WORLD LAUREATES FORUM

世界顶尖科学家论坛

2026

国际工程智能大会

2026 International Conference on AI for Engineering

World Laureates Foundation

上海世界顶尖科学家发展基金会

工程智能研究院

INSTITUTE OF AI FOR ENGINEERING

# LED-BENCH: A Benchmark for Large-format Engineering Drawing Understanding

Ke Li, Tianyi Wang, Enning You, Chaoqun Wan, Zhen Huang, Shuxian Liang, Xian-sheng Hua

Institute of AI for Engineering, Tongji University

## 01 ABSTRACT

Real-world engineering drawings contain dense annotations and evidence scattered across views. Identifying objects, reconstructing relations, and integrating this evidence remain challenging for current MLLMs (Multimodal Large Language Models).

**Research question**

How well do MLLMs understand real-world, large-format engineering drawings?

## 02 BENCHMARK

LED-BENCH evaluates engineering objects, relations, and distributed evidence on complete drawings from three engineering domains.

| 107 | 377 | 3 |
|---|---|---|
| Large-format drawings | Questions | engineering domains |

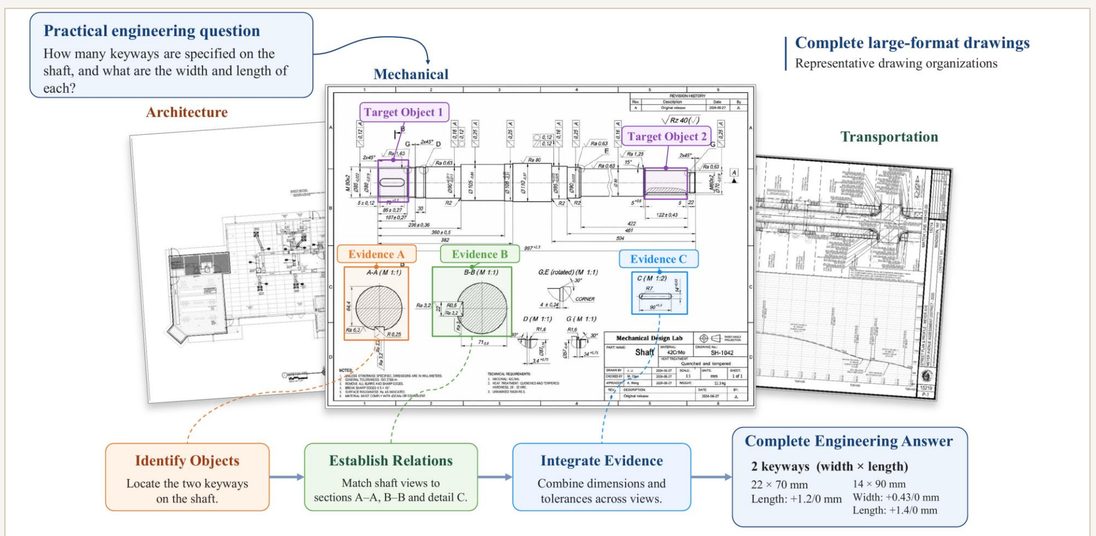



Figure1. Overview of LED-BENCH

**Identify objects → Establish relations → Integrate evidence**

## 03 RESULTS

Extensive evaluation of 11 MLLMs: 8 proprietary and 3 open-weight models.

