SOFT-LABEL TENSOR LEARNING FOR CONTINUAL MULTI-VIEW CLUSTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-view clustering (MVC) partitions data with multiple views in an unsupervised manner. Commonly, it assumes that all data views are available in advance; however, in many real-world scenarios, new data views are accumulated over time, bringing the topic of continual multi-view clustering (CMVC). In this paper, we suggest two key constraints that the task of CMVC should follow: cross-session information utilization, given the limitations on historical data, newly arrived view information and retained historical clustering information should be effectively utilized for cross-session knowledge transfer; structural stability, in an incremental scenario, the clustering structure must be updated in a stable manner. Specifically, to achieve these constraints, we introduce the probability membership module to learn the relationship between current soft labels and historical soft labels, thereby identifying useful information introduced by the current view while leveraging historical clustering information. We further introduce the tensor correction learning module, which jointly models the current soft labels, historical soft labels, and the learned consistency structure as a third-order tensor. It employs the Low-Bias Refined norm to suppress redundant structural components while retaining the dominant ones, thereby enhancing the stability of the clustering structure. Upon these above ideas, we propose a novel CMVC method, namely Soft-Label Tensor Learning for Continual Multi-View Clustering (SLT-CMVC). We conduct extensive experiments to evaluate SLT-CMVC on multiple benchmark datasets. Empirical results demonstrate that SLT-CMVC can significantly outperform existing baseline methods.

1 INTRODUCTION

With the development of multi source sensing and heterogeneous data acquisition technologies, a single research subject can now generate feature descriptions that are both complementary and distinct across multiple dimensions, thereby forming multi-view data Yin et al. (2021). Multi-view clustering (MVC) aims to overcome the limitations of single view approaches by integrating the shared semantics and unique attributes across different views to accurately capture data structures Wang et al. (2025a). In recent years, researchers have proposed numerous advanced MVC algorithms. By efficiently integrating complementary information from multiple viewpoints, these algorithms have demonstrated excellent clustering performance across a wide range of practical applications Wang et al. (2026a), Ouyang et al. (2026).

However, in practice, there are application scenarios for streaming view clustering. Take the task of object recognition as an example: we need to distinguish between targets and the background. Due to limited band information and insufficient feature representation in early stages, clustering results often exhibit blurred boundaries, which affects target extraction and fine-grained classification. With the advancement of imaging technology, multi-band features have become increasingly rich, and data complexity has risen. One intuitive approach is to fuse historical and newly acquired data for unified re-clustering; however, this method incurs significant computational and storage overhead and struggles to adapt promptly to dynamic changes in the feature spaceYan et al. (2025), Wan et al. (2024), Li et al. (2025b), Wang et al. (2024), Wang et al. (2026b).

To address the above tasks, researchers have proposed various algorithms, such as, Feng et al. (2025a) proposed a CDIMVC algorithm that incorporates a knowledge transfer module to control updates to

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

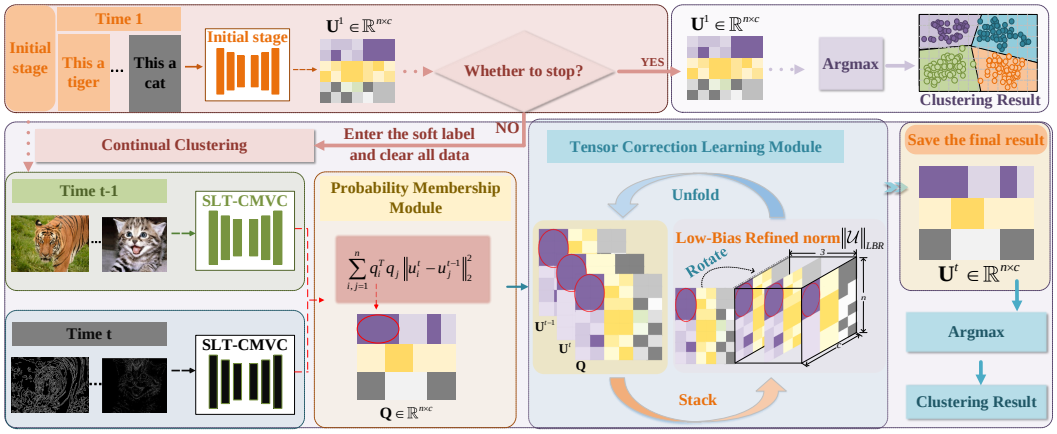


Figure 1: The soft labels for the first view are initialized using the C-means algorithm with an entropy regularization term. Next, a probability objective function is employed to fuse the soft labels from the current stage with those from the previous stage, thereby extracting cross-stage consistency information to effectively suppress noise and stabilize the clustering structure during the evolution process. Subsequently, Stack the clustering results into a tensor and introduce Low-Bias Refined norm as the core tensor constraint to capture higher-order correlations across views. Finally, optimal clustering results are output based on this robust low rank approximation operator.

historical consistency information by new views, ensuring smooth knowledge propagation during incremental processing and maintaining the algorithm's sustained stable clustering capability. and Wang et al. (2025b) proposed a ACMVC employed a teacher-student model architecture to extract feature information from views, using the similarity matrix of the preceding view as the alignment target to achieve clustering.

It is evident that existing methods rely on directly merging the current partition with the previous consensus structure. However, this merging treats the historical state as a fixed clustering reference and fails to explicitly capture the consistency of the structural relationship between the current clustering assignment and the historical one. This limitation becomes even more apparent when newly added views contain noise or incomplete discriminative information, as directly introducing a new clustering structure may propagate errors and lead to structural drift. Therefore, the challenge lies in learning a reliable, cross-session consistent representation from limited clustering information across consecutive sessions.

In this paper, we propose the Continual Multi View Clustering by Soft Label Tensor Learning (SLT-CMVC) to address this problem. For the CMVC task, we propose two constraints:

C1 (Cross-session information utilization): since historical raw features and view-specific representations are inaccessible in subsequent sessions, the current session can only exploit limited historical clustering information, requiring effective utilization of the relationship between the newly arrived view and the retained historical clustering state for cross-session knowledge transfer.

C2 (Structural Stability in Dynamic Environments): From the continual nature of the task, we observe that the clustering structure evolves across sessions and should preserve useful historical structure while adapting to newly arriving information, thereby avoiding excessive structural drift.

Given the above constraints, during the initialization phase, we employ the C-means algorithm with an entropy regularization term to obtain robust initial soft labels, thereby avoiding trivial solutions and establishing a reliable baseline. Subsequently, during the iterative clustering process, we fuse the current soft labels with those from the previous phase via a probabilistic objective function to extract cross-phase consistency information, thereby suppressing noise and stabilizing the clustering structure. To address the bottlenecks in soft label information extraction, we introduce a tensor learning framework that stacks label matrices into a three-dimensional tensor, and propose a novel tensor constraint: Low-Bias Refined norm. Compared to the Tensor Nuclear Norm (TNN), this constraint achieves a superior balance between unbiasedness and sparsity through a log-exponential composition mechanism; while suppressing cross-view redundant noise, it significantly reduces the penalty bias on core topological features, providing a highly robust low rank approximation operator

for multi source data fusion in continual clustering scenarios. Figure 1 illustrates the overall workflow of the proposed method.

The main contributions of this paper can be summarized as follows:

- We formulate a cluster-level memory setting for continual multi-view clustering, where historical raw features and view-specific representations are discarded after each session while the soft clustering structure is retained for subsequent updates. This setting emphasizes continual clustering under restricted historical information rather than relying on repeated access to historical feature representations.
- This paper proposes a probability Membership Module for adaptive cross session consistency learning. By jointly modeling current and historical soft labels, this mechanism is able to learn consistency and inconsistency relationships among clustering structures and provides an information rich and compact state for continuous knowledge transfer.
- We introduce tensor learning to the field of continual clustering and propose a novel tensor constraint. Compared to the traditional tensor nuclear norm (TNN), this constraint achieves a superior balance between unbiasedness and sparsity, effectively suppressing redundant noise while significantly reducing the penalty bias for singular values.

2 Related Work

2.1 Multi-view Clustering

By effectively integrating complementary information from multiple perspectives, the MVC algorithm outperforms single-view methods in terms of performance. Its primary approaches encompass four categories: graph-based Wu et al. (2024), subspace-based Guo et al. (2025), kernel-based Liu et al. (2023), and deep learning-based Chen et al. (2025) methods. Subspace methods map multi-view data to a unified latent subspace and achieve classification through regularization constraints. Graph-based methods extract feature structures from each view and fuse them into a consensus graph for clustering. Kernel-based methods map nonlinear data to high-dimensional spaces using different kernel functions, facilitating data separation and yielding clear clusters. Deep learning-based methods utilize view-specific encoders to extract diverse representations and construct consensus representations through self-supervised tasks, which are then input into the clustering module to generate the final results.

2.2 Incremental Multi-View Clustering

Incremental Multi-View Clustering, as an emerging and significant research direction, has yielded several representative achievements Wan et al. (2025). For example, Qu et al. (2024) proposed the LAIMVC algorithm, which for the first time extended the anchor-based MVC framework into an incremental architecture. This algorithm employs historical view anchors for guidance, thereby preventing misalignment issues across views. It maintains consistent anchor matrices and graph structures across all historical views, iteratively integrating them until the new view ceases to change. Additionally, Feng et al. (2025b) proposed an approximate kernel matrix alignment and fusion algorithm, pioneering the integration of non-negative matrix factorization (NMF) with incremental learning. This approach resolves the challenge of incrementally updating large-scale base kernels, enabling efficient alignment of approximate kernel matrices. These methods demonstrate that compact historical states, such as structural information, anchors, and kernel-related representations, can be maintained to support continual or incremental clustering. Although existing methods avoid repeatedly processing historical views, the retained historical states determine the information that subsequent sessions can utilize. For example, while retaining graph structures, kernel matrices, or view-specific representations can provide relatively rich historical information, they also entail additional storage and maintenance overhead. More importantly, most existing methods update information between new and old sessions through historical information transfer, structural alignment, or incremental fusion, but they lack the ability to model and continuously update cross-session consistency information between current and historical clustering structures. To address this, we have proposed SLT-CMVC to fill the current gap.

