GeoCAD: Local Geometry-Controllable CAD Generation with Large Language Models

Zhanwei Zhang[1]*, **Kaiyuan Liu**[1], **Junjie Liu**[2], **Wenxiao Wang**[4], **Binbin Lin**[4]†,
Liang Xie[3], **Chen Shen**[2], **Deng Cai**[1]
[1] State Key Lab of CAD&CG, Zhejiang University
[2] Alibaba Cloud Computing, [3] Zhejiang University of Technology
[4] School of Software Technology, Zhejiang University
zhanweizhang@zju.edu.cn

Motivation

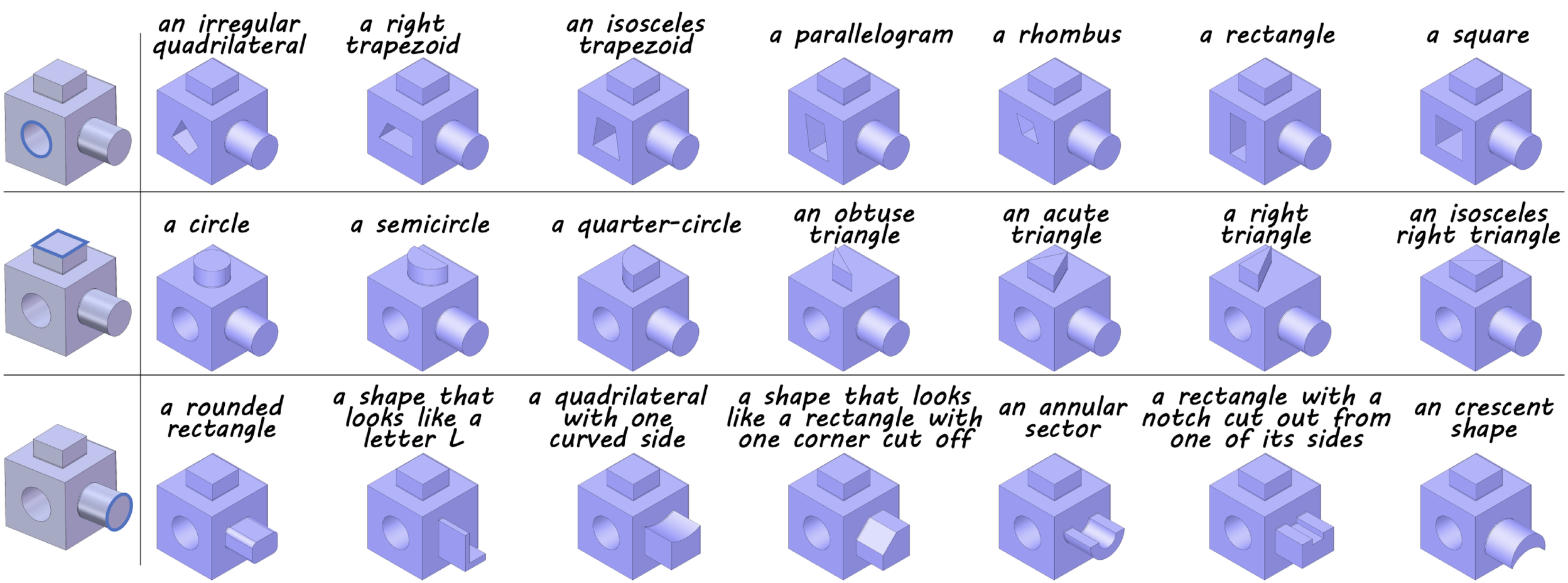


Figure 1: Local geometry-controllable CAD generation achieved by GeoCAD. The input comprises: (1) an original CAD model (the left side), (2) the local part to be modified (highlighted in blue), and (3) user-specific geometric instructions. Subsequently, GeoCAD outputs the revised CAD models where only the target part is altered while adhering to the provided geometric instructions.

