

## Responsible NLP Checklist

Paper title: *Chain-of-Thought Compression Should Not Be Blind: V-Skip for Efficient Multimodal Reasoning via Dual-Path Anchoring*

Authors: *Dongxu Zhang, Yiding Sun, Cheng Tan, Wenbiao Yan, Ning Yang, Jihua Zhu, Haijun Zhang*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- N/A the authors indicated that the question does not apply to their work
- the authors did not respond to the checkbox question

For background on the checklist and guidance provided to the authors, see the [Responsible NLP Checklist](#) page at ACL Rolling Review.

---

### A. Questions mandatory for all submissions.

#### A1. Did you describe the limitations of your work?

*This paper has a Limitations section.*

#### A2. Did you discuss any potential risks of your work?

*The paper focuses on efficiency optimization and mitigating "Visual Amnesia" in existing multimodal large language models. While we discussed technical limitations such as dependence on cross-modal alignment and offline distillation overhead in the Limitations section, this work does not introduce new societal risks, potential for malicious use, or significant ethical hazards beyond those already inherent to the foundational models evaluated.*

### B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

#### B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?

*This research exclusively utilizes widely adopted, publicly available academic benchmark datasets (MMMU and DocVQA). These datasets have been previously curated, anonymized, and vetted by their respective creators. We did not collect or process any new data containing personally identifying information (PII) or offensive content.*

#### B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?

*Section B.3 (Dataset Splits and Evaluation) explicitly states that we strictly utilized the official validation splits for both the MMMU and DocVQA benchmarks for all evaluations.*

### C. Did you run computational experiments?

#### C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?

*The experimental setup is outlined in Section 4.1. Detailed hyperparameter configurations for the LoRA fine-tuning process (e.g., rank, alpha, learning rate, batch size) are provided in Appendix B.4 (Training Configurations). Furthermore, Appendix C.2 provides a detailed hyperparameter search and robustness analysis specifically regarding the LoRA Rank.*

*The Responsible NLP Checklist used at ACL Rolling Review is adopted from NAACL 2022, with the addition of ACL 2023 question on AI writing assistance and further refinements based on ARR practice. ACL 2026 used a subset of ARR checklist form.*

- C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?

*Due to the prohibitive computational cost of evaluating large-scale multimodal models (up to 72B parameters), the results reported in our tables (e.g., Accuracy, ANLS) represent standard metric scores obtained on the official validation splits. We did not perform multiple runs with different random seeds to report error bars or standard deviations.*

**D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?**

- D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

*(left blank)*

- D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

*(left blank)*

- D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?

*(left blank)*

- D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

*(left blank)*

**E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?**

- E1. If you used AI assistants, did you include information about their use?

*We used generative AI assistants to assist with grammar correction, phrasing refinement, and general proofreading of the manuscript text during the writing process.*