ICCV OCT 19-23, 2025 HONOLULU HAWAII

CAD-Assistant: Tool-Augmented VLLMs as Generic CAD Task Solvers

Dimitrios Mallis[1], Ahmet Serdar Karadeniz[1], Sebastian Cavada[1], Danila Rukhovich[1], Niki Foteinopoulou[1], Kseniya Cherenkova[2], Anis Kacem[1], Djamila Aouada[1]

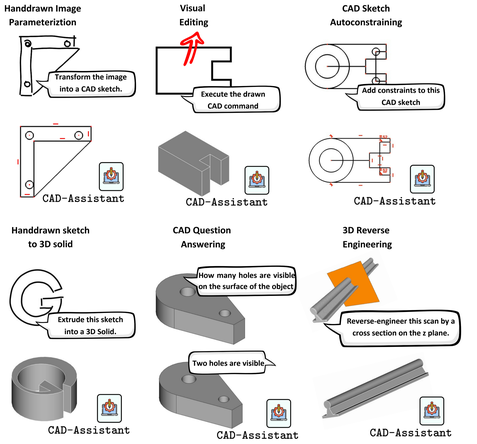
[1]SnT, University of Luxembourg [2]Artec3D, Luxembourg

https://cadassistant.github.io/

Introduction

CAD-Assistant is a tool-augmented VLLM framework for CAD.

Our method integrates FreeCAD and can processes **text** and **multimodal inputs,** such as sketches, drawn commands, and 3D scans.

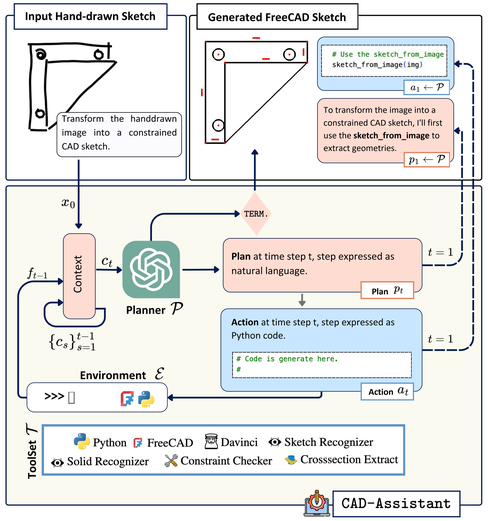


This figure showcases various examples of generic CAD queries and the responses generated by CAD-Assistant.

Contributions:

- **First** tool-augmented framework **for generic CAD tasks.**
- **Enhanced geometric reasoning** via CAD software integration.
- **Training-free** and **highly extensible** framework.
- **Zero-shot method** outperforming task specific baselines.
- Demonstrated potential on diverse **real-world use cases**.

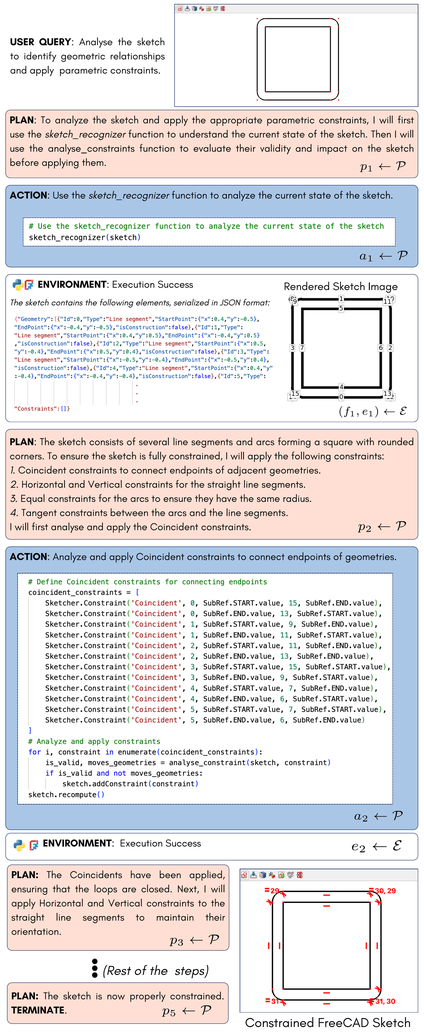
Method



Overview of CAD-Assistant framework.

