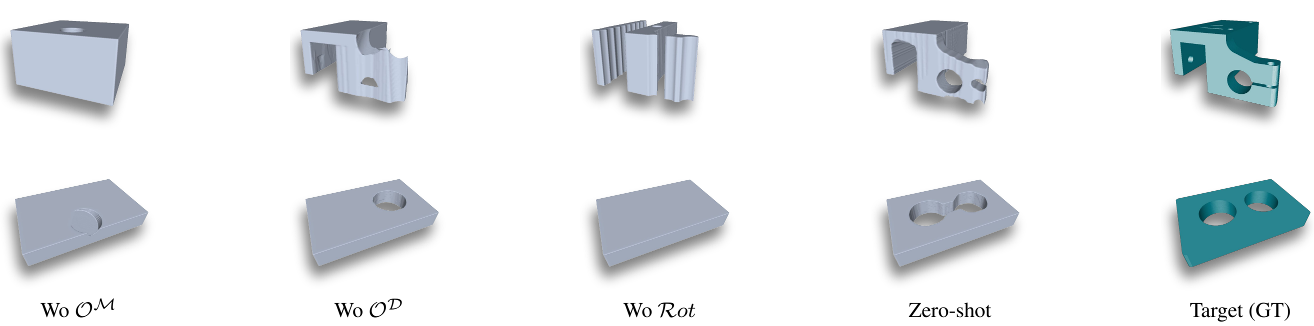

Figure S4. **Qualitative comparison of the results with different operations.** From the 1$^{\text{st}}$ to the 3$^{\text{rd}}$ column, the milling, drilling, and rotation operation is eliminated, respectively.

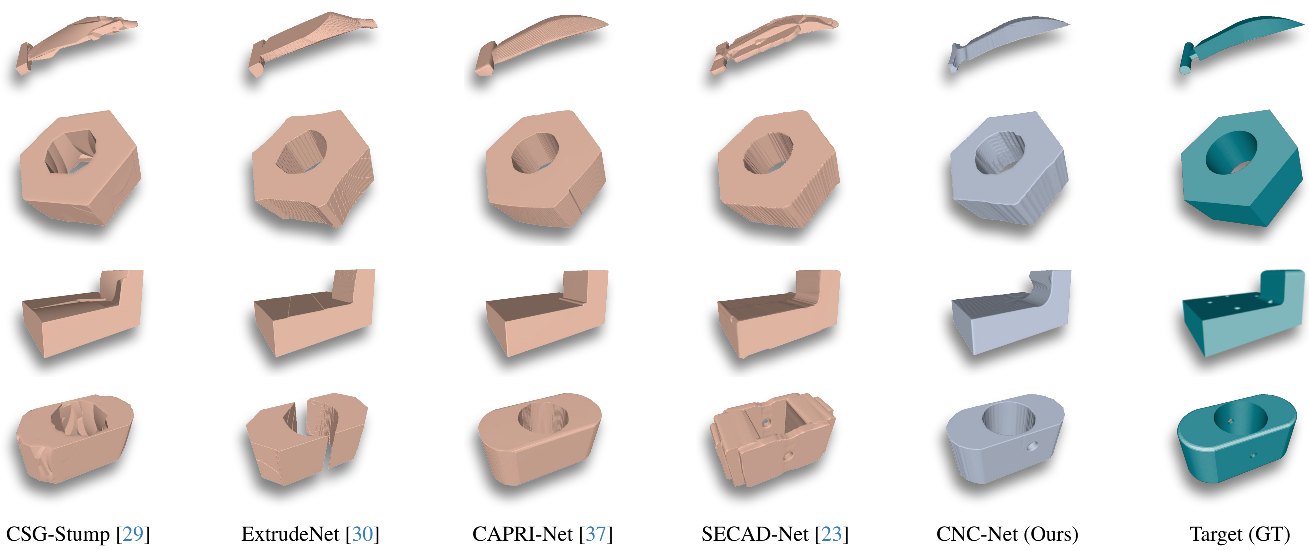

Figure S5. **Qualitative results on ABC [16] dataset.** All shapes are visualized using marching cubes (MC) with 256 resolution.

