Cycle-Verified Gentle Teaching and Confusion Correction for Label-Scarce Cross-Modal Single-Cell Annotation

Zheng Lu, Peng Yu, and Yan Wang(✉)

Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China
71275904035@stu.ecnu.edu.cn, 51275904040@stu.ecnu.edu.cn, ywang@cee.ecnu.edu.cn

Abstract. Accurate cell-type annotation across modalities is critical for single-cell multi-omics studies. However, integrating assays like scATAC-seq with well-annotated scRNA-seq references is hindered by substantial modality shifts and severe label scarcity. Under extreme low-label regimes (e.g., ≈1% source labels), existing semi-supervised transfer methods suffer from severe error accumulation and boundary collapse among closely related cell states due to unconstrained noisy pseudo-labeling. To address this, we propose CADENCE (Cycle-verified Assistant-driven DANCE with confusing pair corrEctioN), a unified semi-supervised framework designed to safely transfer knowledge under extreme label scarcity. For source data, CADENCE expands supervision via optimal transport with a curriculum-guided, class-balanced selection. For unlabeled target data, we replace direct pseudo-labeling with a robust teacher-assistant-student architecture, where an assistant branch absorbs noisy target supervision and distills only clean representation knowledge to the primary student model. To further guarantee reliability, a lightweight cycle-consistent latent translator acts as a structural gate to dynamically filter and reweight ambiguous target cells. Finally, a prototype-based confusing-pair correction mechanism explicitly rectifies systematic pairwise misclassifications. Extensive experiments on public benchmarks demonstrate that CADENCE significantly mitigates boundary collapse and improves annotation robustness among similar cell types, yielding highly reliable cross-modality cell atlases without requiring paired measurements.

Keywords: Single-cell multi-omics · Cross-omics label transfer · Confusion-aware learning · Cycle consistency.

1 Introduction

Single-cell multi-omics technologies, particularly scRNA-seq and scATAC-seq [2, 10], are essential for characterizing gene regulation across complementary molecular layers. While integrating these modalities enables the construction of comprehensive cell atlases, annotating scATAC-seq data remains challenging

due to its high sparsity, high dimensionality, and substantial modality shift from well-curated scRNA-seq references. Consequently, scATAC-seq datasets are often weakly labeled or entirely unlabeled, necessitating cross-modality label transfer pipelines rather than direct supervised learning.

To bridge the modality gap, various computational methods have been developed. Typical approaches tackle cross-modality integration by learning shared embeddings or aligning heterogeneous representations, including scalable kNN-based transfer (e.g., scJoint [12, 7]), reliability-aware alignment (e.g., scBridge [5]), and methods incorporating biological priors or contrastive objectives (e.g., scNCL [14]). Despite their success, most of these methods implicitly assume the availability of abundant, high-quality source labels or rely on expensive paired measurements. To relax these stringent requirements, recent semi-supervised frameworks like DANCE employ optimal transport (OT) [1, 16] for pseudo-labeling and a divide-and-conquer strategy for target learning. However, under extreme label scarcity—a realistic regime where as few as $\approx$1% of source cells are annotated—even these advanced methods rapidly degrade. Early pseudo-labeling errors are easily reinforced, and training directly on ambiguous target samples leaks noise into the prediction head, leading to severe boundary collapse between closely related cell states.

To solve this harsh low-label problem, we propose CADENCE, a robust semi-supervised transfer framework designed to mitigate noise propagation and refine decision boundaries. Building upon OT-based expansion and cross-omic mixup, CADENCE fundamentally restructures the target learning process through a gentle teacher-assistant-student architecture [4]. Instead of exposing the primary classifier to unreliable target pseudo-labels, an assistant network absorbs this noisy supervision, transferring only clean representation knowledge to the student via exponential moving average (EMA) updates. To further guarantee target supervision quality, we introduce a lightweight bidirectional latent translator trained with cycle consistency [6]. By evaluating the cycle-reconstruction error, CADENCE dynamically gates and reweights ambiguous target cells, providing a reliable internal check for cross-modal translation. Furthermore, to address persistent misclassifications among neighboring cell subtypes, we incorporate a prototype-based confusing-pair correction mechanism that dynamically identifies and rectifies soft labels in the feature space [17].

