

Peer Review

Review of: "k2SSL: A Faster and Better Framework for Self-Supervised Speech Representation Learning"

Jian Wang¹

¹. Nanjing University of Aeronautics and Astronautics, China

The paper improved the HuBERT model for Automatic Speech Recognition with ZipFormer instead of Transformer.

1. And, the authors emphasized efficiency and effectiveness. How do you evaluate the scheme in efficiency and effectiveness respectively?
2. You should explain why ZipFormer outperforms Transformer.
3. Perhaps, this improvement may have some negatives. You should give some analyses and discussions.

Declarations

Potential competing interests: No potential competing interests to declare.