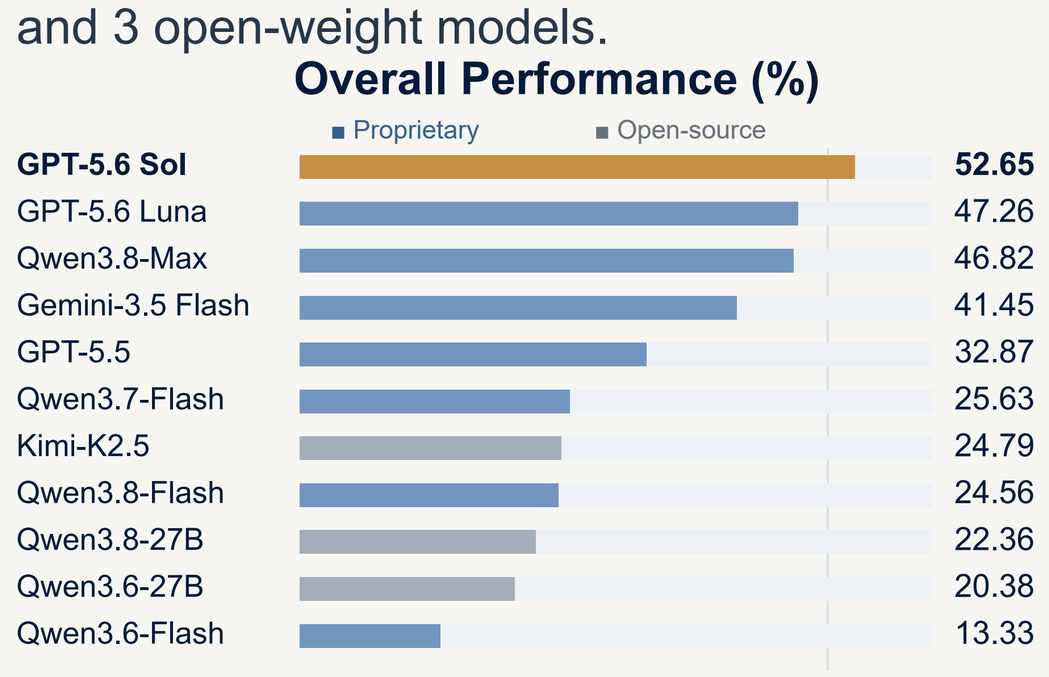
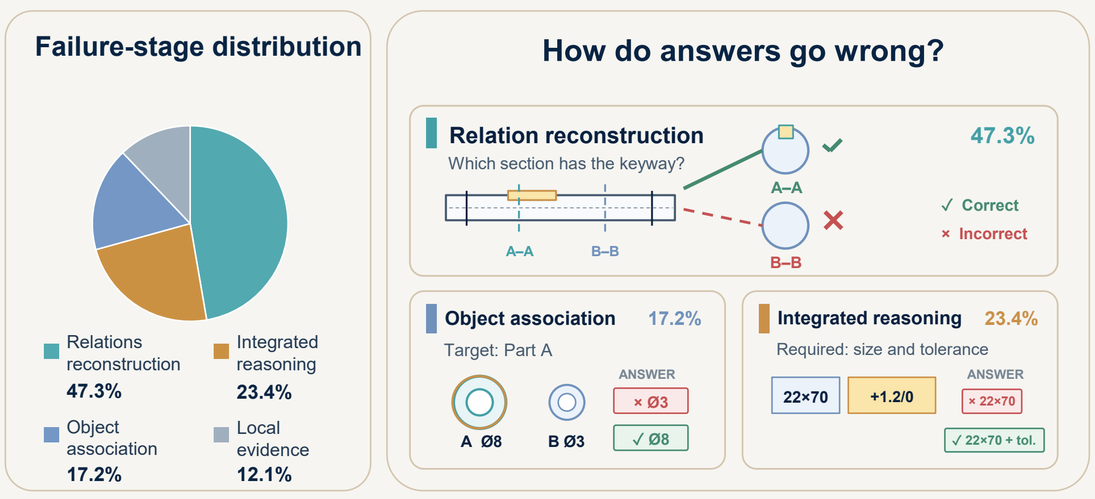



Figure2. Performance across 11 MLLMs

Table1. Top 3 models across engineering domains (%)

| Model | Mech. | Arch. | Transp. |
|---|---|---|---|
| Qwen3.8-Max | 46.36 | 48.00 | 46.09 |
| GPT-5.6 Luna | 50.38 | 47.11 | 44.27 |
| **GPT-5.6 Sol** | **58.81** | **48.89** | **50.26** |

### Failure analysis




## 04 DISCUSSION

Failures most often involve reconstructing relations across views or integrating multiple engineering requirements into a complete answer.

## 05 CONCLUSIONS

LED-Bench reveals substantial gaps in MLLMs' ability to associate and integrate evidence across large-format engineering drawings.

**CONTACT**

Shuxian Liang | sxliang@tongji.edu.cn

**REFERENCES / ACKNOWLEDGEMENTS**

LED-BENCH: A Benchmark for Large-format Engineering Drawing Understanding. arXiv:2610.19571. https://arxiv.org/abs/2610.19571