By systematically controlling pseudo-label reliability rather than merely expanding their quantity, CADENCE prevents noisy supervision from collapsing decision boundaries. In summary, our main contributions are three-fold:

1. **Problem Focus:** We identify and address the severe vulnerability of semi-supervised cross-modality transfer under extreme label scarcity ($\approx$1% source labels), where unconstrained noisy target supervision inevitably leads to boundary collapse.
2. **Methodological Innovation:** We propose CADENCE, a novel framework that prevents noise propagation via a gentle teacher-assistant-student representation distillation, structurally gated by a cycle-consistent reliability metric for ambiguous targets.

3. **Target Refinement & Impact:** We introduce a prototype-based confusing-pair correction mechanism that dynamically rectifies systematic pairwise misclassifications, achieving significantly more robust cell-type annotations without paired multi-omic data.

2 Method

2.1 Problem Definition and Overview

We tackle semi-supervised cell-type transfer from a source scRNA-seq dataset to a target scATAC-seq dataset under extreme label scarcity. Let the source domain be

$$\mathcal{D}^s = \mathcal{D}_l^s \cup \mathcal{D}_u^s, \quad \mathcal{D}_l^s = \{(x_i^s, y_i^s)\}_{i=1}^{N_l}, \quad \mathcal{D}_u^s = \{x_i^s\}_{i=1}^{N_u}, \tag{1}$$

where the labeled fraction is extremely small (e.g., $N_l/(N_l + N_u) \approx 1\%$). The target dataset $\mathcal{D}^t = \{x_i^t\}_{i=1}^{N_t}$ is entirely unlabeled but shares the same C cell types. We employ a shared encoder $F(\cdot)$ and classifier $H(\cdot)$ to extract latent features $z = F(x) \in \mathbb{R}^d$ and predictive probabilities $p = \mathrm{softmax}(H(z)) \in \Delta^C$.

The goal is to maintain discriminability despite the sparse source supervision and transfer robustly across the modality gap. To achieve this, CADENCE partitions target cells into source-like (high-confidence) and ambiguous (low-confidence) subsets. The model is trained via a two-phase process: a supervised warm-up on $\mathcal{D}_l^s$, followed by end-to-end semi-supervised learning comprising optimal transport (OT) source expansion, divide-and-conquer target training, and cross-modality alignment.

2.2 Curriculum OT Expansion on Source

To overcome the selection bias caused by the 1% label regime, we estimate pseudo-labels for $\mathcal{D}_u^s$ from a global distributional perspective. For a mini-batch of B unlabeled source samples with predictions $P \in \mathbb{R}^{B\times C}$, we solve an entropic-regularized OT problem to find a coupling matrix $Q^\star$:

$$Q^\star = \arg\min_{Q\in\mathcal{Q}(\frac{1}{B}\mathbf{1},\,\frac{1}{C}\mathbf{1})} \sum_{i,j} (-\log P_{ij})\, Q_{ij} + \frac{1}{\sigma}\sum_{i,j} -Q_{ij}\log Q_{ij}, \tag{2}$$

where $\mathcal{Q}$ is the transport polytope and σ controls the entropy regularization strength. Normalizing each row of $Q^\star$ yields OT-derived soft pseudo-labels $\hat{y}_i \in \Delta^C$.

To prevent early error amplification, CADENCE employs a curriculum schedule that gradually expands the pseudo-label budget $K_e = \lfloor \rho(e)\cdot N_u \rfloor$ over epochs e, where $\rho(e) \in [0, 1]$ is a monotonically increasing budget ratio. We select a class-balanced subset $\mathcal{I}_e \subset \mathcal{D}_u^s$ based on confidence. The selected samples are merged with the labeled set and trained with a discounted weight w_{pl}:

$$\mathcal{L}_s = \frac{1}{|\mathcal{D}_l^s|}\sum_{(x,y)\in\mathcal{D}_l^s} \mathrm{CE}(p(x), y) + w_{pl}\frac{1}{|\mathcal{I}_e|}\sum_{x\in\mathcal{I}_e} \mathrm{CE}\big(p(x), \arg\max \hat{y}(x)\big). \tag{3}$$