Method: Captioning

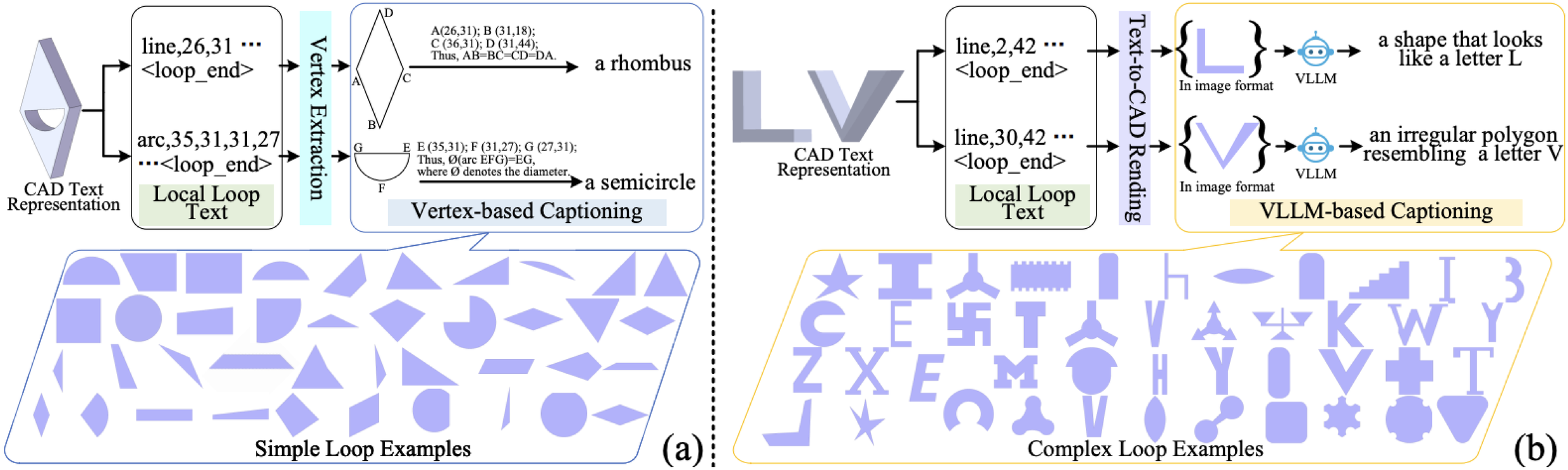


Figure 2: The complementary captioning strategy. (a) Vertex-based captioning for simple local parts. Vertex coordinates are initially extracted, followed by geometric analysis to enable precise captions. (b) VLLM-based captioning for complex local parts. We first convert complex parts into 2D images and subsequently employ powerful VLLMs to produce descriptive captions.

Method: training & Inference

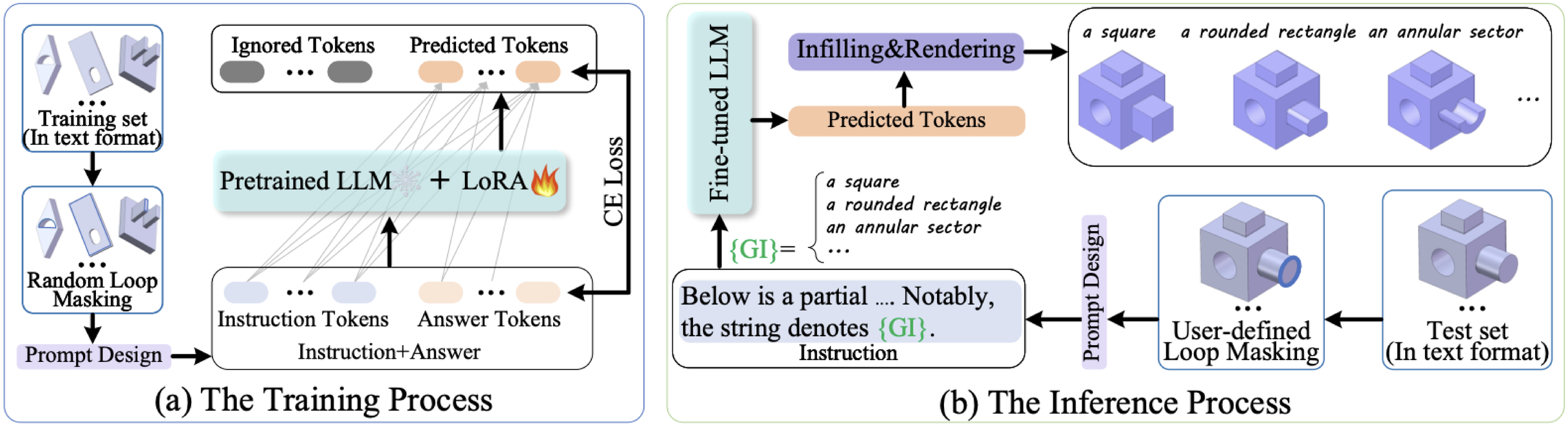
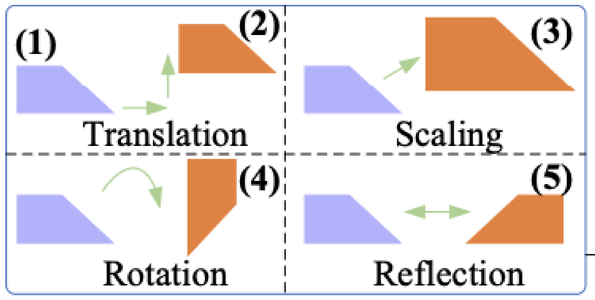


