

ivis: dimensionality reduction in very large datasets using Siamese Networks

Benjamin Szubert¹ and Ignat Drozdov¹

¹ Bering Limited

DOI: [10.21105/joss.01596](https://doi.org/10.21105/joss.01596)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 18 July 2019

Published: 06 August 2019

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

ivis is a dimensionality reduction technique that implements a Siamese Neural Network architecture trained using a novel triplet loss function. Results on simulated and real datasets demonstrate that *ivis* preserves global data structures in a low-dimensional space and adds new data points to existing embeddings using a parametric mapping function.

ivis is easily integrated into standard machine learning pipelines through a scikit-learn compatible API and scales well to out-of-memory datasets. Both supervised and unsupervised dimensionality reduction modes are supported.

Further information on the algorithm and its application to single cell datasets can be found in (Szubert, 2019). Implementation of the *ivis* algorithm is available on [GitHub](#).

Acknowledgements

This work was supported by funding from the European Commission's Seventh Framework Programme [FP7-2007-2013] under grant agreement n°HEALTH-F2-2013-602114 (Athero-B-Cell).

References

Szubert, J., B. Cole. (2019). Structure-preserving visualisation of high dimensional single-cell datasets. *Scientific Reports*, 9, 2045–2322. doi:[10.1038/s41598-019-45301-0](https://doi.org/10.1038/s41598-019-45301-0)