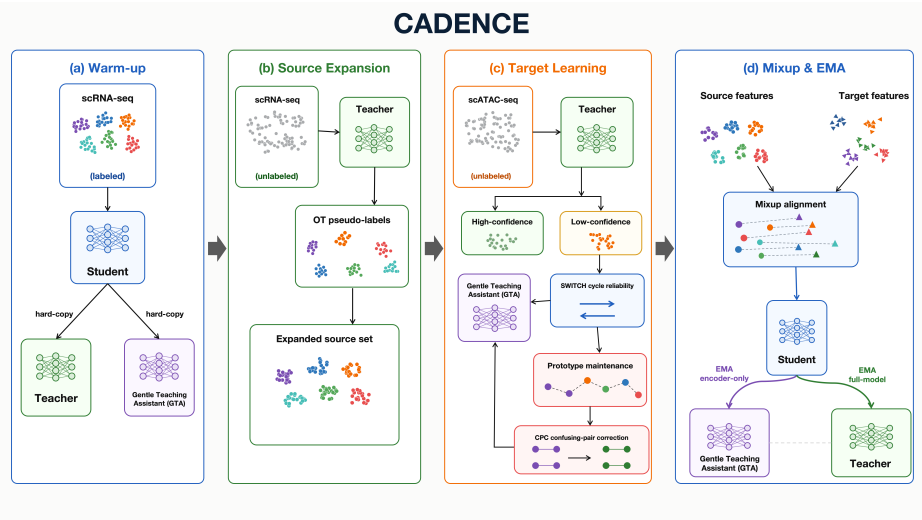

Fig. 1. Overview of CADENCE. (a) Warm-up initializes the Student, Teacher, and Gentle Teaching Assistant (GTA). (b) Source expansion generates OT-based pseudo-labels for unlabeled scRNA-seq. (c) Target learning splits scATAC-seq cells into high-/low-confidence subsets; SWITCH estimates cycle reliability and CPC performs prototype-based confusing-pair correction. (d) Mixup alignment reduces cross-omic shift, with GTA→Student encoder EMA and Student→Teacher full-model EMA.

Here, $\arg\max \hat{y}(x)$ converts the OT-derived soft assignment into a hard pseudo-label for source CE training, rather than using the model prediction $\arg\max p(x)$ directly.

2.3 Reliability-Aware Divide-and-Conquer on Target

Target pseudo-labels are inherently noisier due to modality shift. We partition target samples using the teacher model's confidence $s_i = \max_c p_i^T[c]$:

$$\mathcal{D}_{SL}^t = \{x_i^t \mid s_i > \tau\}, \quad \mathcal{D}_{AMB}^t = \{x_i^t \mid s_i \le \tau\}, \tag{4}$$

where τ is a confidence threshold. For source-like targets ($\mathcal{D}_{SL}^t$), we apply consistency regularization against perturbations $\tilde{x}$:

$$\mathcal{L}_{SL} = \frac{1}{|\mathcal{D}_{SL}^t|} \sum_{x \in \mathcal{D}_{SL}^t} \mathrm{CE}\big(p^S(\tilde{x}), \arg\max p^T(x)\big)\,. \tag{5}$$

Ambiguous targets ($\mathcal{D}_{AMB}^t$), conversely, require a dedicated reliability framework to prevent noise from corrupting the classifier.

Gentle Teacher–Assistant–Student Distillation. We isolate noisy supervision by introducing an assistant branch A alongside the teacher T and student S. The assistant learns from ambiguous targets using weak partial labels. To

shield the student's prediction head from corrupted gradients, the student only absorbs representation knowledge via an exponential moving average (EMA) of the assistant's encoder θ_f^A:

$$\theta_f^S \leftarrow \alpha\theta_f^S + (1-\alpha)\theta_f^A, \qquad \theta^T \leftarrow \beta\theta^T + (1-\beta)\theta^S. \tag{6}$$

Partial-Label Learning and Cycle Gating. For each $x_i^t \in \mathcal{D}_{AMB}^t$, we initialize a candidate mask m_i using the teacher's top-k predictions Y_i, yielding a normalized soft pseudo-label q_i. The assistant minimizes the partial-label objective:

$$\mathcal{L}_{AMB} = \frac{1}{|\mathcal{D}_{AMB}^t|} \sum_{x_i^t \in \mathcal{D}_{AMB}^t} w_i \sum_{c=1}^{C} -q_i[c] \log p_i^A[c]. \tag{7}$$