- **A multimodal user request** is provided to a VLLM planner.
- The planner has access to a **Python environment equipped with CAD software** and **CAD specific tools**.
- The planner analyses the context and generates plan and action.
- Action is executed on the CAD environment.
- The environment **provides feedback** on the current design state.
- Process completes with the generation of the *TERMINATE* plan.

Autoconstraining Trajectory



Tool Use for Effective Geometric Reasoning

VLLMs have shown limited ability to geometrically comprehend CAD programs. We examine CAD representations **that can be derived using tools** and improve geometric reasoning.

Parametrization Strategy: Geometries can be represented by different parameter sets. *We compare the implicit parametrization to the point-based and over-parametrization.*

Serialization Strategy: We compare commonly used formats such as CSV, Markdown, HTML, and JSON.

Rendering-based Reasoning: We provide the VLLM planner with 2D renderings of the CAD sketch or 3D solid.

Text-based Reasoning

Serialization	Parametarization	2D Acc
Common CAD Sketch formats		
Serialized Graph	Implicit	0.674
DXF		0.671
OCA		0.707
Serialization Strategy (Tabular formats)		
CSV	Point-based	0.703
Markdown	Point-based	0.706
HTML	Point-based	0.710
Serialization Strategy (Schema-embedded formats)		
Serialized Graph	Point-based	0.744
JSON	Point-based	**0.748**
JSON	Overparametarized	0.747

Rendering-based Reasoning

CAD Sketch Image Type	
Hand-drawn Sketch	0.616
Precise Rendering	**0.754**

Evaluation on CAD Benchmarks

The CAD-Assistant can solve a wide range of tasks. We propose a CAD agent evaluation setting based on the following tasks:

- **CAD Question Answering**: Answering open-ended questions about a CAD sketches and 3D CAD models.
- **Autoconstraining**: Automatically inferring and applying parametric constraints to primitives in a CAD sketch.
- **Hand-drawn Sketch Parameterization**: Converting a hand-drawn sketch image into a parametric CAD sketch.

CAD Question Answering

Method	Planner	2D Acc	3D Acc
SGPBench [1]	GPT-4 mini	0.594	0.737
	GPT-4 Turbo	0.674	0.762
	GPT-4o	0.686	0.782
CAD-ASSISTANT	GPT-4 mini	0.614	0.783
	GPT-4 Turbo	0.741	0.825
	GPT-4o	**0.791**	**0.857**

CAD Sketch Autoconstraining

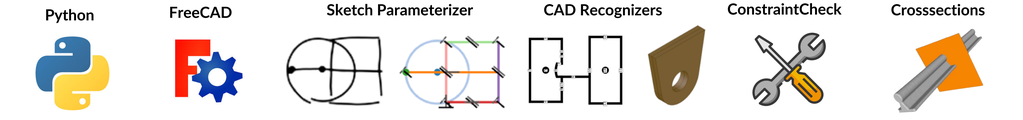
Method	Type	PF1 ↑	CF1 ↑
GPT-4o	*zero-shot*	0.693	0.274
Vitruvion [2]	*supervised*	0.706	0.238
CAD-ASSISTANT	*zero-shot*	**0.979**	**0.484**

Hand-drawn Sketch Image Parametrization

Method	Acc ↑	CD ↓
Vitruvion [2]	0.659	1.586
Davinci [3]	**0.789**	1.184
CAD-ASSISTANT	0.784	**0.680**

The CAD-Assistant is a **zero-shot method** that outperforms baselines and task-specific approaches trained on large datasets.

CAD Tools



Conclusions

We propose CAD-Assistant, a generic tool-augmented CAD agent using CAD-specific tools. We outperform task-specific methods and demonstrate the potential of tool-augmented VLLMs in real-world CAD workflows.

[1] Qiu, et al. "Can Large Language Models Understand Symbolic Graphics Programs?" ICLR (2025)
[2] Seff et al. "Vitruvion: A Generative Model of Parametric CAD Sketches." ICLR (2022)
[3] Karadeniz et al. "DAVINCI: A Single-Stage Architecture for Constrained CAD Sketch Inference." BMVC (2024)

Project Page

SCAN ME

Github

SCAN ME