3 THE PROPOSED METHOD

3.1 PROBLEM STATEMENT

Problem Definition: For CMVC, we define a sequence $\left\{\mathrm{X}^1, \mathrm{X}^2, ..., \mathrm{X}^t\right\}_{t=1}^{T}$ comprising t clustering sessions, where $\mathrm{X}^t \in \mathbb{R}^{d_t \times n}$ represents the input data for the session t, T denotes the final session Here, n denotes the number of samples and d_t denotes the feature dimension. After data is transmitted to the server, the system executes the algorithm and returns the results; subsequently, the server clears the original data, retaining only the current clustering results. According to the CMVC definition above, the feature space evolves continually throughout the session, and X^t is directly accessible only within session t. Therefore, when processing newly added multi-view data, the algorithm must continually update the clustering structure based on soft labels generated from the previous session to enhance clustering performance. The ultimate goal is to achieve clustering performance superior to any single-view method, or at least maintain the stability of the original clustering structure, when new view information is limited.

3.2 FORMULATION

Given multi-view data X^t, we learn its soft label representation through fuzzy clustering. To mitigate clustering ambiguity caused by excessively high membership degrees and enhance inter-class separability, we introduce entropy regularization as a penalty term in the objective function to obtain clearer and more discriminative clustering results. Its formulation is as follows:

$$\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik}^t \left\| {x_i}^t - {v_k}^t \right\|_2^2 + \lambda u_{ik}^t \ln u_{ik}^t \quad s.t.\ 0 < u_{ik}^t < 1 \tag{1}$$

where, element u_{ik}^t represents the probability that the i-th sample belongs to the k-th cluster, vector v_k^t denotes the k-th cluster center, and λ is the hyperparameter.

A1 (Probability Membership Module): When the model receives the clustering results from the previous stage, in order to address the constraint on cross-session information utilization, we introduce the dot product operation on probability vectors to extract consistency relationships across views, thereby uncovering effective clustering information. Formally, we multiply the dot product result by the corresponding distance metric to obtain the membership relationship between two soft labels. The Probability Membership Module (PMM) is defined as follows:

$$\begin{gathered}\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik}^t \left\| x_i^t - v_k^t \right\|_2^2 + \lambda u_{ik}^t \ln u_{ik}^t + \lambda_1 \left(\sum_{i=1}^{n}\sum_{j=1}^{n} q_i^T q_j \left\| u_i^t - u_j^{t-1} \right\|_2^2 + q_{ik} \ln q_{ik} \right) \\ s.t.\ \left\{ u_{ik}^t, u_{ik}^{t-1}, q_{ik} \right\} \in (0,1)\,.\end{gathered} \tag{2}$$

here, $q_i = \{q_{i1}, q_{i2}, ..., q_{ik}\}$ denotes the probability vector, and $q_i^T q_j$ can be interpreted as the probability weight for soft label consistency between two sessions. By minimizing this objective function Eq. (2), $q_i^T q_j$ becomes negatively correlated with the distance metric $\left\| u_i^t - u_j^{t-1} \right\|_2^2$, thereby simultaneously learning the probabilities of consistency and in-consistency in soft labels across the two sessions.

Definition 1 (Low-Bias Refined norm): Given a third-order tensor $\mathcal{U} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the Low-Bias Refined norm (LBR) of $\mathcal{U}$ is defined as:

$$\|\mathcal{U}\|_{LBR} = \frac{1}{n_3} \sum_{k=1}^{n_3} \sum_{i=1}^{h} \ln \left(2 - e^{\frac{-\mathcal{P}_f^k(i,i)}{\varepsilon}} \right) \tag{3}$$

where $h = min(n1, n2)$ and ε is a positive scalar that controls the flexibility of the norm. Here, $\mathcal{P}_f^k$ denotes the k-th frontal slice of the tensor $\mathcal{P}_f$, which is obtained through the tensor singular value decomposition (t-SVD Kilmer et al. (2013)) of $\mathcal{P}$'s k-th frontal slice $\mathcal{U} = \mathcal{S} * \mathcal{P} * \mathcal{V}^T$. A full exposition of LBR is presented in Appendix 8.1

A2 (Tensor Correction Learning Module): To ensure that the current clustering results can correctly guide the clustering of subsequent view information during the continuous clustering process, we

have introduced tensor learning into our continuous clustering framework. Specifically, we stack the probability matrices corresponding to the previous session, the current session, and the consistency information between the two sessions into a three-dimensional tensor. This approach fully extracts the implicit information within views, effectively learning the complementarity and consistency of multi-view data. The Tensor Correction Learning Module (TCLM) is defined as follows:

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik}^{t}\left\|x_i^t - v_k^t\right\|_2^2 + \lambda u_{ik}^t \ln u_{ik}^t \\ &+\lambda_1\left(\sum_{i=1}^{n}\sum_{j=1}^{n} q_i^T q_j \left\|u_i^t - u_j^{t-1}\right\|_2^2 + q_{ik}\ln q_{ik}\right) + \|\mathcal{U}\|_{LBR} \\ &s.t.\,\mathcal{U} = \left\{\mathbf{U}^t, \mathbf{U}^{t-1}, \mathbf{Q}\right\}, \left\{u_{ik}^t, u_{ik}^{t-1}, q_{ik}\right\} \in (0,1). \end{aligned} \tag{4}$$

By performing tensor learning on $\mathcal{U}$, complementary information exchange can be facilitated while maintaining consistency in the clustering structure, thereby correcting biases within the clustering structure. Traditional $||\mathcal{U}||_*$ constraint imposes an unbounded linear penalty on large singular values, leading to numerical bias in the optimization results. The $||\mathcal{U}||_{LBR}$ constraint we propose ensures that large singular values are not excessively compressed after algorithmic processing, while small singular values guide the tensor to converge rapidly to a low-rank state.

Eq. (4) represents the final objective function for SLT-CMVC. Compared to existing methods, SLT-CMVC leverages tensor learning to extract implicit information from soft labels and utilizes higher-order correlations to guide the correct clustering structure, thereby effectively enhancing clustering performance.

4 Optimization

To solve the proposed optimization problem, we adopt the Alternating Direction Method of Multipliers (ADMM) Lin et al. (2011). The corresponding augmented Lagrangian function is formulated as follows:

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik}^{t}\left\|x_i^t - v_j^t\right\|_2^2 + \lambda u_{ik}^t \ln u_{ik}^t + \|\mathcal{M}\|_{LBR} \\ &+\lambda_1\left(\sum_{i=1}^{n}\sum_{j=1}^{n} q_i^T q_j \left\|u_i^t - u_j^{t-1}\right\|_2^2 + q_{ik}\ln q_{ik}\right) + \frac{\mu}{2}\left\|\mathcal{M} - \mathcal{U} + \mathcal{A}/\mu\right\|_F^2 \\ &s.t.\mathcal{U} = \mathcal{M}, \mathcal{U} = \left\{\mathbf{U}^t, \mathbf{U}^{t-1}, \mathbf{Q}\right\}, \left\{u_{ik}^t, u_{ik}^{t-1}, q_{ik}\right\} \in (0,1). \end{aligned} \tag{5}$$

In the above formulation, $\mathcal{A}$ is the Lagrange multipliers and $\mu > 0$ is the penalty parameters. The detailed optimization procedure is provided in Appendix 8.2 and 8.3.

4.1 Complexity Analysis

Memory complexity: The primary memory overhead of SLT-CMVC stems from the matrices $\mathbf{U}^{t-1} \in \mathbb{R}^{c\times n}$, $\mathbf{U}^t \in \mathbb{R}^{c\times n}$, $\mathbf{V}^t \in \mathbb{R}^{d\times c}$ and $\mathbf{Q} \in \mathbb{R}^{c\times n}$. Thus the space complexity of our SLT-CMVC is $\mathcal{O}(cn + dc)$, which proves its linear space complexity.

Time complexity: As previously stated, the computational complexity of SLT-CMVC consists of five steps. The time consumed when updating $\mathbf{U}^t$, $\mathbf{V}^t$, $\mathcal{M}$, and $\mathcal{A}$ is $\mathcal{O}(cnd + dn + dc)$. And the variable $\mathbf{Q}$ involves the Euclidean distance, so it will cost $\mathcal{O}(cn^2)$ and the overall time complexity is $\mathcal{O}(cn^2 + cnd + dn + dc)$.

5 Experiment

The performance of the SLT-CMVC algorithm is validated on 7 common multi-view datasets and compared with 10 state-of-the-art clustering methods. The detailed summary of them is provided in Appendix 8.4

5.1 EXPERIMENT SETTINGS

This paper selected 5 state-of-the-art single-view clustering and incremental multi-view clustering algorithms, all of which were parameterized according to the original literature recommendations to ensure optimal performance. For the single-view algorithms, the final result was determined by the highest clustering accuracy achieved across all views.

Table 1: Clustering results (%) w.r.t three metrics on seven datasets. "–" indicates out of memory.