To determine the reliability weight w_i, we introduce a lightweight bidirectional latent translator ($G_{s\to t}$, $G_{t\to s}$) trained via cycle consistency:

$$\mathcal{L}_{\text{cyc}} = \mathbb{E}\Big[\big\|G_{t\to s}\big(G_{s\to t}(z^s)\big) - z^s\big\|_2^2\Big] + \mathbb{E}\Big[\big\|G_{s\to t}\big(G_{t\to s}(z^t)\big) - z^t\big\|_2^2\Big]. \tag{8}$$

The cycle error $e_i = \|G_{s\to t}\big(G_{t\to s}(z_i^t)\big) - z_i^t\|_2^2$ provides the per-cell reliability score $w_i = \exp(-e_i/\sigma_c)$ and a binary gate $g_i = \mathbb{I}[e_i < \delta]$, where σ_c controls the reliability weighting scale and δ is the cycle-error threshold. We iteratively refine q_i using the assistant's predictions ($q_i \leftarrow \text{Normalize}((m_i \odot q_i) \odot p_i^A)$) and progressively drop unlikely candidates based on a running average $\bar{q}_i$, shrinking the candidate set safely based on g_i.

Prototype-Based Confusing-Pair Correction. To mitigate persistent pairwise errors among neighboring cell states, CADENCE dynamically maintains EMA class prototypes μ_c using reliable source features:

$$\mu_c \leftarrow \eta\mu_c + (1-\eta) \cdot \text{Mean}\{z_i \mid \hat{y}_i = c\}. \tag{9}$$

By tracking the teacher's top-2 predictions for ambiguous targets passing the gate g_i, we identify the most frequently confused class pair (a, b). For affected samples, we evaluate the distances $d_a = \|G_{t\to s}(z_i^t) - \mu_a\|_2^2$ and $d_b = \|G_{t\to s}(z_i^t) - \mu_b\|_2^2$. Letting $m = q_i[a] + q_i[b]$ be their combined probability mass, we reallocate it toward the closer prototype:

$$q_i[a] \leftarrow (1-\gamma)q_i[a] + \gamma m, \qquad q_i[b] \leftarrow m - q_i[a] \quad \text{if } d_a \le d_b, \tag{10}$$

with the symmetric operation applied if $d_b < d_a$, where γ controls the correction strength.

2.4 Cross-Modality Mixup and Overall Objective

To align the latent spaces without paired data, we form semantic embeddings u_i^t and u_j^s using the classifier weights W, compute their similarity S_{ij}, and construct a multi-sample Mixup feature $\tilde{z}_i^s = \sum_j \text{softmax}(S) z_j^s$. The alignment loss is:

$$\mathcal{L}_{\text{align}} = \frac{1}{B} \sum_{i=1}^{B} \big\|z_i^t - \tilde{z}_i^s\big\|_2^2. \tag{11}$$

The entire framework is trained end-to-end. The student network minimizes the primary objective (with an optional entropy regularizer $\mathcal{R}_{\text{ent}}$):

$$\mathcal{L}_{\text{student}} = \mathcal{L}_s + \lambda_{SL}\mathcal{L}_{SL} + \lambda_{\text{align}}\mathcal{L}_{\text{align}} + \lambda_{\text{ent}}\mathcal{R}_{\text{ent}}, \tag{12}$$

while the assistant optimizes $\lambda_{\text{AMB}}\mathcal{L}_{\text{AMB}}$ and the translators optimize $\mathcal{L}_{\text{cyc}}$.

Unless otherwise specified, the default loss weights are set to $\lambda_{SL} = 0.1$, $\lambda_{\text{align}} = 0.1$, and $\lambda_{\text{ent}} = 1.0$.

3 Experiments

3.1 Datasets Processing and Experiment Setup

To validate the effectiveness and robustness of CADENCE, we conduct extensive experiments on two widely used benchmark single-cell multi-omics datasets. In our setup, scRNA-seq serves as the source domain and scATAC-seq as the unlabeled target domain.

