Review of: "Enhancement of Network Architecture Alignment in Comparative Single-Cell Studies"

Rehan Ali Khan[1,2]

1 Department of Electrical Engineering, University of Science & Technology Bannu, Pakistan
2 College of Electrical Engineering, Zhejiang University, China

Potential competing interests: No potential competing interests to declare.

Report

I believe this paper has potential for publication, subject to minor revisions.

The paper introduces scSpecies, a novel deep learning-based approach to align single-cell RNA sequencing (scRNA-seq) datasets across species. The proposed methodology addresses key challenges in cross-species data analysis, such as non-orthologous genes and variations in gene expression patterns. Below is a structured review of the paper, with comments on its strengths, areas for improvement, and suggestions for further exploration.

The step-by-step explanation of the workflow, including pretraining, architecture transfer, fine-tuning, and latent space alignment, is clear and easy to follow.

Comments and Suggestions:

1. The model struggles with identifying cell types unique to the target dataset, instead aligning them with similar cell types in the context dataset. This could result in misinterpretation of species-specific cell populations, which is a critical limitation for studies focused on unique cell types.

Suggestion: Incorporate a mechanism to flag and cluster potential unique cell types during the alignment process.

1. While the paper highlights improvements in clustering metrics, it does not provide detailed quantitative comparisons with competing methods (e.g., Cell BLAST, scArches) on shared benchmarks.

Suggestion: Include a comprehensive evaluation table comparing scSpecies with existing methods across datasets and metrics.

1. The paper acknowledges that scSpecies currently handles only scRNA-seq data and does not support multimodal datasets such as CITE-seq (RNA and protein data).

Suggestion: Extend the model to accommodate multimodal datasets for broader applicability in integrative biology.

1. The computational requirements for aligning large datasets are not discussed. Scalability is a concern when dealing with high-dimensional data.

Suggestion: Provide benchmarks for runtime and memory usage compared to existing methods.

1. The paper presents a well-structured and innovative approach to cross-species scRNA-seq data alignment. While the methodology is robust and addresses several existing challenges, limitations related to cell type identification, context dataset dependency, and scalability need further exploration. With improvements and extensions, **scSpecies** has the potential to become a valuable tool for the biomedical research community, bridging the gap between model organisms and human biology.

This work is a significant step forward, and with some refinements, it could have far-reaching implications for advancing translational research.