

FMSNE: FAST MULTI-SCALE NEIGHBOUR EMBEDDING IN R

fmsne: fast multi-scale neighbour embedding in R

Author(s): Laurent Gatto, Cyril de Bodt

Affiliation(s): UCLouvain, Belgium

Social media: <https://fosstodon.org/@lgatto> (<https://fosstodon.org/@lgatto>)

Dimensionality reduction (DR) has been a workhorse of large scale, multivariate omics data analysis from the early days. Since the advent of single-cell RNA sequencing, non-linear approaches have taken the front stage, with t-distributed stochastic neighbour embedding (t-SNE) [1,2] being one of, if not the main player. Packages such as `Rtsne` [3] and `scater` [4] have made it easy to apply t-SNE in R/Bioconductor workflows. One sticking point with t-SNE is the single perplexity parameter, that controls the number of nearest high-dimensional (HD) neighbours that are taken into account when constructing the low-dimensional (LD) embedding: small (resp. large) values only enable preserving small (resp. large) neighbourhoods from HD to LD during DR, impairing the reproduction of large (resp. small) neighbourhoods. It is thus a key parameter, especially if the LD embedding is used for interpretation, which is often the case in omics-based applications. Multi-scale neighbour embedding [5] is an extension to single-scale approaches such as t-SNE, that exempt users from having to set a single perplexity (scale) arbitrarily. Multi-scale approaches maximise the LD embedding quality at all scales, preserving both local and global HD neighbourhoods [6]. They have been shown to better capture the structure of data and to significantly improve DR quality [7]. Here, we present `fmsne` (<https://github.com/lgatto/fmsne>), an R package that relies on the `basiliks` package [8] to provide Bioconductor-friendly interface to fast multi-scale methods implemented in python. `fmsne` implements fast multi-scale functions such as `runFMSTSNE()` and `plotFMSTSNE()`, based on `scater`'s `scater::run*()` and `scater::plot*()` interface [4]. It also exposes the `drQuality()` function to assess DR quality using rank-based criteria [7]. Finally, we illustrate fast multi-scale methods on various single-cell datasets. [1] van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(Nov), 2579-2605. [2] van der Maaten, L. (2014). Accelerating t-SNE using tree-based algorithms. *Journal of Machine Learning Research*, 15(1), 3221-3245. [3] Jesse H. Krijthe (2015). *Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut Implementation*, URL: <https://github.com/jkrijthe/Rtsne>

[4] McCarthy DJ, Campbell KR, Lun ATL, Willis QF (2017). Scater: pre-processing, quality control, normalisation and visualisation of single-cell RNA-seq data in R. *_Bioinformatics_*, 33, 1179-1186. doi:10.1093/bioinformatics/btw777 [5] C. de Bodt, D. Mulders, M. Verleysen and J. A. Lee, 'Fast Multiscale Neighbor Embedding,' in *_IEEE Transactions on Neural Networks and Learning Systems_*, 2020, doi: 10.1109/TNNLS.2020.3042807. [6] Lee, J. A., Peluffo-Ordóñez, D. H., & Verleysen, M. (2015). Multi-scale similarities in stochastic neighbour embedding: Reducing dimensionality while preserving both local and global structure. *_Neurocomputing_*, 169, 246-261. [7] Lee, J. A., & Verleysen, M. (2009). Quality assessment of dimensionality reduction: Rank-based criteria. *_Neurocomputing_*, 72(7-9), 1431-1443. [8] Lun ATL (2022). basilisk: a Bioconductor package for managing Python environments. *_Journal of Open Source Software_*, 7, 4742. doi:10.21105/joss.04742.

TAGS

🔗 [BIRDS-OF-A-FEATHER \(/TAGS/BIRDS-OF-A-FEATHER\)](/TAGS/BIRDS-OF-A-FEATHER)

🔗 [PACKAGE-DEMO \(/TAGS/PACKAGE-DEMO\)](/TAGS/PACKAGE-DEMO)

🔗 [POSTER \(/TAGS/POSTER\)](/TAGS/POSTER)

🔗 [POSTERS \(/TAGS/POSTERS\)](/TAGS/POSTERS)

🔗 [SHORT-TALK \(/TAGS/SHORT-TALK\)](/TAGS/SHORT-TALK)

🔗 [SHORT-TALKS-TRACK \(/TAGS/SHORT-TALKS-TRACK\)](/TAGS/SHORT-TALKS-TRACK)

ABOUT US

Bioconductor conference highlights current developments within and beyond the Bioconductor project.

CONTACT

eurobioc@embl.org

Copyright (c) 2020 - 2023, Bioconductor.

Template by Bootstrapious (<https://bootstrapious.com/p/universal-business-e-commerce-template>). Ported to Hugo by DevCows (<https://github.com/devcows/hugo-universal-theme>).

This site is powered by Netlify (<https://www.netlify.com/>).