Mouse Atlas Data. Derived from the Tabula Muris consortium, this dataset provides comprehensive multi-omics profiles of mouse tissues. Following standard evaluation protocols, the dataset is partitioned into multiple sub-tasks based on various sequencing technologies. We utilize the scRNA-seq gene expression matrices and the corresponding scATAC-seq quantitative gene activity score matrices for cross-modality transfer [15].

CITE-ASAP PBMC Data. This PBMC dataset contains cells profiled under both control and stimulated conditions. CITE-seq captures antibody-derived tag (ADT) and gene expression matrices, whereas ASAP-seq simultaneously captures ADT and chromatin accessibility matrices. For our experiments, we extract the gene expression matrices from CITE-seq (source) and the chromatin accessibility matrices from ASAP-seq (target) to construct the cross-modality integration tasks [10].

Evaluation Protocol under Label Scarcity. To rigorously simulate the extreme label scarcity scenario, we randomly sample only $\approx 1\%$ of the cells in the source scRNA-seq dataset to serve as the labeled set $\mathcal{D}_l^s$, leaving the remaining 99% of the source data and the entirety of the target scATAC-seq data unlabeled.

3.2 Baselines

We evaluate CADENCE against a comprehensive suite of baselines under the same $\approx 1\%$ source label regime. General domain adaptation methods include **DAN** [8] (minimizing MMD), **CDAN** [9] (conditional adversarial networks), and **MCC** [3] (minimizing pairwise class confusion), which address domain shifts but are not explicitly tailored for single-cell multi-omics sparsity. General semi-supervised learning is represented by **FixMatch** [13], demonstrating the vulnerability of naive thresholding under severe cross-modality shifts. Targeted single-cell integration methods include **scJoint** [7] (joint embeddings and kNN), **scNCL** [14] (neighborhood-aware contrastive learning), and **scBridge**

Table 1. Accuracy (%) under the 1% source-label setting. Best and second-best results are shown in **bold** and underlined. T1–T4 denote scRNA_SMARTer-snmC, scRNA_SMARTer-ATAC, snRNA_SMARTer-ATAC, and CITE-ASAP.

Method	Task 1	Task 2	Task 3	Task 4
DAN [8]	41.96	21.23	22.02	29.68
CDAN [9]	34.66	22.05	20.99	27.32
MCC [3]	32.09	21.73	20.99	22.53
FixMatch [13]	31.55	22.07	21.25	25.71
scJoint [7]	23.94	19.62	11.05	23.42
scNCL [14]	45.28	19.93	19.89	22.95
scBridge [5]	13.13	28.32	26.44	26.76
DANCE [16]	51.56	46.40	47.43	46.83
Ours	**55.80**	**49.82**	**50.22**	**55.82**

Table 2. Ablation study on CITE-ASAP.

Seg-GTA	SWITCH	CPC	Acc (%)
			46.83
✓			48.82
	✓		48.95
		✓	51.62
✓	✓	✓	**55.82**

[5] (reliability-aware alignment), which often implicitly assume abundant source supervision. Finally, **DANCE** [16], our direct predecessor, expands source labels via optimal transport and employs a confidence-based divide-and-conquer strategy. It serves as our primary baseline to explicitly demonstrate that our reliability-enhancing modules (SWITCH, CPC, and Seg-GTA) are essential to prevent boundary collapse under extreme label scarcity.

3.3 Results

Main Results on Benchmark Datasets. Table 1 presents the quantitative classification accuracy across four cross-modality transfer tasks under the extreme $\approx 1\%$ source label regime. Overall, our proposed framework achieves the highest accuracy across all tasks, demonstrating superior robustness against severe modality shifts and extreme label scarcity.

