

Performance

100%

80%

60%

40%

20%

0%

2024

2025

2026

2027

2028

2029

2030

- GeoBench
- CAD-Eval
- Visual Task Assessment (VISTA)
- VPCT
- AudioMultiChallenge
- VisualToolBench

