

Responsible NLP Checklist

Paper title: *CLIP-UP: A Simple and Efficient Mixture-of-Experts CLIP Training Recipe with Sparse Upcycling*

Authors: *Xinze Wang, Chen Chen, Yinfei Yang, Hong-You Chen, Bowen Zhang, Aditya Pal, Xiangxin Zhu, Xianzhi Du*

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 did not explicitly discuss potential risks because our work focuses on improving model efficiency using existing publicly available datasets without collecting new sensitive data or developing new capabilities that pose direct ethical concerns.

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

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

3

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

We used only publicly available models and datasets (CLIP, COCO, and ImageNet) under their standard academic licenses, but due to space constraints and the common usage of these resources, we did not explicitly discuss the licenses 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 used all existing artifacts strictly for research purposes, consistent with their intended use. We did not explicitly mention this in the paper due to space constraints

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 used publicly available datasets (COCO, ImageNet) that do not explicitly contain personally identifying information as per their documentation. Our work did not collect new data or handle offensive content directly, so no additional anonymization steps were taken.

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.

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

We used standard publicly available datasets (COCO, ImageNet) and the pretrained CLIP model, which have extensive documentation publicly available. Due to space constraints, we did not include detailed dataset documentation in the paper.

- 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?

3

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?

3

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

3

- 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?

3

- 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 did not use any external packages.

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)

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

(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 AI tools solely for grammar correction and did not consider the changes substantive enough to warrant an explicit acknowledgment in the manuscript.