Figure 5: The overall framework of GeoCAD. (a) Training process. Given a CAD model, we randomly mask a local loop within it. During stages 1 and 2, we design the corresponding prompts (as introduced in Fig. 3 and Fig. 4), and fine-tune LLMs. (b) Inference process. Users can optionally mask any local part for modification, driven by various geometric instructions (GI). The mask part is then infilled with the predicted local parts to construct new CAD models.

Method: Prompt used in the pretraining stage

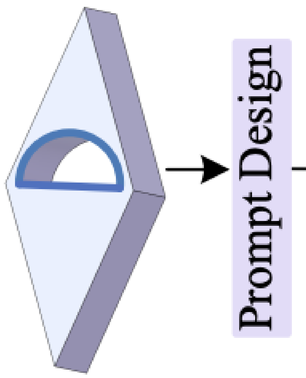


Instruction: Generate a string that denotes a right trapezoid.

Answer (1): line,31,31 <curve_end> line,31,39 <curve_end> line,39,39 <curve_end> line,47,31 <curve_end> <loop_end>
Answer (2): line,41,37 <curve_end> line,41,45 <curve_end> line,49,45 <curve_end> line,57,37 <curve_end> <loop_end>
Answer (3): line,31,31 <curve_end> line,31,43 <curve_end> line,43,43 <curve_end> line,55,31 <curve_end> <loop_end>
Answer (4): line,31,31 <curve_end> line,39,31 <curve_end> line,39,23 <curve_end> line,31,15 <curve_end> <loop_end>
Answer (5): line,31,31 <curve_end> line,31,39 <curve_end> line,23,39 <curve_end> line,15,31 <curve_end> <loop_end>

Figure 3: The prompt template used in stage 1. Local parts are first augmented through translation, scaling, rotation, and reflection. Subsequently, we construct the corresponding prompt that incorporates the geometric instruction, and ask LLMs to predict both the initial and augmented parts.