Dataset	Metric	Single view clustering					Incremental multi-view clustering					Ours
		DEC	MLKM	UFCER	UCFCM	CEHM	ACMVC	INMKC	CDIMVC	LAIMVC	ASIAMVC	
3Sources	ACC	64.29	50.41	62.72	59.18	35.53	58.79	55.03	53.25	47.34	66.27	**73.96**
	NMI	62.70	27.82	49.46	35.94	20.15	56.36	48.80	41.07	27.90	52.02	**64.85**
	Purity	70.61	52.13	62.72	56.19	41.12	62.54	68.91	60.95	56.80	72.78	**75.15**
MSRC	ACC	69.75	62.34	72.86	73.33	70.21	77.89	40.90	68.10	40.95	70.17	**84.76**
	NMI	52.68	57.71	68.73	55.11	50.98	75.16	31.06	50.18	24.20	60.39	**76.18**
	Purity	67.51	63.81	70.13	62.29	65.19	80.00	43.33	70.18	40.95	70.17	**85.76**
Synthetic3d	ACC	79.68	88.17	80.17	68.12	84.33	88.98	90.12	92.00	87.33	80.72	**94.67**
	NMI	65.92	66.98	66.13	33.29	61.69	67.78	70.59	79.91	63.36	66.36	**80.31**
	Purity	79.68	88.17	79.12	37.65	63.42	88.98	80.15	92.10	87.33	79.94	**94.67**
Wiki	ACC	43.73	54.01	50.91	29.18	53.21	55.42	41.17	44.28	57.15	56.02	**61.51**
	NMI	25.37	53.45	49.13	18.55	39.90	51.29	26.22	27.81	54.71	50.12	**56.26**
	Purity	45.18	61.74	59.18	26.89	55.21	59.83	45.01	48.19	60.46	58.67	**62.53**
Scene15	ACC	20.51	28.18	30.15	24.01	32.74	29.65	32.54	33.73	26.27	30.14	**34.14**
	NMI	13.99	27.52	29.52	20.32	19.33	18.92	33.25	31.59	18.37	32.01	**35.62**
	Purity	24.32	31.79	31.55	29.81	36.10	29.87	30.78	36.97	29.36	32.59	**38.93**
Fashion	ACC	48.38	51.74	58.15	62.31	58.21	–	**75.59**	72.31	60.00	59.31	74.51
	NMI	46.52	51.24	60.32	55.18	41.04	–	69.87	70.35	57.11	60.28	**73.25**
	Purity	52.11	60.18	62.81	62.31	62.12	–	**75.59**	72.31	64.08	64.22	75.13
Reuters	ACC	30.98	35.72	50.79	34.78	–	–	34.37	–	40.96	–	**55.96**
	NMI	18.97	29.93	34.47	26.32	–	–	15.65	–	29.98	–	**37.81**
	Purity	32.38	39.14	59.84	39.98	–	–	41.40	–	50.65	–	**60.74**

Datasets: To further validate the effectiveness of the proposed method, we conduct experiments on 7 real-world multi-view datasets, including **3Sources**, **MSRC**, **Wiki**, **Reuters**, **Scene15**, **Fashion** and **Synthetic3d**.

Compared methods: (1) **DEC** Xie et al. (2015) (2) **MLKM** Wang et al. (2025c); (3) **UFCER** Nie et al. (2025a); (4) **UC-FCM** Nie et al. (2025b); (5) **CEHM** Li et al. (2025a); (6) **ACMVC** Wang et al. (2025b); (7) **INMKC** Feng et al. (2025b); (8) **CDIMVC** Feng et al. (2025a); (9) **LAIMVC** Qu et al. (2024); (10) **ASIA-MVC** Qu et al. (2025).

Clustering performance is evaluated using three common metrics: accuracy (ACC), normalized mutual information (NMI), and Purity. To mitigate the impact of random initialization in k-means, all experiments were independently run 50 times with random initialization, and the average results for each metric are reported.

5.2 EXPERIMENT RESULTS

The clustering performance of 10 state-of-the-art compared approaches and ours on 7 benchmark datasets are listed in Table 1. Based on the observations in the table, the following conclusions can be drawn: Table 1 lists the clustering performance of 10 advanced comparison methods and SLT-CMVC on 7 benchmark datasets. The following conclusions can be drawn:

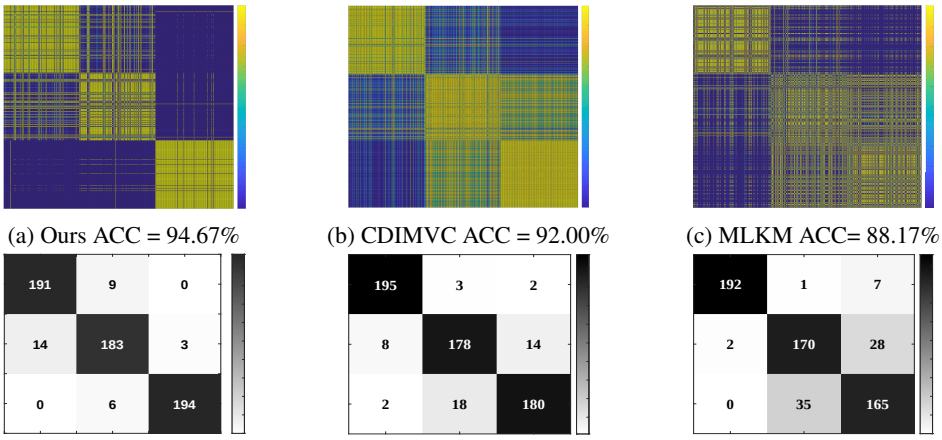
- The proposed SLT-CMVC significantly outperforms all comparison methods on the 3Sources, MSRC, Synthetic3d, and Reuters datasets, while also achieving optimal results on other datasets. Specifically, SLT-CMVC achieved ACC improvements of 7.69%, 4.11%, 3.73%, and 5.97% over the second-best method across the aforementioned datasets, with consistent enhancements observed across other evaluation metrics, fully demonstrating the superiority of this approach.

- Compared to existing incremental multi-view clustering methods, the proposed SLT-CMVC demonstrates significant advantages in clustering performance. Although baseline methods retain more view feature information, SLT-CMVC still achieves superior results on most datasets. Taking the Wiki dataset as an example, SLT-CMVC demonstrates significantly superior performance compared to these methods. This is achieved through the introduction of tensor technology in SLT-CMVC, which enables the extraction and integration of complementary information from multiple sources within limited data, thereby effectively preserving the correctness and stability of the clustering structure.
- Although single-view methods typically select the clustering results from the optimal view as the final output, they exhibit significant limitations: First, in practical scenarios, the quality of views is often difficult to assess beforehand, limiting the applicability of such methods. Second, single-view methods fail to leverage the inherent complementary and consistency information within multi-view data, thereby constraining further improvements in clustering performance. SLT-CMVC integrates multi-view consistency information through a probabilistic objective function and leverages tensor learning to extract complementary information from soft labels. This approach guides the correct clustering structure using consistency information, significantly enhancing overall clustering performance. Experimental results further validate this conclusion.

5.3 Qualified Study

To further validate the superiority of SLT-CMVC, we visualized similarity matrices and confusion matrices on the Synthetic3d dataset and compared them with the best-performing algorithms in single-view and incremental multi-view clustering. The results are shown in Figure 2. The visualization reveals that SLT-CMVC more effectively identifies marginal samples and extracts consistency information across views. In contrast, CDIMVC and MLKM exhibit insufficient marginal sample identification, and CDIMVC fails to fully exploit view consistency information, resulting in an unclear block-diagonal structure. By comparison, SLT-CMVC demonstrates a distinct block-diagonal structure.

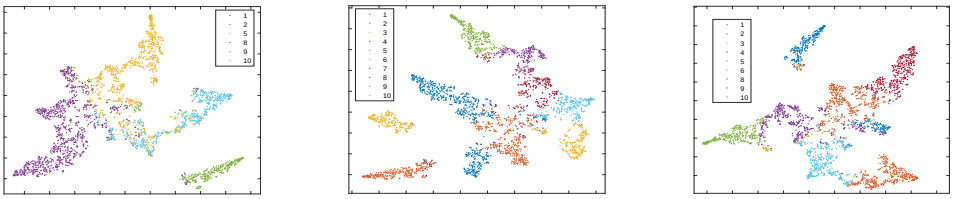
To further validate SLT-CMVC's advantages in continual clustering tasks, we performed t-SNE van der Maaten and Hinton (2008) visualization of its pseudo-labels on the Wiki and 3Sources datasets and conducted an intuitive comparison with other suboptimal algorithms, as shown in Figure 3. The figure demonstrates that SLT-CMVC fuses view consistency and complementary information through tensor learning while uncovering higher-order correlations in the data. Consequently, only a small number of data points are misclassified into other clusters. In contrast, the cluster structures formed by other comparison methods exhibit noticeable data point drift (e.g., Figure (d)) or partial structural loss (e.g., Figure (a)).



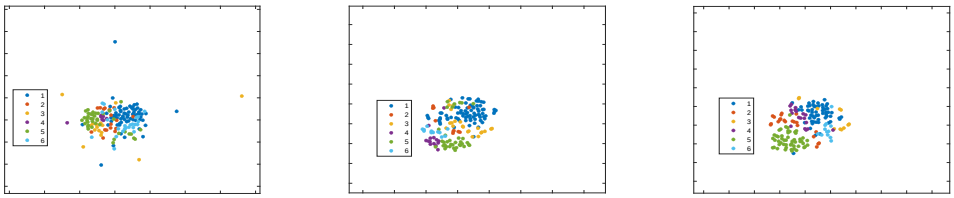
(a) Ours ACC = 94.67% (b) CDIMVC ACC = 92.00% (c) MLKM ACC= 88.17%

(d) Ours ACC = 94.67% (e) CDIMVC ACC = 92.00% (f) MLKM ACC= 88.17%

Figure 2: Visualization results of similarity matrices and confusion matrices for SLT-CMVC, CDIMVC, and MLKM on dataset Synthetic3d.