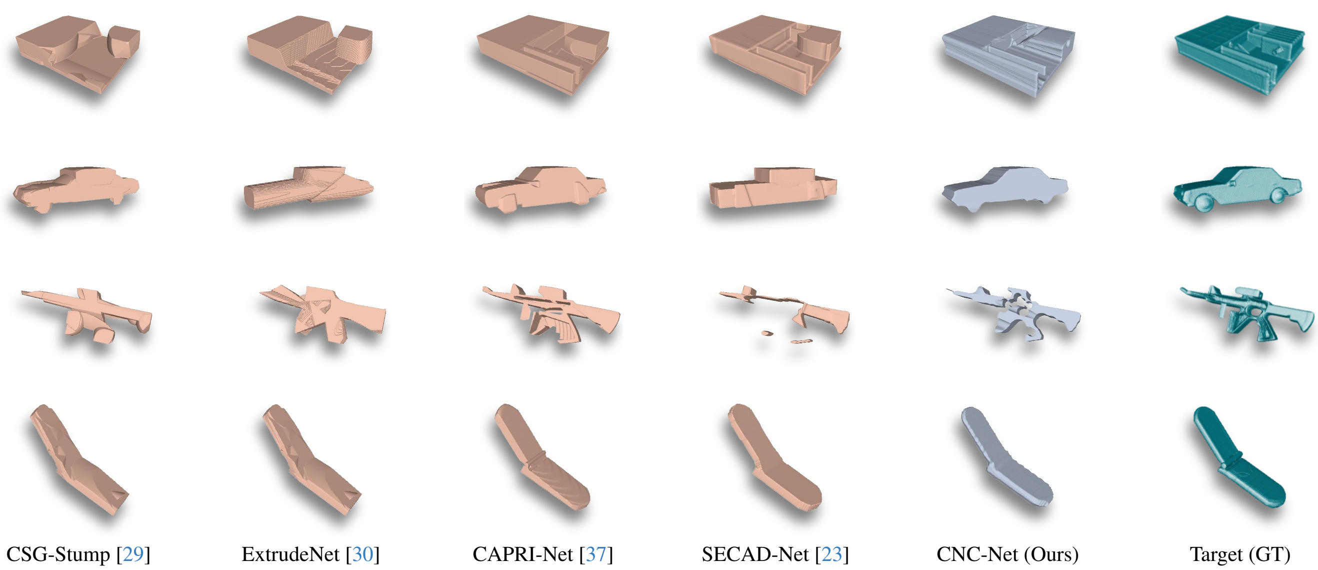

Figure S6. **Qualitative results on ShapeNet [3] dataset.** All shapes are visualized using marching cubes (MC) with 256 resolution.

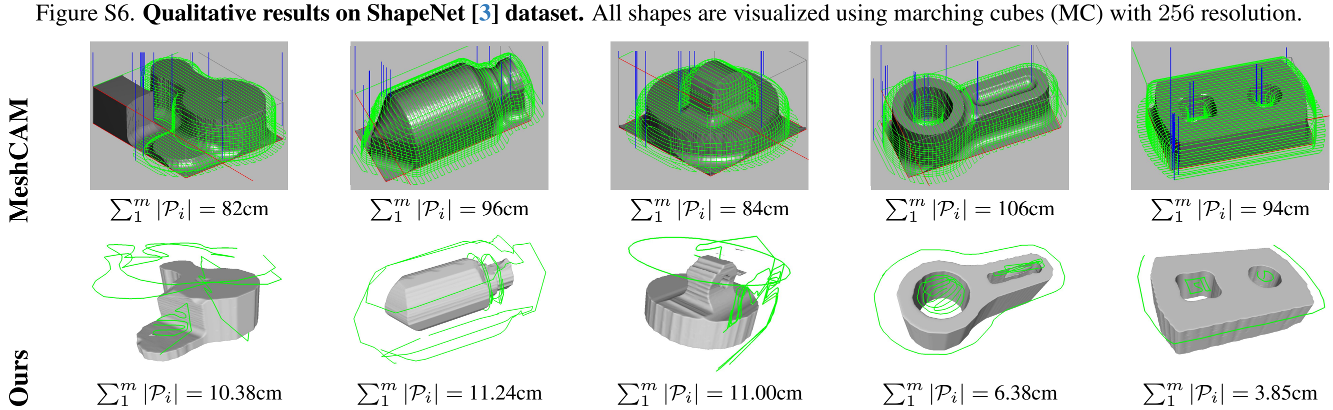

Figure S7. **3D visualization of millling path compared to MeshCAM**