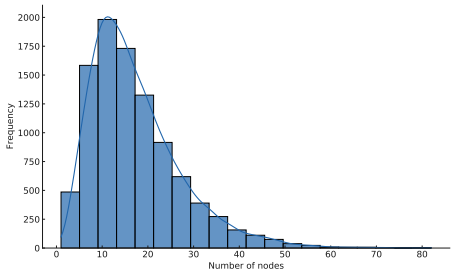


Figure 1: Distribution of node number in CSQA.

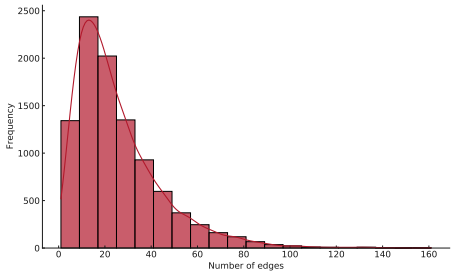


Figure 2: Distribution of edge number in CSQA.

Table 1: Performance of LLaVA-v1.6 before and after fine-tuning on each graph comprehension task. N. denotes node and E. denotes Edge. For node description and triple listing, we consider the average accuracy of each test example. We use exact matching to determine the accuracy, which may be a stricter evaluation.

Model	N. description	N. degree	Highest N. degree	N. number	E. number	Triple listing
Original	1.4	15.3	3.3	16.7	9.7	0.6
After fine-tuning	12.8 (+11.4)	27.0 (+11.7)	11.6 (+8.3)	27.5 (+10.8)	16.2 (+9.7)	8.2 (+7.6)

Table 2: Performance of GraphVis compared with different baselines.

	CSQA (%)
Base LVLM	70.5
Base w/ FT	76.3
KAPING	67.7
KAPING w/ FT	78.1
GraphVis	**82.8**

Table 3: Performance of LLaVA-v1.6 on ScienceQA compared with GraphVis and GraphVis (joint fine-tuning).

	ScienceQA (%)
Base LVLM	68.86
GraphVis (Joint)	71.94
GraphVis	**73.18**

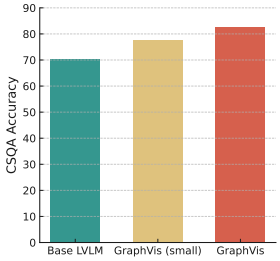


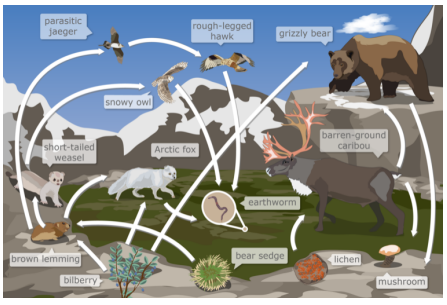
Figure 3: Effect of lower resolution graph images.

Table 4: Effect of synthetic question format number.

Format number	1	5
CSQA	81.9	82.8

Table 5: Effect of different visualization.

	CSQA (Set 1)	CSQA (Set 2)
GraphVis	82.8	82.6



Query: Context: Below is a food web from a tundra ecosystem in Nunavut, a territory in Northern Canada. A food web models how the matter eaten by organisms moves through an ecosystem. The arrows in a food web represent how matter moves between organisms in an ecosystem.
Which of these organisms contains matter that was once part of the lichen?
A. mushroom B. short-tailed weasel C. brown lemming
D. rough-legged hawk E. bilberry
Answer with the option's letter from the given choices directly.

Base (LLaVA-v1.6 7B): C **GraphVis (LLaVA-v1.6 7B)**: A

Figure 4: Example of model output on ScienceQA (VQA task). Note that after finetuned the base LVLM with GraphVis on CSQA with synthetic KG images, the model can successfully traverse the graph to locate the correct answer.