General domain adaptation methods (e.g., DAN [8], MCC [3]) and standard semi-supervised frameworks (e.g., FixMatch [13]) exhibit significant performance degradation, with accuracies dropping to the 20%–30% range on Tasks 2, 3, and 4. This indicates that standard global distribution alignment and naive confidence thresholding are insufficient to overcome the high background noise

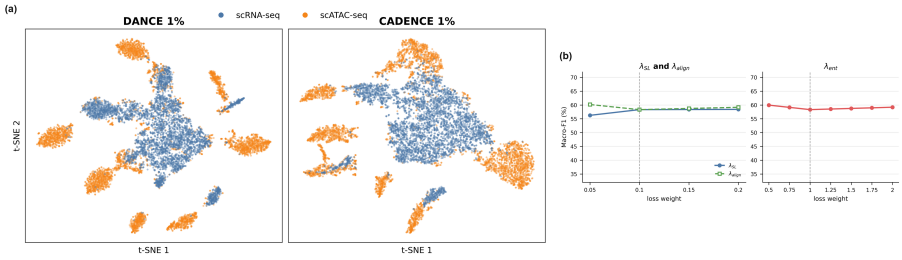

Fig. 2. Additional analysis on Task 4 under the 1% source-label setting. (a) t-SNE visualization of source scRNA-seq and target scATAC-seq embeddings. (b) Hyperparameter sensitivity of CADENCE.

and modality discrepancies inherent in single-cell multi-omics when supervision is severely limited.

While traditional single-cell integration methods (e.g., scNCL [14], scJoint [7]) typically perform well under standard settings, they suffer from severe boundary collapse in our extreme low-label scenario. For instance, although scNCL [14] achieves a moderate 45.28% on Task 1, its performance drastically falls to 19.89% on Task 3. Similarly, scJoint [7] collapses to 11.05% on Task 3, and scBridge [5] dramatically fails on Task 1 (13.13%). These failures highlight that existing single-cell methods heavily rely on abundant source labels or high-quality cross-modality pseudo-pairs; without them, the models are easily misled by erroneous pseudo-labels, leading to local class collapse.

Compared to the recent semi-supervised baseline, DANCE [16], which establishes a relatively stable performance, our framework achieves consistent and substantial performance gains. Notably, on the highly challenging Task 3, our method is the only one to exceed the 50% accuracy threshold (50.22%). It also yields a substantial absolute improvement of exactly 8.99% on Task 4 (increasing from 46.83% to 55.82%). On Task 4, CADENCE achieves a Macro-F1 of 58.27%, supporting the accuracy-based comparison under class imbalance.

Ablation Study. We conduct an ablation study on Task 4 (CITE-ASAP) to validate each proposed component (Table 2). Adding Seg-GTA [4] or SWITCH [6] to the base architecture (DANCE [16]) yields improvements of +1.99% (48.82%) and +2.12% (48.95%), respectively. This confirms that isolating the classifier from noisy gradients and filtering distorted samples effectively enhance pseudo-label reliability. Notably, CPC [17] improves accuracy by +4.79 percentage points to 51.62%, highlighting the importance of prototype based confusing-pair correction. Finally, the full CADENCE model synergizes all components to achieve a peak performance of 55.82%. The modules are highly complementary: Seg-GTA and SWITCH filter generalized noise, while CPC refines challenging local boundaries, forming a robust solution against extreme label scarcity.

Visualization and Sensitivity Analysis. Fig. 2 provides qualitative and sensitivity analyses on Task 4. Compared with DANCE, CADENCE produces a more connected cross-modal embedding, with target scATAC-seq cells better

aligned to the source scRNA-seq manifold. The sensitivity curves show that CADENCE remains stable around the default loss weights.