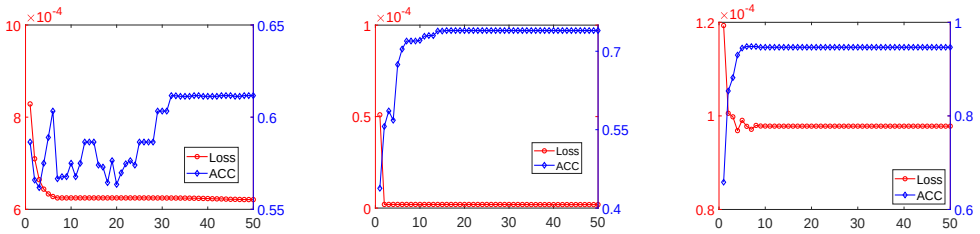


(a) MLKM ACC = 54.01% (b) LAIMVC ACC = 57.15% (c) SLT-CMVC ACC = 61.51%



(d) DEC ACC = 64.29% (e) ASIAMVC ACC = 66.27% (f) SLT-CMVC ACC = 73.96%

Figure 3: t-SNE processed visualization results for each method



(a) Wiki (b) 3Sources (c) Synthetic3d

Figure 4: The objective function values of the proposed method across iterations

5.4 Convergence Analysis

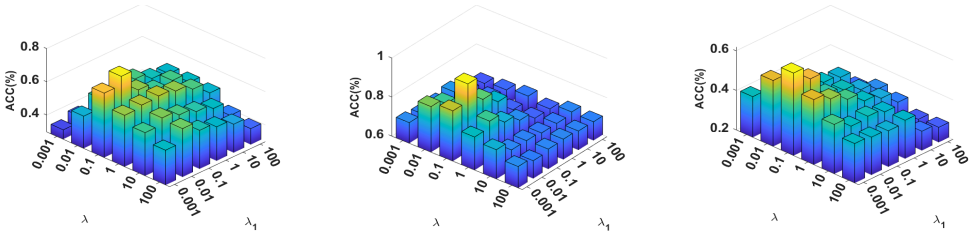
This paper proposes an alternating optimization algorithm for efficiently solving the proposed SLT-CMVC algorithm. By iteratively optimizing each subproblem sequentially, the algorithm achieves a closed-form solution while ensuring lower bound convergence. Figure 4 displays the convergence curves of this algorithm on three datasets: As can be seen from the figure, the loss value decreases rapidly and stabilizes below the lower bound, while the ACC steadily improves as the number of iterations increases, indicating that the algorithm exhibits excellent convergence performance.

5.5 Parameters Analysis

The model includes two balancing hyperparameters: λ and λ_1. To analyze the parameter sensitivity of the model, this paper fixes one parameter while performing a grid search on the other, with the parameter ranges set to $\{0.001, 0.01, 0.1, 1, 10, 100\}$, and uses ACC as the quantitative evaluation metric. As can be seen from the Figure 5, this method exhibits strong robustness with respect to λ and λ_1, maintaining stable performance across most datasets.

5.6 Ablation Study

We conducted ablation experiments on SLT-CMVC to evaluate the effectiveness of each module. Specifically, to evaluate the contributions of the PMM and TCLM modules to model performance, we conducted ablation experiments on 3Sources, Synthetic3d, and Wiki datasets; the corresponding results are shown in Table 2.



(a) 3Sources (b) Synthetic3d (c) Wiki

Figure 5: Clustering performance of the proposed method when parameters λ and λ_1 take different values

Table 2: Ablation results of the SLT-CMVC.

Components \ Datasets		3Sources			Synthetic3d			Wiki		
PMM	TCLM	ACC	NMI	PUR	ACC	NMI	PUR	ACC	NMI	PUR
✓		51.03	38.94	58.14	70.19	45.81	60.79	35.21	28.15	26.51
	✓	56.12	41.21	57.65	75.18	54.79	62.78	42.18	32.42	29.51
✓	✓	**73.96**	**64.85**	**75.15**	**94.67**	**80.31**	**94.67**	**61.51**	**56.26**	**62.53**

After removing the PMM module, the objective function for SLT-CMVC is defined as:

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik}\left\|x_i^t - v_j^t\right\|_2^2 + \lambda u_{ik}^t \ln u_{ik}^t + \|\mathcal{U}\|_{LBR} \\ &s.t.\, \mathcal{U} = \left\{\mathbf{U}^t, \mathbf{U}^{t-1}\right\}, \left\{u_{ik}^t, u_{ik}^{t-1}\right\} \in (0,1)\,. \end{aligned} \tag{6}$$

After removing the TCLM module, the objective function for SLT-CMVC is defined as:

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik}\left\|x_i^t - v_j^t\right\|_2^2 + \lambda u_{ik}^t \ln u_{ik}^t + \lambda_1\left(\sum_{i=1}^{n}\sum_{j=1}^{n} q_i^T q_j \left\|u_i^t - u_j^{t-1}\right\|_2^2 + q_{ik}\ln q_{ik}\right) \\ &s.t.\, \left\{u_{ik}^t, u_{ik}^{t-1}, q_{ik}\right\} \in (0,1)\,. \end{aligned} \tag{7}$$

The results in Table 2 demonstrate that the complete SLT-CMVC consistently outperforms other degenerate algorithms, confirming that tensor learning and consistency information learning both exert positive effects on continual clustering.

6 Conclusion

This paper proposes the SLT-CMVC algorithm to address the issue of feature information loss in CMVC tasks. Specifically, SLT-CMVC learns consistency information across different views through the inner product of probability vectors and pioneers the use of tensor techniques to uncover higher-order correlations between views. This approach corrects errors in the clustering structure while preserving valid clustering components. The algorithm relies solely on the final clustering results to guide subsequent clustering processes, eliminating the need to retain any feature information. Extensive experiments on multiple benchmark datasets demonstrate the effectiveness of SLT-CMVC in utilizing newly arriving view information under restricted historical access while maintaining stable clustering evolution.

7 Limitations

A limitation of the algorithm is that its performance degrades in scenarios involving misaligned samples. In such cases, a viable approach is to use the clustering results to analyze the similarities among historical samples and align the samples in the stream view through similarity propagation. Therefore, future research will focus on continuous clustering in misaligned scenarios.

AI USE STATEMENT

No generative AI tools were used in conducting this research or preparing this manuscript.

ETHICS STATEMENT

This study does not involve human subjects, animal experiments, or the collection of personally identifiable information. All data used in this study are publicly available benchmark datasets used for continual multi-view clustering research. We have followed the original usage guidelines for these datasets and have used them solely for the academic evaluation of machine learning algorithms. As this study focuses on representation learning and clustering methodology, it does not involve decision-making applications that could directly impact individuals or social groups. We have made every effort to ensure that the proposed methods are developed and evaluated in a responsible and transparent manner.

REPRODUCIBILITY STATEMENT

The data used in this study is publicly available, and the source code will be released after it has been reviewed.

REFERENCES

Pham Dinh Ato and Le Thi Hoai An. Convex analysis approach to d.c. programming: Theory, algorithms and applications. *Acta Mathematica Vietnamica*, 22(1):p.289–356, 1997.

Jianpeng Chen, Yawen Ling, Jie Xu, Yazhou Ren, Shudong Huang, Xiaorong Pu, Zhifeng Hao, Philip S. Yu, and Lifang He. Variational graph generator for multiview graph clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(6):11078–11091, 2025. doi: 10.1109/TNNLS.2024.3524205.

Yu Feng, Weixuan Liang, Xinhang Wan, Jiyuan Liu, Miaomiao Li, and Xinwang Liu. Incremental multi-view clustering: Exploring stream-view correlations to learn consistency and diversity. *IEEE Transactions on Knowledge and Data Engineering*, 37(12):7226–7239, 2025a. doi: 10.1109/TKDE.2025.3605594.

Yu Feng, Weixuan Liang, Xinhang Wan, Jiyuan Liu, Suyuan Liu, Qian Qu, Renxiang Guan, Huiying Xu, and Xinwang Liu. Incremental nyström-based multiple kernel clustering. In *Association for the Advancement of Artificial Intelligence*, pages 16613–16621. AAAI Press, 2025b. ISBN 978-1-57735-897-8. doi: 10.1609/aaai.v39i16.33825. URL https://doi.org/10.1609/aaai.v39i16.33825.

Shengzhao Guo, Zhenyu Ma, Jingyu Wang, Feiping Nie, and Xuelong Li. Fast multiview co-clustering in unified subspace. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 1–1, 2025. doi: 10.1109/TCSVT.2025.3613381.

Misha E. Kilmer, Karen Braman, Ning Hao, and Randy C. Hoover. Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIAM Journal on Matrix Analysis and Applications*, 34(1):148–172, 2013. doi: 10.1137/110837711. URL https://doi.org/10.1137/110837711.

Feijiang Li, Jieting Wang, Liuya zhang, Yuhua Qian, Shuai jin, Tao Yan, and Liang Du. k-hyperedge medoids for clustering ensemble. In *Association for the Advancement of Artificial Intelligence*, pages 18271–18278. AAAI Press, 2025a.

Haoran Li, Zhenwen Ren, Yulan Guo, Jiali You, and Xiaojian You. Lsvc: A lifelong learning approach for stream-view clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(5):8714–8727, 2025b. doi: 10.1109/TNNLS.2024.3439394.

