Int4 Quantization Of Llava-Uhd Impact Accuracy On Seed-Bench Visual Reasoning Subtasks Compared To Fp16 Precision

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: How does INT4 quantization of LLaVA-UHD impact accuracy on SEED-Bench visual reasoning subtasks compared to FP16 precision. Recent advancements in Chain of Thought (COT) generation have significantly improved the reasoning capabilities of Large Language Models (LLMs), with reinforcement learning (RL) emerging as an effective post-training approach. Multimodal Large Language Models (MLLMs) inherit. 17 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Exploring the Effect of Reinforcement Learning on Video Understanding: Insights from SEED-Bench-R1. Research question: How does INT4 quantization of LLaVA-UHD impact accuracy on SEED-Bench visual reasoning subtasks compared to FP16 precision?.

2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.2/10.

3 Results

12 papers retrieved. 17 claims extracted; 2 independently verified. Quality review score: 4.2/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
SEED-Bench-R1 consists of a training set containing 50,269 samples.	×	0.06
The SEED-Bench-R1 training set focuses on the 'Daily life' domain.	×	0.06
The SEED-Bench-R1 Val-L1 validation set contains 2,432 samples focused on 'Daily life'.	×	0.04
The SEED-Bench-R1 Val-L1 split does not include cross-environment or cross-task evaluations.	×	0.11
The SEED-Bench-R1 Val-L2 validation set contains 923 samples and includes cross-environment evaluation but not cross-tas	×	0.09
The SEED-Bench-R1 Val-L3 validation set contains 1,321 samples covering Hobbies, Daily life, Recreation, and Work domain	×	0.04
The SEED-Bench-R1 Val-L3 split includes both cross-environment and cross-task evaluations.	×	0.13
In the provided L1 example, the correct action to add cream cheese to the soup is to throw away the cream cheese contain	×	0.01
In the provided L2 example, the correct action to prepare a flour mixture is to pour milk into the flour bowl.	×	0.01
In the provided L3 example, the correct action to take measurements around the room is to mark a point on the wall.	×	0.01
The study uses Qwen2-VL-Instruct-7B as the base model.	✓	0.17
The study adopts GRPO as the representative reinforcement learning algorithm.	×	0.05
Both the RL and SFT experiments utilized 6,000 out of the 50,000 training samples from SEED-Bench-R1.	×	0.09
The maximum number of sampled frames per input video was limited to 16 for training efficiency.	×	0.03
The frame resolution used in the study was 25.	×	0.03
Reinforcement learning (specifically GRPO) outperforms supervised fine-tuning (SFT) in data efficiency and generalizatio	✓	0.18
The RL-trained model exhibits strong generalization capabilities on the LongVideoBench benchmark.	×	0.07

References

- http://arxiv.org/abs/2307.16125v2
- http://arxiv.org/abs/2503.24376v1
- http://arxiv.org/abs/2311.17092v1