Review of: "Harnessing Self-Supervision in Unlabelled Data for Effective World Representation Learning in AI Models"

Shigang Wang[1]

1 Northwestern Polytechnical University

This paper proposes a method for contrastive self-supervised learning on unlabeled images. Through comparative experimental results, it is shown that this method enables the model to achieve superior representation learning compared to supervised learning, especially under scarce labeled data conditions. The research topic of this study is novel, but the current results are still preliminary. It is recommended that the author further enrich the experiments to demonstrate the effectiveness of the method. After summarizing the key points, my specific comments are as follows.

1. The method description in the paper lacks details. It is recommended to supplement more experimental details, such as network architecture, loss function, etc.

2. The experimental results in this paper are few. It is suggested to increase comparison experiments with other supervised methods and self-supervised methods to demonstrate the advantages of the method proposed in this paper.

3. It is recommended to combine the method in this paper with efficient downstream tuning and multi-task representation learning for transfer evaluation, in order to demonstrate the validity of this method. For example, the statement in the paper that fine-tuning can efficiently adapt to downstream tasks is reasonable. I suggest supplementing more quantitative comparative experimental results, which will make this part of the statement more complete.