Zhouchen Lin, Risheng Liu, and Zhixun Su. Linearized alternating direction method with adaptive penalty for low-rank representation. In *Proceedings of the 25th International Conference on Neural Information Processing Systems*, NIPS'11, page 612–620, Red Hook, NY, USA, 2011. Curran Associates Inc. ISBN 9781618395993.

Jiyuan Liu, Xinwang Liu, Yuexiang Yang, Qing Liao, and Yuanqing Xia. Contrastive multi-view kernel learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8):9552–9566, 2023. doi: 10.1109/TPAMI.2023.3253211.

Feiping Nie, Runxin Zhang, Yu Duan, and Rong Wang. Unconstrained fuzzy c-means based on entropy regularization: An equivalent model. *IEEE Transactions on Knowledge and Data Engineering*, 37(2):979–990, 2025a. doi: 10.1109/TKDE.2024.3516085.

Feiping Nie, Runxin Zhang, Weizhong Yu, and Xuelong Li. Unconstrained fuzzy c-means algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(5):3440–3451, 2025b. doi: 10.1109/TPAMI.2025.3532357.

Jihong Ouyang, Qingyi Meng, Ximing Li, Zhengjie Zhang, Bo Fu, and Wenting Wang. Correlation matters in deep clustering: Transforming clustering into self-supervised multi-label learning. *IEEE Transactions on Big Data*, 12(2):496–508, 2026.

Qian Qu, Xinhang Wan, Weixuan Liang, Jiyuan Liu, Yu Feng, Huiying Xu, Xinwang Liu, and En Zhu. A lightweight anchor-based incremental framework for multi-view clustering. In *ACM International Conference on Multimedia*, MM '24, page 8652–8661, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400706868. doi: 10.1145/3664647.3680837. URL https://doi.org/10.1145/3664647.3680837.

Qian Qu, Xinhang Wan, Jiyuan Liu, Xinwang Liu, and En Zhu. Anchor-guided sample-and-feature incremental alignment framework for multi-view clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 1–1, 2025. doi: 10.1109/TCSVT.2025.3614344.

Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008. URL http://jmlr.org/papers/v9/vandermaaten08a.html.

Wouter Van Gansbeke, Simon Vandenhende, Stamatios Georgoulis, Marc Proesmans, and Luc Van Gool. Scan: Learning to classify images without labels. In *European Conference on Computer Vision*, page 268–285, Berlin, Heidelberg, 2020. Springer-Verlag. ISBN 978-3-030-58606-5. doi: 10.1007/978-3-030-58607-2_16. URL https://doi.org/10.1007/978-3-030-58607-2_16.

Xinhang Wan, Bin Xiao, Xinwang Liu, Jiyuan Liu, Weixuan Liang, and En Zhu. Fast continual multi-view clustering with incomplete views. *IEEE Transactions on Image Processing*, 33:2995–3008, 2024. doi: 10.1109/TIP.2024.3388974.

Xinhang Wan, Jiyuan Liu, Hao Yu, Qian Qu, Ao Li, Xinwang Liu, Ke Liang, Zhibin Dong, and En Zhu. Contrastive continual multiview clustering with filtered structural fusion. *IEEE Transactions on Neural Networks and Learning Systems*, 36(8):14829–14842, 2025. doi: 10.1109/TNNLS.2024.3502455.

Huibing Wang, Mingze Yao, Yawei Chen, Yunqiu Xu, Haipeng Liu, Wei Jia, Xianping Fu, and Yang Wang. Manifold-based incomplete multi-view clustering via bi-consistency guidance. *IEEE Transactions on Multimedia*, 26:10001–10014, 2024. doi: 10.1109/TMM.2024.3405650.

Huibing Wang, Yawei Chen, Mingze Yao, Wenzhe Liu, Jinjia Peng, and Xianping Fu. Tensor completion framework by graph refinement for incomplete multi-view clustering. *IEEE Transactions on Multimedia*, 27:9385–9398, 2025a. doi: 10.1109/TMM.2025.3613125.

Huibing Wang, Yawei Chen, Mingze Yao, Qian Liu, Guangqi Jiang, and Xianping Fu. Tensorized parameter-free multi-view spectral clustering based on fair representation learning. *IEEE Transactions on Multimedia*, pages 1–13, 2026a. doi: 10.1109/TMM.2026.3668481.

Huibing Wang, Luyan Cui, Mingze Yao, and Xianping Fu. Tensorized fine-grained incomplete multiview clustering via intrinsic structure recovery. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pages 1–14, 2026b. doi: 10.1109/TSMC.2026.3678289.

Jing Wang, Songhe Feng, Kristoffer Knutsen Wickstrøm, and Michael C. Kampffmeyer. Adaptcmvc: Robust adaption to incremental views in continual multi-view clustering. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10285–10294, 2025b.

Qianqian Wang, Mengping Jiang, Zhengming Ding, and Quanxue Gao. Unified k-means clustering with label-guided manifold learning. In *International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pages 63186–63199. PMLR, 13–19 Jul 2025c. URL https://proceedings.mlr.press/v267/wang25ao.html.

Danyang Wu, Zhenkun Yang, Jitao Lu, Jin Xu, Xiangmin Xu, and Feiping Nie. Ebmgc-gnf: Efficient balanced multi-view graph clustering via good neighbor fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):7878–7892, 2024. doi: 10.1109/TPAMI.2024.3398220.

Junyuan Xie, Ross B. Girshick, and Ali Farhadi. Unsupervised deep embedding for clustering analysis. *ArXiv*, abs/1511.06335, 2015. URL https://api.semanticscholar.org/CorpusID:6779105.

Xiaoqiang Yan, Yiqiao Mao, Yangdong Ye, and Hui Yu. Incremental multiview clustering with continual information bottleneck method. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 55(1):295–306, 2025. doi: 10.1109/TSMC.2024.3465039.

Hongwei Yin, Wenjun Hu, Zhao Zhang, Jungang Lou, and Minmin Miao. Incremental multi-view spectral clustering with sparse and connected graph learning. *Neural Networks*, 144:260–270, 2021. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2021.08.031. URL https://www.sciencedirect.com/science/article/pii/S0893608021003440.

8 APPENDIX

The supplementary material provides deeper insights into the SLT-CMVC model through detailed optimization procedures, additional experimental results, and further analyses.

8.1 ANALYSIS OF THE LOW-BIAS REFINED NORM

Theoretical analysis: Let the function $f_{LBR}(x) = ln(2 - e^{\frac{-x}{\varepsilon}})$. As $x \to 0$, we have $f_{LBR}(x) \to 0$. When $x \to +\infty$, the function converges to $f_{LBR}(x) \to ln2$. This bounded saturation property effectively suppresses the penalty ceiling for large singular values, thereby preserving the more dominant structural features in the data. Consequently, the LBR norm operator suppresses the noise represented by small singular values while eliminating excessive shrinkage bias on significant features, striking a balance between tensor sparsity and the preservation of structural information. It is worth noting that, in all experiments, we set $\varepsilon = 2$ to make the curve flatter and the optimization process smoother.

8.2 OPTIMIZATION PROCEDURES

The optimization problem in Eq. (4) is not jointly convex when all variables are considered simultaneously. Therefore, we employ an alternating optimization algorithm to optimize each variable sequentially, presenting an overall optimization framework along with its time/space complexity analysis:

Initial stage:

$$\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik} \left\| x_i - v_j \right\|_2^2 + \lambda u_{ik} \ln u_{ik} \quad s.t. u_{ik} \in (0,1)\,. \tag{8}$$

1) u_{ik} **Sub-Step:** Fixing the other variables, the optimization problem for u_{ik} in the Eq. (8) can be rewritten as:

$$\min_{u} \sum_{i=1}^{n}\sum_{k=1}^{c} d_{ik}u_{ik} + \lambda u_{ik} \ln u_{ik} \quad s.t. \forall i, \sum_{k=1}^{c} u_{ik} = 1. \tag{9}$$

where $d_{ik} = \left\| x_i - v_j \right\|_2^2$. We then construct Lagrange function as:

$$\mathcal{L}\left(u_{ik}, \alpha\right) = \sum_{k=1}^{c} d_{ik}u_{ik} + \lambda u_{ik} \ln u_{ik} + \alpha \left(\sum_{k=1}^{c} u_{ik} - 1 \right) \tag{10}$$

where α is the Lagrange multiplier for the equality constraint. The derivatives of Eq. (10)with respect to u_{ik} and α should be zero, specifically as follows:

$$\frac{\partial \mathcal{L}\left(u_{ik}, \alpha\right)}{\partial u_{ik}} = d_{ik} + \lambda \left(\ln u_{ik} + 1 \right) + \alpha = 0 \tag{11}$$

$$\frac{\partial \mathcal{L}\left(u_{ik}, \alpha\right)}{\partial \alpha} = \sum_{k=1}^{c} u_{ik} - 1 = 0 \tag{12}$$

According to Eq. (11), we express u_{ik} as:

$$u_{ik} = e^{-\frac{1}{\lambda}(d_{ik}+\alpha)-1} \tag{13}$$

we substitute the expression related to α (Eq. (12)) into Eq. (13):

$$u_{ik} = \frac{e^{-\frac{1}{\lambda}d_{ik}}}{\sum_{m=1}^{c} e^{-\frac{1}{\lambda}d_{im}}} \tag{14}$$

2) v_i **Sub-Step:** Fixing the other variables, the optimization problem for v_i in the Eq. (8) can be rewritten as:

$$\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik} \left\| x_i - v_j \right\|_2^2 \tag{15}$$

Taking the derivative of Eq. (15) with respect to v_i and setting as zero, the solution is obtained as:

$$v_i = \frac{\sum_{i=1}^{n} u_{ik} x_i}{\sum_{i=1}^{n} u_{ik}} \tag{16}$$

Continual Clustering Stage: The overall objective function is as follows:

