

## Responsible NLP Checklist

Paper title: *SAM3-I: Segment Anything with Instructions*

Authors: *Jingjing Li, Yue Feng, Yuchen Guo, Jincan Huang, Wei Ji, Qi Bi, Yongri Piao, Miao Zhang, Xiaoqi Zhao, Qiang Chen, Shihao Zou, Huchuan Lu, Li Cheng*

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.*

- N/A* A2. Did you discuss any potential risks of your work?

*This work focuses on instruction-following image segmentation using publicly available datasets and does not introduce new application domains or deployment scenarios beyond existing segmentation models. It does not involve personal data, decision-making about individuals, or safety-critical applications. As such, we do not identify significant potential risks beyond those commonly associated with general-purpose vision models.*

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

- N/A* 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?

*HMPL-Instruct does not introduce new personally identifiable information. All images originate from existing public datasets and retain their original privacy and ethical characteristics. No additional personal data is collected, inferred, or annotated in the construction process (Section Ethical Considerations).*

- 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?

*Table 1*

### C. Did you run computational experiments?

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

*Section 6 and Appendix C*

- 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?

*All reported results are obtained from deterministic single-run evaluations on fixed benchmarks using*

---

*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.*

standard segmentation metrics (gIoU and P@50). We do not report error bars or variance since the evaluation protocol does not involve repeated runs with different random seeds, and the results are directly comparable across methods under the same settings.

**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.?

*Annotation was conducted internally by the authors or project collaborators as part of dataset construction, rather than a human-subject study involving external participants. Therefore, no participant-facing experimental instructions or risk disclosures were required.*

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)?

*Human annotation was conducted internally by the authors or project collaborators using research funding. No external recruitment, crowdsourcing platforms, or participant compensation were involved.*

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)?

*The dataset is constructed by augmenting existing publicly available datasets (e.g., PACO-LVIS) that were originally collected and released under appropriate licenses. No new personal data was collected, and no additional consent beyond the original dataset licenses was required.*

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

*This work does not involve human-subject research as defined by institutional ethics guidelines. Therefore, ethics review board approval was not required.*

**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?

*No AI assistance was involved in coding or research idea generation.*