Chair of Computational Modeling and Simulation
TUM School of Engineering and Design
Technical University of Munich

Using Multi-Modal Large Language Model for BIM/Bridge Point Cloud Data.

Project Characteristics

Modeling:
Mathematics:
Programming:
Science:

Task

This project aims to develop a point cloud-aware multi-modal language model (MLLM) by building a specialized point cloud encoder trained on BIM/bridge point cloud datasets and fine-tuning a language model.

- Develop a pipeline for handling and pre-processing bridge/BIM specific point cloud data.
- Conduct an in-depth review of existing research on multi-modal language models and their application to point cloud data, identifying relevant architectures and methods.
- Experiment with state-of-the-art architectures tailored for bridge/BIM point cloud data. This includes designing, fine-tuning, and implementing a point cloud encoder. Writing efficient, scalable code for this task will be critical.
- Finally develop a pipeline to fine-tune the language model and evaluate the results based on different classification and segmentation benchmarks.

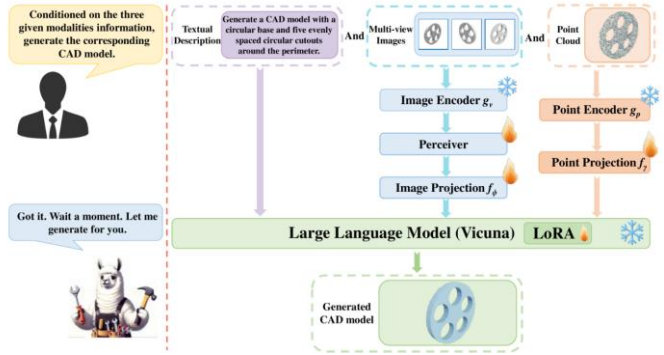

Figure 1: CAD-MLLM architecture. For Point cloud encoder. [1]

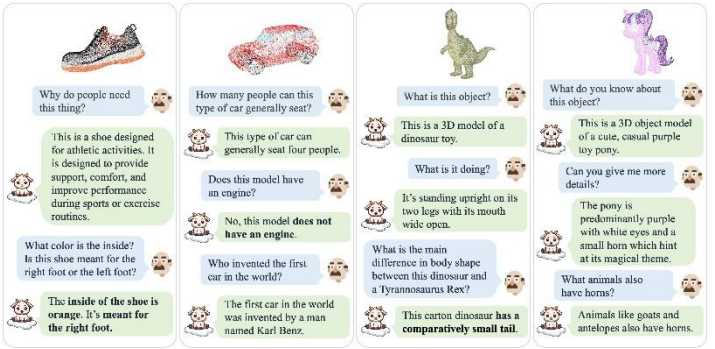

Figure 2: PointLLM [2]

[1] Xu, Jingwei, et al. "CAD-MLLM: Unifying Multimodality-Conditioned CAD Generation With MLLM." *arXiv preprint arXiv:2411.04954* (2024).

[2] Xu, Runsen, et al. "Pointllm: Empowering large language models to understand point clouds." *European Conference on Computer Vision*. Springer, Cham, 2025.