Method: Prompt used in the training stage



Prompt Design

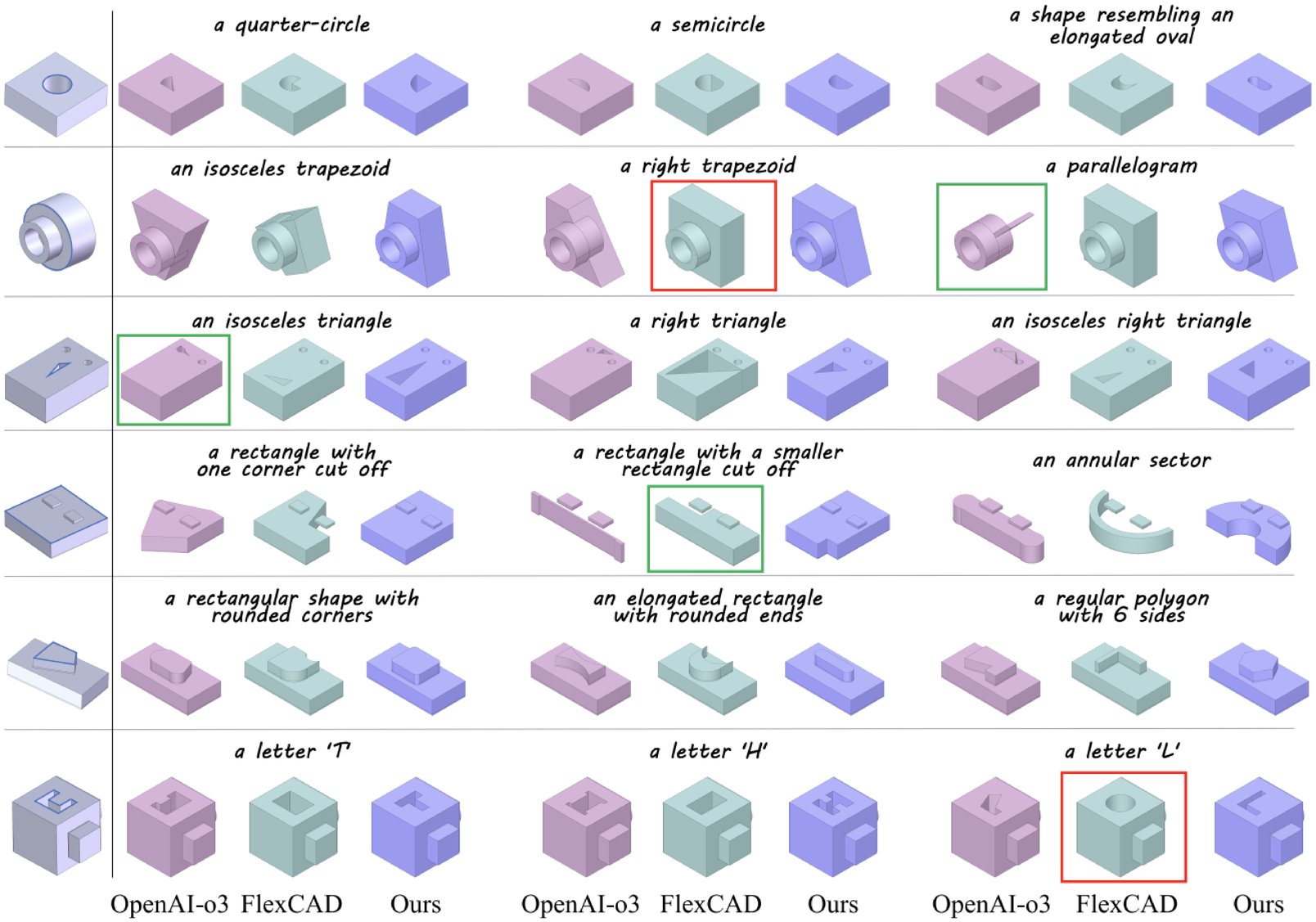
Instruction: Below is a partial description of a CAD sequence where one command has been replaced with the string "[loop mask]":
line,26,31 <curve_end> line,31,18 <curve_end> line,36,31 <curve_end> line,31,44 <curve_end> <loop_end> <face_end> <sketch_end> add,31,39,31,31,31,1,0,0,0,1,0,0,0,1,37,24,28 <extrusion_end> [loop mask] <face_end> <sketch_end> cut,31,39,31,31,31,1,0,0,0,1,0,0,0,1,37,24,28 <extrusion_end>
Generate a string that could replace "[loop mask]" in the CAD sequence. Notably, the string denotes a semicircle.

Answer: arc,35,31,31,27 <curve_end> line,27,31 <curve_end> <loop_end>

Figure 4: The prompt template used in stage 2. Given a local part (highlighted in blue) in a CAD model, we formulate the prompt that integrates the geometric instruction (highlighted in green) and the remaining parts of the CAD model, and require LLMs to predict this local part.

Experiments & Results

Model	COV↑	MMD↓	JSD↓	PV↑	Ver-score↑	VLLM-score↑	Realism↑
OpenAI-o3 (five-shot)	53.6%	1.64	1.49	65.7%	33.6%	22.1%	18.7%
FlexCAD	58.3%	1.40	1.58	86.7%	19.8%	6.93%	13.6%
FlexCAD (five-shot)	59.4%	1.37	1.34	88.1%	43.5%	26.8%	20.2%
GeoCAD	64.9%*	**1.13**	0.98*	90.5%*	76.4%*	65.7%*	40.9%*
GeoCAD (five-shot)	**66.0%**	*1.16	**0.80**	**92.3%**	**82.2%**	**68.2%**	**43.6%**



