

Responsible NLP Checklist

Paper title: *See More, Store Less: Memory-Efficient Resolution for Video Moment Retrieval*

Authors: *Mingyu Jeon, Sungjin Han, Jinkwon Hwang, Minchol Kwon, Jonghee Kim, Junyeong Kim*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- 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?

We discussed technical limitations in Section 6, but we did not anticipate any significant ethical risks or negative societal impacts associated with this work, as it focuses on a memory-efficient framework for video retrieval.

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

B1. Did you cite the creators of artifacts you used?

We cited the creators of the datasets (QVHighlights, Charades-STA, ActivityNet-Captions) in Section 4 and the models (Flan-T5, InstructBLIP, BLIP2, CLIP) in Section 4.1.

B2. Did you discuss the license or terms for use and/or distribution of any artifacts?

We utilized publicly available datasets and open-source models which are widely used in the research community, but we did not explicitly discuss their specific license terms in the paper.

B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?

We employed widely recognized public benchmarks for their designed purpose of video moment retrieval and did not include a specific discussion on intended use in the paper.

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?

We utilized established public datasets that are standard in the field and did not perform a separate analysis for personally identifying information or offensive content.

B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?

We described the characteristics and statistics of the datasets (QVHighlights, Charades-STA, ActivityNet-Captions) in Section 4 under the 'Datasets' heading.

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.

- 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?

We reported relevant statistics, such as the number of videos and annotations for QVHighlights, Charades-STA, and ActivityNet-Captions, in Section 4 under the 'Datasets' heading.

C. Did you run computational experiments?

- C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?

We detailed the number of trainable parameters, the computing infrastructure (A6000 GPUs), and the total computational budget (GPU hours) in Appendix A.

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

We described the experimental setup in Section 4.1 and Appendix A, and provided a detailed hyperparameter analysis and the best-found values in Appendix B.

- 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?

We reported the final performance metrics (Recall@K, mAP) on the standard test/validation sets for our proposed method, consistent with the evaluation protocols of prior works in VMR.

- C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation, such as NLTK, SpaCy, ROUGE, etc.), did you report the implementation, model, and parameter settings used?

We provided detailed parameter settings and implementation details for the packages and models used (e.g., Flan-T5, LoRA, InstructBLIP, CLIP) in Section 4.1 and Appendix A.

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.?

N/A

- 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)?

N/A

- 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)?

N/A

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

N/A

- D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?

N/A

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?

N/A