[BIOINFO 2025]

scFNCCL: Dual Contrastive Learning with False-negative correction at Cell level for single-cell RNA sequence Clustering

Gunho Choi[1], and Minsik Oh[2,*]

[1]Department of Data Technology, Myongji University
[2]Department of Convergence Software, Myongji University
*Corresponding author: msoh@mju.ac.kr

Single-cell RNA sequencing (scRNA-seq) has enabled the characterization of cellular processes at single-cell resolution, but its high dimensionality, sparsity, and technical noise present major challenges for analysis. Clustering remains essential for identifying cell types and uncovering intrinsic cellular patterns, and recent studies have increasingly adopted contrastive learning frameworks to improve representation learning. Dual contrastive learning—which integrates instance-level and cluster-level objectives—can capture both cell–cell differences and inter-cluster heterogeneity. However, conventional InfoNCE formulations treat all samples, except for paired views, as negatives. In the context of scRNA-seq, this assumption produces false negatives by forcing cells of the same type apart, thereby collapsing intra-cluster consistency and distorting the embedding space.

To address this, we propose scFNCCL, a dual contrastive learning framework that mitigates false negatives through similarity-based thresholding at the instance level, preserving intra-cluster consistency while enhancing separation across cell types. At the cluster level, where limited cluster numbers weaken contrastive learning, we reformulate the loss using a correlation matrix, promoting consistent assignments while enforcing orthogonality between clusters. We evaluate scFNCCL on four benchmark scRNA-seq datasets using both clustering metrics (ACC, NMI, ARI) and manual cell type annotations, and compare it against four representative methods: Seurat, scDCCA, scCCL, and JojoSCL. Our method consistently outperforms existing approaches, demonstrating improved robustness and biologically meaningful embeddings for single-cell data analysis.