scCAN: A SINGLE CELL CLUSTERING METHOD USING AUTOENCODER AND NETWORK FUSION

Bang Tran*, Duc Tran*, Ha Nguyen*, Tin Nguyen*
*Department of Computer Science and Engineering, University of Nevada, Reno

Abstract:

Motivation: Single-cell RNA-seq (scRNA-Seq) is the latest high throughput sequencing technology that shifted genomics research from the analysis of bulk tissues toward a comprehensive characterization of individual cells. Unlike the traditional technology, scRNA-Seq genes expression can be used to study several different tissues and organs at different stages. Therefore, identifying the cell types present in the sample from the single cell transcriptome data is essential to study cell heterogeneity. Several methods have been developed to address this problem. However, identification and characterization of cell types remains challenging. Here, we propose scCAN, a completed pipeline using autoencoder and network-based clustering method to identify cell population scRNA-seq data. scCAN is a fast and accurate method to segregate cell types from single cell transcriptome data. In extensive analysis, we measure the performance of the scCAN on 26 publicly available scRNA-seq datasets with known cell types and simulation datasets.

Results: We compare the proposed method with five other existing methods: CIDR, SEURAT3, Monocle3, SINCERA, SHARP, and SCANPY. The results show that scCAN method outperforms these methods.