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik} \left\| x_i^t - v_j^t \right\|_2^2 + \lambda u_{ik}^t \ln u_{ik}^t + \|\mathcal{M}\|_{LBR} \\ &+\lambda_1 \left(\sum_{i=1}^{n}\sum_{j=1}^{n} q_i^T q_j \left\| u_i^t - u_j^{t-1} \right\|_2^2 + q_{ik} \ln q_{ik} \right) + \frac{\mu}{2} \left\| \mathcal{M} - \mathcal{U} + \mathcal{A}/\mu \right\|_F^2 \\ &s.t. \mathcal{U} = \mathcal{M}, \mathcal{U} = \left\{ \mathbf{U}^t, \mathbf{U}^{t-1}, \mathbf{Q} \right\}, \left\{ u_{ik}^t, u_{ik}^{t-1}, q_{ik} \right\} \in (0,1). \end{aligned} \tag{17}$$

1) u_{ik}^t **Sub-Step:** Fixing the other variables, the optimization problem for u_{ik}^t in the Eq. (17) can be rewritten as:

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik} \left\| x_i^t - v_j^t \right\|_2^2 + \lambda u_{ik}^t \ln u_{ik}^t \\ &+\lambda_1 \sum_{i=1}^{n}\sum_{j=1}^{n} q_i^T q_j \left\| u_i^t - u_j^{t-1} \right\|_2^2 + \frac{\mu}{2} \left\| \mathcal{M} - \mathcal{U} + \mathcal{A}/\mu \right\|_F^2 \\ &s.t. \forall i, \sum_{k=1}^{c} u_{ik}^t = 1, 0 < u_{ik}^t < 1. \end{aligned} \tag{18}$$

By setting the derivative of Eq. (18) to zero during the derivation process of Eq. (14), we obtain:

$$d_{ik} + \lambda \left(\ln u_{ik} + 1 \right) + \alpha + \lambda_1 h_{ij} + \mu o_{jk} = 0 \tag{19}$$

here $h_{ij} = \sum_{j=1}^{n}\sum_{k=1}^{c} s_{ij} * u_{jk}^{t-1}$, $o_{jk} = \sum_{j=1, j\neq i}^{n} (u_{jk}^t/2 - b_{jk})$, $\mathcal{B} = \mathcal{M} - \mathcal{A}/\mu$ and $s_{ij} = q_i^T * q_j$. Based on Eq .(14), we can obtain the final result:

$$u_{ik}^t = \frac{e^{-\frac{1}{\lambda}(d_{ik} + \lambda_1 h_{ij} + \mu o_{jk})}}{\sum_{m=1}^{c} e^{-\frac{1}{\lambda}(d_{im} + \lambda_1 h_{im} + \mu o_{jm})}} \tag{20}$$

2) v_i^t **Sub-Step:** Similar to the update for Eq. (16):

$$v_i^t = \frac{\sum_{i=1}^{n} u_{ik}^t {x_i}^t}{\sum_{i=1}^{n} u_{ik}^t} \tag{21}$$

3) $\mathcal{M}$ **Sub-Step:** By implementing ALM, the objective function with respect to $\mathcal{M}$ becomes:

$$\|\mathcal{M}\|_{LBR} + \frac{\mu}{2} \left\| \mathcal{M} - \mathcal{U} + \mathcal{A}/\mu \right\|_F^2 \tag{22}$$

Accordingly, the solution is:

$$\mathcal{M}^* = \mathcal{S} * ifft\left(\Omega_f\left(P_f \right), [\,], 3 \right) * \mathcal{V}^T \tag{23}$$

In this formulation, $ifft\left(\Omega_f\left(P_f \right), [\,], 3 \right)$denotes a tensor whose frontal slices are diagonal matrices. Each diagonal element $\Omega_f\left(P_f \right)$ is obtained by solving the following optimization problem:

$$\Omega_f\left(P_f^k(i,i) \right) = \arg\min_{x \geq 0} \frac{1}{2} \left(x - P_f^k(i,i) \right)^2 + f(x) \tag{24}$$

where the function $f(x)$ is defined as $ln(2 - e^{\frac{-x}{\varepsilon}})$.

Eq. 24 comprises both convex and concave components, and can be addressed using Difference of Convex (DC) programming Ato and An (1997), leading to the following closed-form solution:

$$\eta^{iter+1} = \left(P_f^k(i,i) - \frac{\partial f\left(\eta^{iter+1} \right)}{\mu} \right)_+ \tag{25}$$

Algorithm 1 : SLT-CMVC

Input: The data matrices $\{\mathbf{X}^t\}_{t=1}^m$ and cluster number c.
Output: Soft Label $\mathbf{U}^t$

```
for t = 1 to m do
  if t = 1 then
    while not converged do
      Update U^1, V^(1) using Eq. (14), Eq. (16);
    end while
  else
    while not converged do
      Update U^t, V^t using Eq. (20), Eq. (21);
      Update M, A using Eq. (22), Eq. (26);
      Update Q via solving Eq. (29);
    end while
  end if
end for
return Clustering Results: max(U^t)
```

where $\eta = \Omega_f\left(P_f^k(i,i)\right)$ and $iter$ denotes the iteration index.

4) $\mathcal{A}$ **Sub-Step:** By setting its partial derivative to zero, the optimal value of $\mathcal{A}$ is given by:

$$\mathcal{A} = \mathcal{A} + \mu(\mathcal{U} - \mathcal{M}) \tag{26}$$

5) Q **Sub-Step:** Fixing the other variables, the optimization problem for Q in the Eq. (17) can be rewritten as:

$$\begin{aligned}\mathcal{J}_q &= \lambda_1\left(\sum_{i=1}^{n}\sum_{j=1}^{n} q_i^T q_j \left\|u_i^t - u_j^{t-1}\right\|_2^2 + q_{ik}\ln q_{ik}\right) + \frac{\mu}{2}\left\|\mathcal{M} - \mathcal{U} + \mathcal{A}/\mu\right\|_F^2 \\ &\Leftrightarrow \min_{\mathbf{Q}} \lambda_1(Tr(\mathbf{Q}^T\mathbf{D}\mathbf{Q}) + q_{ik}\ln q_{ik}) + \frac{\mu}{2}\left\|\mathcal{M} - \mathcal{U} + \mathcal{A}/\mu\right\|_F^2\end{aligned} \tag{27}$$

where $d_{ij} = \left\|u_i^t - u_j^{t-1}\right\|_2^2$. Since Eq. (27) represents a constrained optimization problem, inspired by the SCAN Van Gansbeke et al. (2020) method, we transform it into a new objective function based on the Lagrange multiplier approach:

$$\begin{aligned}\mathcal{J}_q &= \lambda_1(\sum_{i=1}^{n}\sum_{j\neq i}^{n}\varphi\left(\mathbf{Q}_{,:k}\right)d_{ij} + q_{ik}lnq_{ik}) + \mu\sum_{j=1,j\neq i}^{n}(q_{jk}/2 - b_{jk})d_{ij} \quad s.t.\forall i, \sum_{k=1}^{c} q_{ik} = 1. \\ \Leftrightarrow \mathcal{J}_q &= \lambda(\sum_{i=1}^{n}\sum_{j\neq i}^{n}\varphi\left(\mathbf{Q}_{,:k}\right)d_{ij} + q_{ik}\ln q_{ik}) + \alpha\left(1 - \sum_{k=1}^{c} q_{ik}\right) + \mu\sum_{j=1,j\neq i}^{n}(q_{jk}/2 - b_{jk})d_{ij}\end{aligned} \tag{28}$$

where $\varphi\left(\mathbf{Q}_{,:k}\right) = \sum_{k=1}^{c} q_{ik}q_{jk}$ Taking the derivative of Eq. (28) results in:

$$q_{ik} = \frac{e^{-\frac{1}{\lambda_1}(q_{jk}d_{ik} + \mu(q_{jk}/2 - b_{jk}))}}{\sum_{m=1}^{c} e^{-\frac{1}{\lambda_1}q_{jm}d_{im}}} \tag{29}$$

Finally, the whole optimization procedure is summarized in the Algorithm 1, which updates each variable iteratively until convergence.

The experimental environment consisted of a Windows 10 PC equipped with an Intel Core i5-8400 2.80GHz processor and 16GB of memory. The compiler is MATLAB R2024a.

8.3 Convergence proof

In this section, we provide a rigorous proof of the convergence of the proposed algorithm. Let $\mathcal{U} = \text{stack}(\mathbf{U}^t, \mathbf{U}^{t-1}, \mathbf{Q}) \in \mathbb{R}^{n\times c\times 3}$ be a third-order tensor constructed by stacking the membership

and transition matrices. The Augmented Lagrangian Function (ALF) is defined as:

$$\begin{aligned}\mathcal{L}(\mathcal{U},\mathcal{V},\mathcal{M},\mathcal{A}) = &\sum_{i=1}^{n}\sum_{k=1}^{c} u_{ik}\|x_i^t - v_j^t\|_2^2 + \lambda u_{ik}\ln u_{ik} + \|\mathcal{M}\|_{LBR} \\ &+ \lambda_1\left(\sum_{i,j} q_i^T q_j \|u_i^t - u_j^{t-1}\|_2^2 + q_{ik}\ln q_{ik}\right) + \frac{\mu}{2}\|\mathcal{M} - \mathcal{U} + \mathcal{A}/\mu\|_F^2\end{aligned} \tag{30}$$

where $\mathcal{A}$ is the Lagrange multiplier tensor and $\mu > 0$ is the penalty parameter.