Experiments & Results

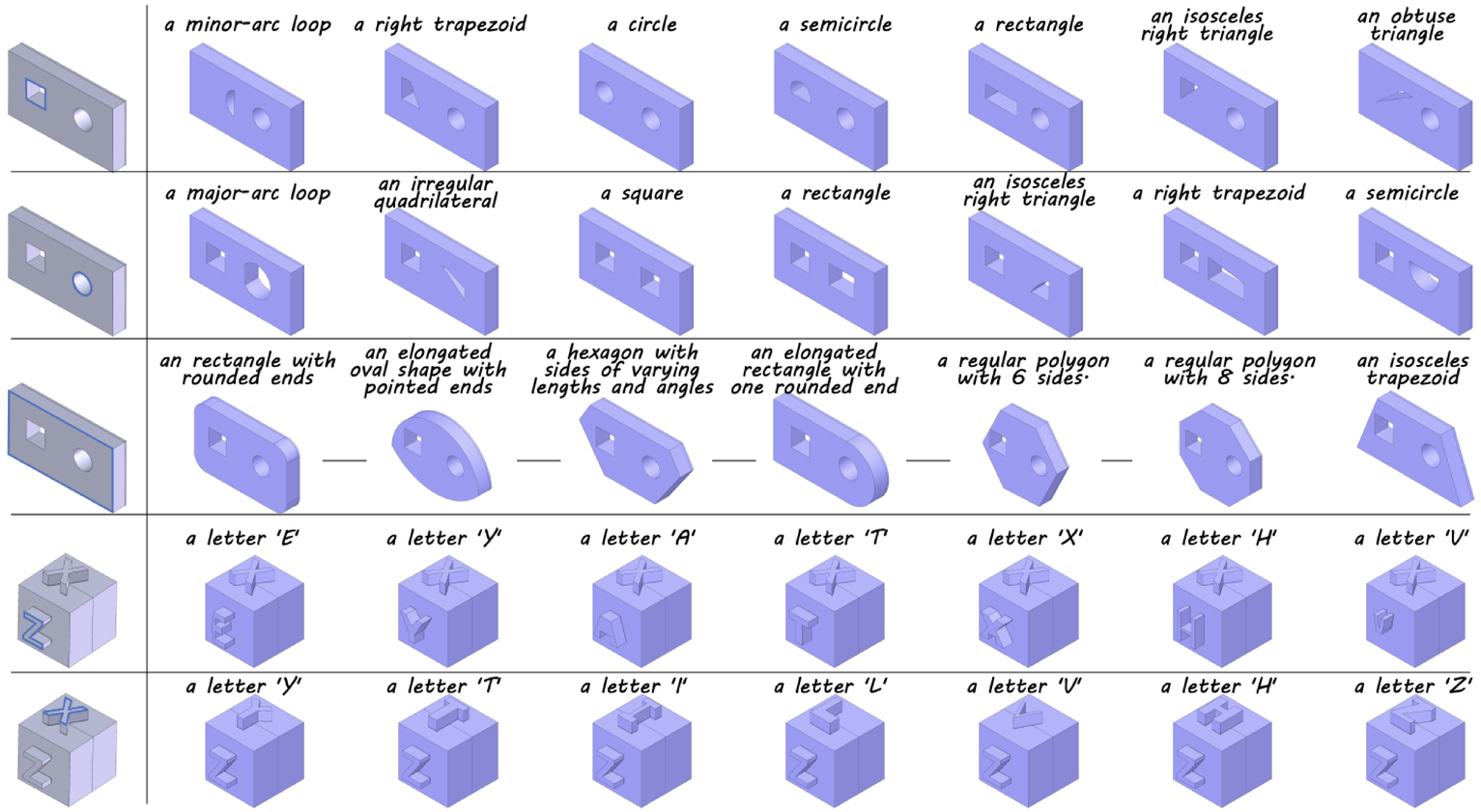


Figure 7: Additional qualitative results for GeoCAD. On the right, the upper section shows the user-defined instructions, while the lower section presents the newly generated CAD models.