4 Conclusion

In this paper, we introduced CADENCE, a robust semi-supervised framework designed to tackle the severe challenges of cross-modality cell-type annotation under extreme label scarcity (≈1% source labels). By shifting the focus from merely expanding pseudo-labels to strictly controlling their reliability, our method effectively mitigates cross-omic error accumulation. Specifically, the synergistic integration of a gentle teacher-assistant-student distillation scheme, cycle-consistent target gating, and prototype-based confusing-pair correction successfully prevents the classification boundary collapse that plagues existing methods. Extensive experiments across diverse multi-omics benchmarks demonstrate that CADENCE significantly outperforms previous state-of-the-art baselines. Ultimately, this work provides a highly reliable and practical computational tool for constructing comprehensive single-cell atlases in realistic, resource-constrained biomedical scenarios.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Grant No. 62471182), Shanghai Rising-Star Program (Grant No. 24QA2702100), Science and Technology Commission of Shanghai Municipality Basic Research Program (Grant No. 25JD1401300), and the Science and Technology Commission of Shanghai Municipality (Grant No. 22DZ2229004).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chang, W., Shi, Y., Wang, J.: CSOT: Curriculum and structure-aware optimal transport for learning with noisy labels. In: Advances in Neural Information Processing Systems, vol. 36, pp. 8528–8541 (2023)
2. Chen, H., Lareau, C., Andreani, T., Vinyard, M.E., Garcia, S.P., Clement, K., Andrade-Navarro, M.A., Buenrostro, J.D., Pinello, L.: Assessment of computational methods for the analysis of single-cell ATAC-seq data. Genome Biol. **20**(1), 1–25 (2019)
3. Jin, Y., Wang, X., Long, M., Wang, J.: Minimum class confusion for versatile domain adaptation. In: Computer Vision–ECCV 2020, LNCS, vol. 12366, pp. 464–480. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58571-6_28
4. Jin, Y., Wang, J., Lin, D.: Semi-supervised semantic segmentation via gentle teaching assistant. In: Advances in Neural Information Processing Systems (2022)
5. Li, Y., Zhang, D., Yang, M., Peng, D., Yu, J., Liu, Y., Lv, J., Chen, L., Peng, X.: scBridge embraces cell heterogeneity in single-cell RNA-seq and ATAC-seq data integration. Nat. Commun. **14**(1), 6045 (2023)

6. Li, Z., Qu, S., Liang, H., Tang, R., Zhang, X., Lu, F., Yang, J., Gan, Z., Gao, S., Zhang, Y., Chen, G.: Integrative deep learning of spatial multi-omics with SWITCH. Nature Computational Science (2025). https://doi.org/10.1038/s43588-025-00891-w
7. Lin, Y., Wu, T.Y., Wan, S., Yang, J.Y., Wong, W.H., Wang, Y.R.: scJoint integrates atlas-scale single-cell RNA-seq and ATAC-seq data with transfer learning. Nat. Biotechnol. **40**(5), 703–710 (2022)
8. Long, M., Cao, Y., Wang, J., Jordan, M.: Learning transferable features with deep adaptation networks. In: International Conference on Machine Learning, pp. 97–105. PMLR (2015)
9. Long, M., Cao, Z., Wang, J., Jordan, M.I.: Conditional adversarial domain adaptation. Advances in Neural Information Processing Systems **31** (2018)
10. Mimitou, E.P., Lareau, C.A., Chen, K.Y., Zorzetto-Fernandes, A.L., Hao, Y., Takeshima, Y., Luo, W., Huang, T.-S., Yeung, B.Z., Papalexi, E., et al.: Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. Nature Biotechnology **39**(10), 1246–1258 (2021)
11. Rozenblatt-Rosen, O., Stubbington, M.J., Regev, A., Teichmann, S.A.: The human cell atlas: from vision to reality. Nature **550**(7677), 451–453 (2017)
12. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., Satija, R.: Comprehensive integration of single-cell data. Cell **177**(7), 1888–1902 (2019)
13. Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C.A., Cubuk, E.D., Kurakin, A., Li, C.-L.: Fixmatch: Simplifying semisupervised learning with consistency and confidence. In: Proceedings of the Conference on Neural Information Processing Systems (2020)
14. Yan, X., Zheng, R., Chen, J., Li, M.: scNCL: transferring labels from scRNA-seq to scATAC-seq data with neighborhood contrastive regularization. Bioinformatics **39**(8), btad505 (2023)
15. Yao, Z., Liu, H., Xie, F., Fischer, S., Adkins, R.S., Aldridge, A.I., Ament, S.A., Bartlett, A., Behrens, M.M., Van den Berge, K., et al.: A transcriptomic and epigenomic cell atlas of the mouse primary motor cortex. Nature **598**(7879), 103–110 (2021)
16. Zhang, F., Liu, T., Chen, Z., Peng, X., Chen, C., Hua, X.S., Luo, X., Zhao, H.: Semi-supervised Knowledge Transfer Across Multi-omic Single-cell Data. In: Proceedings of the 38th Conference on Neural Information Processing Systems (NeurIPS 2024) (2024)
17. Zhi, C., Zhuo, J., Wang, S.: Confusing pair correction based on category prototype for domain adaptation under noisy environments. In: The Thirty-Eighth AAAI Conference on Artificial Intelligence, pp. 17060–17068. AAAI Press (2024)