Lemma 1 (Sufficient Descent)**.** *For a sufficiently large penalty parameter μ, the sequence $\{\mathcal{L}(\Theta^{iter})\}$ generated by the algorithm, where $\Theta = \{\mathcal{U},\mathcal{V},\mathcal{M},\mathcal{A}\}$, satisfies the sufficient descent property over a full iteration cycle. Specifically, for the non-convex LBR norm, we employ the Majorization-Minimization (MM) strategy by linearizing $f(\sigma) = \ln(2 - e^{-\sigma/\epsilon})$ at σ^{iter}:*

$$f(\sigma) \le f(\sigma^i ter) + w^{iter}(\sigma - \sigma^{iter}), \quad w^{iter} = \frac{e^{-\sigma^{iter}/\epsilon}}{\epsilon(2 - e^{-\sigma^{iter}/\epsilon})} \tag{31}$$

The minimization of this surrogate upper bound guarantees $\mathcal{L}(\ldots,\mathcal{M}^{iter+1},\mathcal{A}^{iter}) \le \mathcal{L}(\ldots,\mathcal{M}^{iter},\mathcal{A}^{iter})$.

Lemma 2 (Sequence Boundedness)**.** *The sequence of iterates $\{\mathcal{U}^{iter},\mathcal{V}^{iter},\mathcal{M}^{iter},\mathcal{A}^{iter}\}$ is bounded.*

Proof. Since $u_{ik}, q_{ik} \in (0,1)$, $\mathcal{U}$ is naturally bounded within a compact simplex. The update of $\mathcal{M}$ via the weighted SVT operator (a proximity operator) on a compact manifold ensures its boundedness. Consequently, the multiplier $\mathcal{A}$ and centroids $\mathcal{V}$ remain within bounded regions. □

Theorem 1. *The sequence $\{\Theta^{iter}\}$ generated by the algorithm converges to a KKT stationary point.*

Proof. First, $\mathcal{L}$ is a semi-algebraic function consisting of polynomials, logs, and exponentials, thus satisfying the **KL property**. Second, we establish the **Relative Error Condition**. As detailed in the Appendix, there exists a constant $\rho > 0$ such that:

$$\text{dist}(0, \partial\mathcal{L}(\Theta^{iter+1})) \le \rho\|\Theta^{iter+1} - \Theta^{iter}\| \tag{32}$$

Given the sufficient descent and boundedness (Lemmas 1-2), the sequence $\{\Theta^{iter}\}$ is a Cauchy sequence. According to KL theory, the iterate sequence converges to a stationary point Θ^* satisfying the KKT conditions. □

To prove $\text{dist}(0, \partial\mathcal{L}(\Theta^{iter+1})) \le \rho\|\Theta^{iter+1} - \Theta^{iter}\|$, we analyze the partial subgradients:

1. Subgradient of $\mathcal{U}$: From the optimality condition $\nabla_{\mathcal{U}}\mathcal{L}(\mathcal{U}^{iter+1},\mathcal{V}^{iter+1},\mathcal{M}^{iter},\mathcal{A}^{iter}) = 0$, the subgradient at the next full state Θ^{iter+1} is:

$$\begin{aligned}\|\nabla_{\mathcal{U}}\mathcal{L}(\Theta^{iter+1})\| &= \|\mu(\mathcal{M}^{iter} - \mathcal{M}^{iter+1}) + (\mathcal{A}^{iter+1} - \mathcal{A}^{iter})\| \\ &\le \mu\|\mathcal{M}^{iter+1} - \mathcal{M}^{iter}\| + \mu\|\mathcal{M}^{iter+1} - \mathcal{U}^{iter+1}\|\end{aligned} \tag{33}$$

2. Subgradient of $\mathcal{M}$: Using the optimality of the $\mathcal{M}$-step and the update rule $\mathcal{A}^{iter+1} = \mathcal{A}^{iter} + \mu(\mathcal{M}^{iter+1} - \mathcal{U}^{iter+1})$, we have:

$$\|\partial_{\mathcal{M}}\mathcal{L}(\Theta^{iter+1})\| = \|\mathcal{A}^{iter+1} - \mathcal{A}^{iter}\| = \mu\|\mathcal{M}^{iter+1} - \mathcal{U}^{iter+1}\| \tag{34}$$

By summing these bounds and utilizing the Lipschitz continuity of the smooth components, we conclude that the norm of the total subgradient is bounded by the distance between successive iterates, which satisfies the KL requirement for global convergence.

Table 3: Summary of the used MVC benchmark datasets

Dataset	Samples	Views	Clusters	Dimension
3Sources	169	3	6	3560, 3631, 3068
MSRC	210	4	7	576, 512, 256, 254
Synthetic3d	600	3	3	600, 600, 600
Wiki	2866	2	10	128, 10
Scene15	4485	3	15	20, 59, 40
Fashion	10000	3	10	784, 784, 784
Reuters	18758	5	6	21531, 24892, 34251, 15506, 11547

8.4 Additional Experimental Results

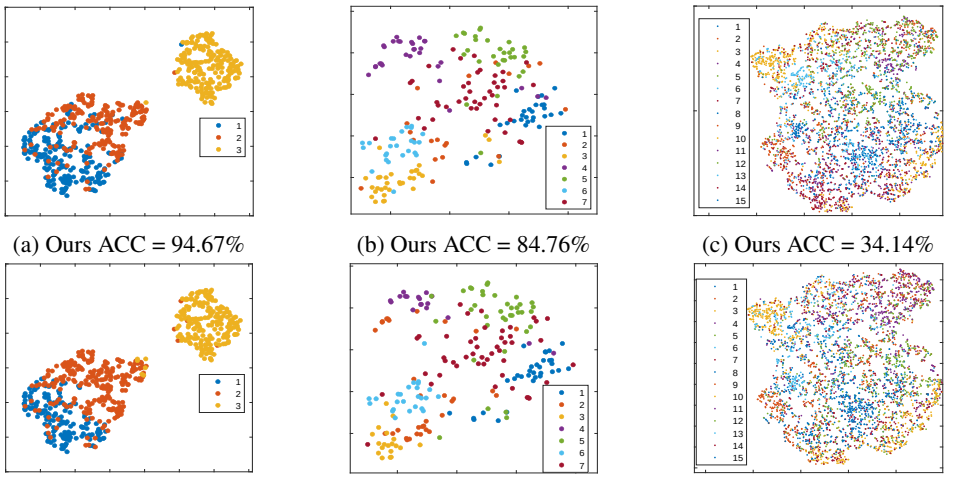
Only a subset of the experimental figures was presented in the main text. In this subsection, we provide the remaining figures to offer a more complete view of the experimental results. Specific information for all datasets is presented in Table 3.

The ablation experiments on the remaining datasets are shown in Table 4.

Figure 8 shows a parameter-sensitivity analysis across four datasets, revealing that the proposed method is sensitive outside the specified interval.

The convergence results for the remaining datasets are shown in Figure 9. Figure 6 and Figure 10 show the t-SNE visualizations of the proposed algorithm and the suboptimal algorithm on the five datasets: MSRC, Synthetic3d, Scene15, Fashion, and Reuters.

The Figure 7 shows the confusion matrices for the proposed algorithm and the suboptimal algorithm across the six datasets 3Sources, MSRC, Synthetic3d, Scene15, Fashion, and Reuters.



(a) Ours ACC = 94.67% (b) Ours ACC = 84.76% (c) Ours ACC = 34.14%

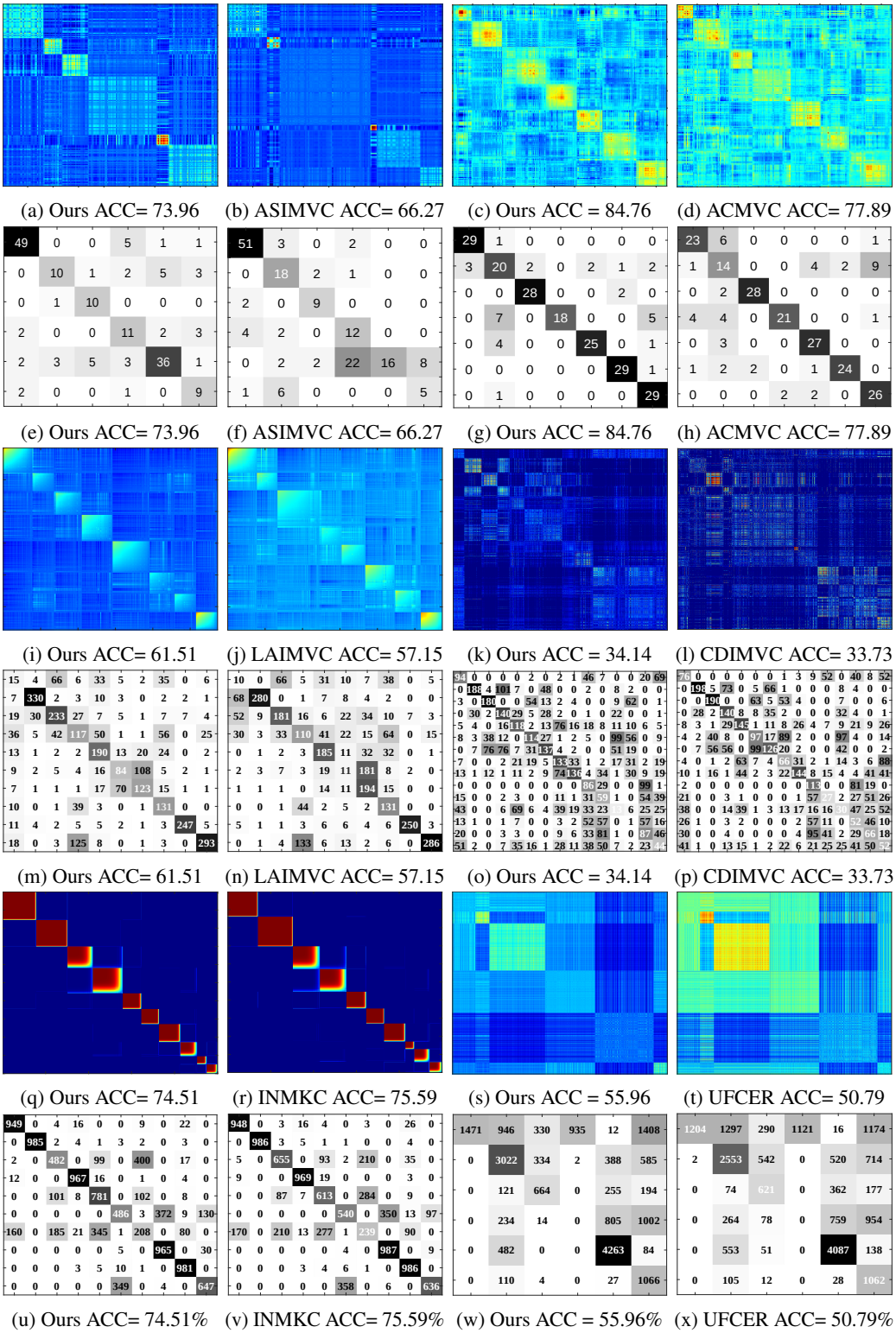
(d) CDIMVC ACC= 92.00% (e) ACMVC ACC= 77.98% (f) CDIMVC ACC = 33.73%