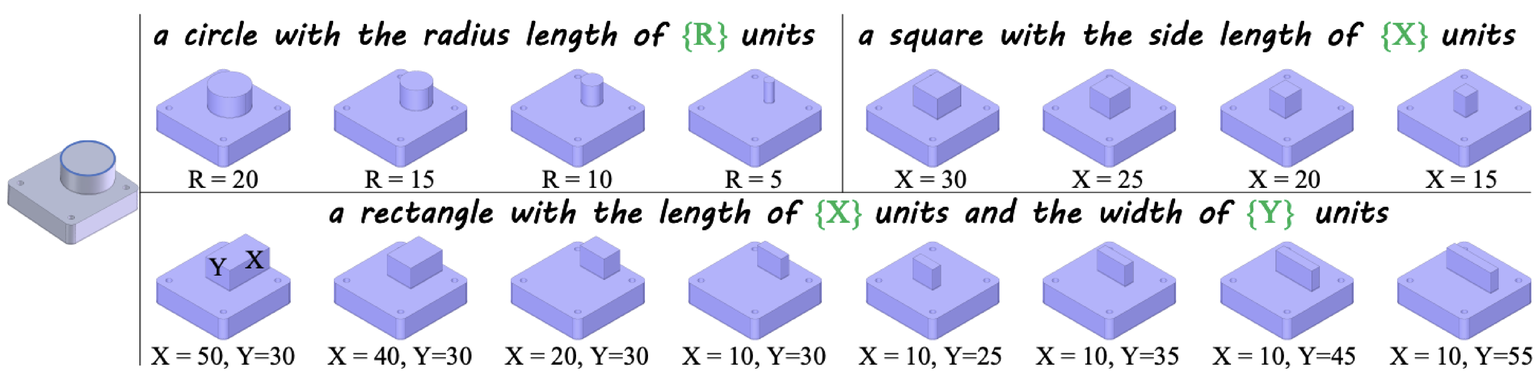


Figure 8: GeoCAD is capable of precisely controlling the key dimensional parameters. The right side displays the newly generated CAD models and the corresponding geometric instructions.

Experiments & Results

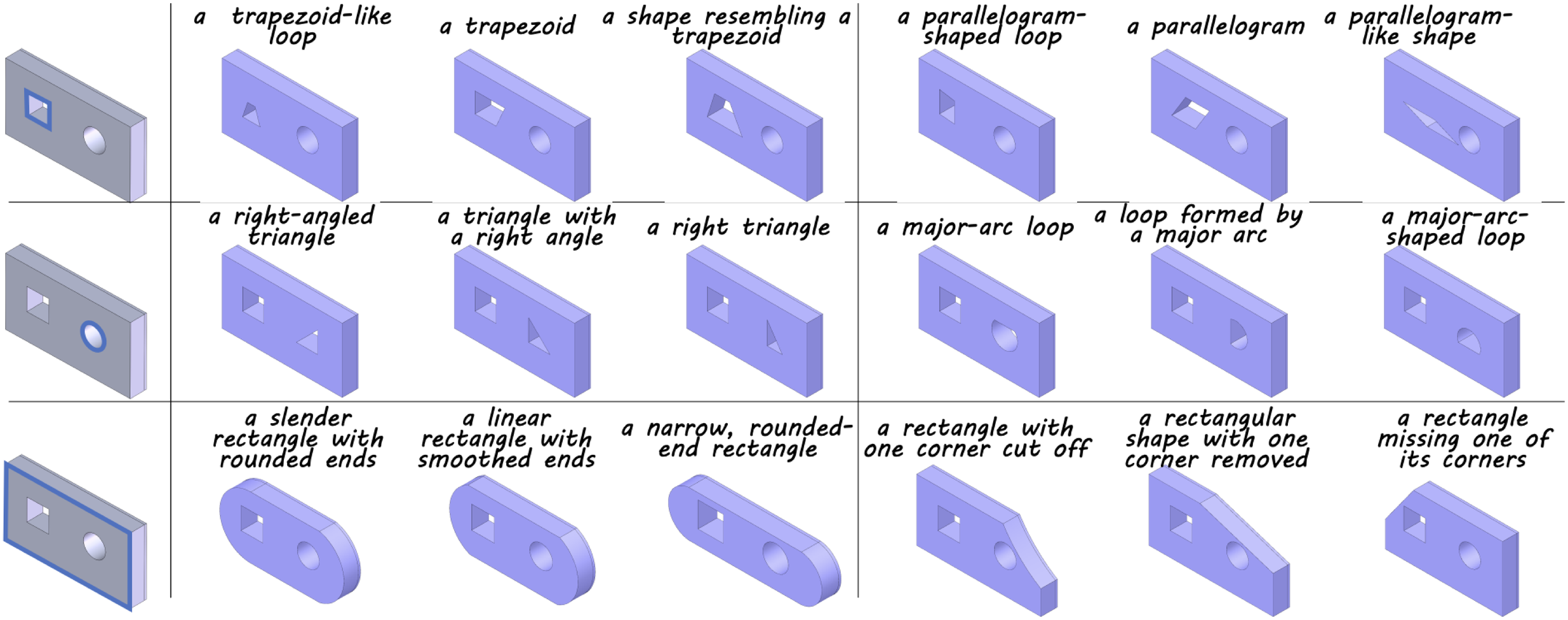


Figure 9: Generalization ability of GeoCAD. On the right, each row contains two groups, with each group comprising three examples generated based on semantically similar instructions.

Thank you!