Interactive Visual Analysis of Mass Cytometry Data by Hierarchical Stochastic Neighbor Embedding Reveals Rare Cell Types

Thomas Höllt[1,2], Vincent van Unen[3], Nicola Pezzotti[1], Na Li[3], Marcel Reinders[4], Elmar Eisemann[1], Frits Koning[3], Anna Vilanova[1], Boudewijn Lelieveldt[4,5]
[1]Computer Graphics and Visualization Group, TU Delft, the Netherlands; [2]Computational Biology Center, Leiden University Medical Center, the Netherlands; [3]Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, the Netherlands
[4]Pattern Recognition and Bioinformatics Group, TU Delft, the Netherlands; [5]Division of Image Processing, Department of Radiology, Leiden University Medical Center, the Netherlands

Mass cytometry allows high-resolution dissection of the cellular composition of the immune system. However, the high-dimensionality, large size, and non-linear structure of the data poses considerable challenges for data analysis.

We introduce Hierarchical Stochastic Neighbor Embedding (HSNE) for single-cell analysis, a computational approach that constructs a hierarchy of non-linear similarities, allowing the analysis of millions of cells via different levels of detail up to single-cell resolution within minutes. We integrated HSNE into the Cytosplore [+HSNE] framework to facilitate interactive exploration and analysis of the hierarchy by a set of corresponding two-dimensional plots with stepwise increase in detail up to the single-cell level. This divide and conquer approach minimizes computation time and, thereby, allows efficient and interactive visualization.

We validated the discovery potential of Cytosplore[+HSNE] by re-analyzing a recent study on gastrointestinal disorders as well as two other publicly available mass cytometry datasets. We found that Cytosplore[+HSNE] efficiently identifies both abundant and rare cell populations, without resorting to downsampling of the data, including rare cell populations that were missed in a previous analysis due to downsampling. Taken together, Cytosplore [+HSNE] offers unprecedented possibilities for visual exploration and analysis of millions of cells measured in mass cytometry studies.