Figure 6: t-SNE processed visualization results for each method.

Table 4: Ablation results of the SLT-CMVC.

Components		MSRC			Scene15			Fashion			Reuters		
PMM	TCLM	ACC	NMI	PUR	ACC	NMI	PUR	ACC	NMI	PUR	ACC	NMI	PUR
✓		74.52	69.80	75.01	25.29	19.32	27.81	62.55	54.21	60.13	40.15	24.38	49.17
	✓	79.83	74.34	80.15	26.97	18.95	28.11	63.24	58.72	62.15	41.24	26.01	50.00
✓	✓	**84.76**	**76.18**	**85.76**	**34.14**	**35.62**	**38.93**	**74.51**	**73.25**	**75.13**	**55.96**	**37.81**	**60.74**

To provide a more comprehensive evaluation of the practical efficiency of different continual multi-view clustering methods, we have added wall-clock runtime and peak-memory comparisons on three



(a) Ours ACC= 73.96 (b) ASIMVC ACC= 66.27 (c) Ours ACC = 84.76 (d) ACMVC ACC= 77.89

(e) Ours ACC= 73.96 (f) ASIMVC ACC= 66.27 (g) Ours ACC = 84.76 (h) ACMVC ACC= 77.89

(i) Ours ACC= 61.51 (j) LAIMVC ACC= 57.15 (k) Ours ACC = 34.14 (l) CDIMVC ACC= 33.73

(m) Ours ACC= 61.51 (n) LAIMVC ACC= 57.15 (o) Ours ACC = 34.14 (p) CDIMVC ACC= 33.73

(q) Ours ACC= 74.51 (r) INMKC ACC= 75.59 (s) Ours ACC = 55.96 (t) UFCER ACC= 50.79

(u) Ours ACC= 74.51% (v) INMKC ACC= 75.59% (w) Ours ACC = 55.96% (x) UFCER ACC= 50.79%

Figure 7: Visualization results of similarity matrices and confusion matrices.

medium- and large-scale datasets: Scene15, Fashion, and Reuters. All methods were evaluated under the same hardware and software environment. The results are shown in Table 5.

Ours runs stably on three datasets and achieves a reasonable trade-off between performance, runtime, and memory usage; on the Reuters dataset in particular, some methods have already achieved OM.

To ensure a controlled and practically relevant comparison, we selected TNN, Log-Sum, and Schatten-p, which are representative spectral regularizers commonly adopted in tensor low-rank modeling. They cover three distinct families: linear, logarithmic, and power-type spectral penalties. In all variants, we retain the same tensor construction, optimization framework, initialization, stopping criterion, and parameter-selection protocol, and replace only the spectral penalty. This setting allows the effect of the low-rank surrogate itself to be isolated as clearly as possible. The results are shown in Table 6.

Experimental results show that LBR demonstrates superior overall performance across various datasets and consistently outperforms the standard TNN model. This behavior is consistent with LBR's design. TNN imposes a linear penalty on all singular values, which may lead to excessive compression of dominant singular components.

In the CMVC settings, the algorithm processes different views sequentially. Therefore, we examined the impact of the view presentation order on the algorithm. Specifically, we ranked the views in each dataset by quality and set their order from highest to lowest and from lowest to highest. We selected the suboptimal algorithm and the second-best algorithm for comparison; the results are shown in Table 7.

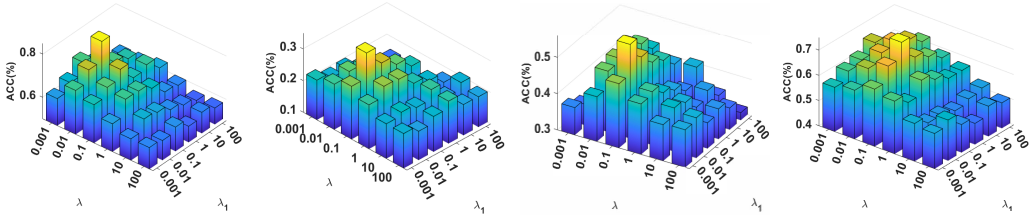
It can be observed that our method outperforms other CMVC methods across various view orders, and the results are robust.

Table 5: Runtime and peak memory consumption of different methods on three datasets.

Dataset	**Metric**	**ACMVC**	**INMKC**	**CDIMVC**	**LAIMVC**	**ASIA-MVC**	**Ours**
Scene15	Runtime (s)	890.17	395.90	710.10	410.57	907.68	563.45
	Peak Memory (MB)	2078.14	853.20	1980.72	724.21	2025.76	1256.02
Fashion	Runtime (s)	OM	83.54	219.84	101.24	302.54	138.06
	Peak Memory (MB)	OM	3279.06	10254.75	2097.85	4798.70	3617.48
Reuters	Runtime (s)	OM	490.40	OM	507.84	OM	984.45
	Peak Memory (MB)	OM	7375.96	OM	6798.25	OM	11737.53

Table 6: Comparison of different regularizers on seven datasets.

Regularizer	3Sources	MSRC	Synthetic3d	Wiki	Scene15	Fashion	Reuters	Average
TNN	70.27	80.78	93.94	57.14	30.24	70.29	50.74	64.77
Log-Sum	72.34	83.49	94.71	58.91	31.77	72.98	52.33	66.65
Schatten-p	71.89	84.57	**95.17**	56.35	33.31	71.45	54.79	66.79
LBR	**73.96**	**84.76**	94.67	**61.51**	**34.14**	**74.51**	**55.96**	**68.50**

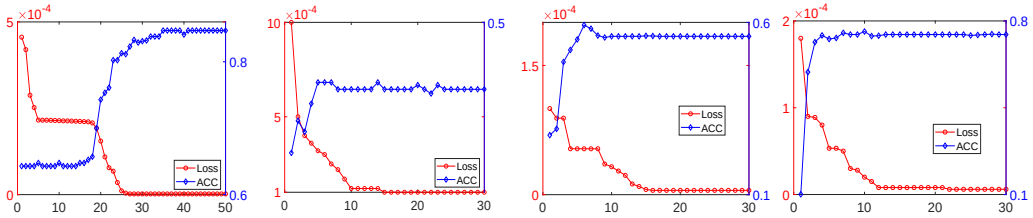


(a) MSRC (b) Scene15 (c) Reuters (d) Fashion

Figure 8: Clustering performance of the proposed method when parameters λ and λ_1 take different values.

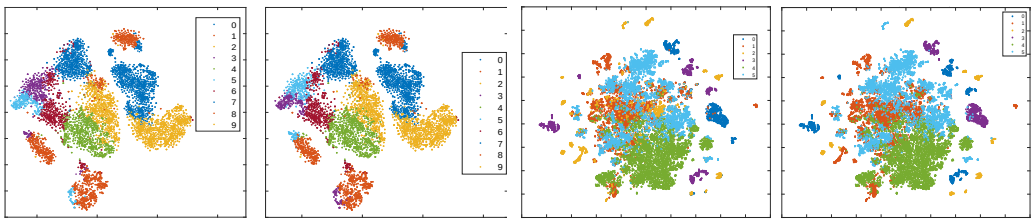
Table 7: Clustering results (ACC%) on four datasets under different view orders

Synthetic3d	INMKC	CDIMVC	Ours	Wiki	ASIAMVC	LAIMVC	Ours
High to Low	90.12	92.00	**94.67**	High to Low	56.02	57.15	**61.51**
Low to High	87.54	88.91	**91.13**	Low to High	52.65	51.34	**57.18**
Scene15	INMKC	CDIMVC	Ours	Reuters	INMKC	LAIMVC	Ours
High to Low	32.54	33.73	**34.14**	High to Low	34.37	40.96	**55.96**
Low to High	30.27	31.44	**33.98**	Low to High	32.73	38.41	**52.35**



(a) MSRC (b) Scene15 (c) Reuters (d) Fashion

Figure 9: The objective function values of the proposed method across iterations.



(a) Ours ACC = 74.51 (b) INMKC ACC = 75.59 (c) Ours ACC = 55.96 (d) UFCER ACC= 50.79

Figure 10: t-SNE processed visualization results for each method.