Are SSMs better than Transformers at Reasoning?

author names withheld

Abstract

Place abstract here

1. Main

Main text limited to 4 pages, excluding references and appendices.

1.1. Citations (`natbib package`)

Attention [1] is pervasive in deep learning.

References

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.