Multi-View Clustering by Inter-cluster Connectivity Guided Reward

Hao Dai [1 2] **Yang Liu** [1 2] **Peng Su** [1 2] **Hecheng Cai** [1 2] **Shudong Huang** [1 2] **Jiancheng Lv** [1 2]

Abstract

Multi-view clustering has been widely explored for its effectiveness in harmonizing heterogeneity along with consistency in different views of data. Despite the significant progress made by recent works, the performance of most existing methods is heavily reliant on strong priori information regarding the true number of clusters k, which is rarely feasible in real-world scenarios. In this paper, we propose a novel graph-based multi-view clustering algorithm to infer unknown k through a graph consistency reward mechanism. To be specific, we evaluate the cluster indicator matrix during each iteration with respect to diverse k. We formulate the inference process of unknown k as a parsimonious reinforcement learning paradigm, where the reward is measured by inter-cluster connectivity. As a result, our approach is capable of independently producing the final clustering result, free from the input of a predefined cluster number. Experimental results on multiple benchmark datasets demonstrate the effectiveness of our proposed approach in comparison to existing state-of-the-art methods.

1. Introduction

In real-world settings, data are likely derived from varied domains and represented in a multitude of forms, or views (Bickel & Scheffer, 2004). For instance, a document may be represented through audio, images, videos, and text themselves can be available in multiple languages (Wang et al., 2015). With a predominant portion of such data being unlabeled, the role of multi-view clustering becomes pivotal. This approach incorporates information from various views to effectively segregate data into separate categories (Liu et al., 2021).

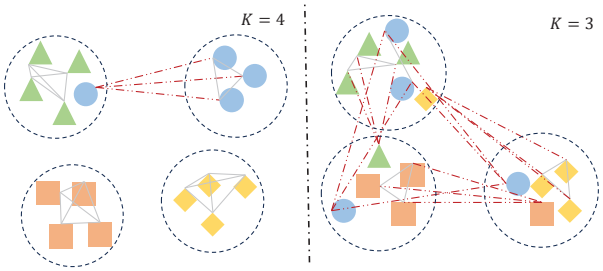


Figure 1. The motivation for designing the reward function for finding unknown k. It is obvious that an erroneous k ($k = 3$ in this case) would dramatically increase the inter-cluster connectivity (as shown by the red dashed line). Hence it is intuitive and reasonable to measure the reward by minimizing the inter-cluster connectivity.

In recent years, significant advancements have been witnessed in multi-view clustering algorithms, concomitant with the burgeoning development of the machine learning domain. Despite considerable advancements in multi-view clustering, a majority of these methods are parametric, necessitating a predefined number of clusters as input. This requirement often proves impractical in real-world scenarios. A natural thought is to adopt non-parametric solutions from single-view clustering. Early approaches primarily relied on conducting k experiments repeatedly and empirically and then determining optimal k by observing unsupervised criteria (Kodinariya et al., 2013; Yuan & Yang, 2019). However, employing these methods in a multi-view setting can be time-consuming. Furthermore, they are primarily developed for convex data. Inspired by (Liu et al., 2023), a feasible approach is to form the decision-making problem of k as a paradigm of reinforcement learning, which is fundamentally a trial-and-error process. Compared to early approaches, clustering under a reinforcement learning paradigm can infer optimal k during the iterative process, significantly reducing time costs. Besides, its capability to automatically infer k provides an advantage over empirical decision, especially in scenarios where certain criteria (i,e, ELBOW) are ambiguous (Umargono et al., 2020; Onumanyi et al., 2022). Nevertheless, transferring the method of (Liu et al., 2023) to a multi-view setting also proves problematic. For one, it demands an additional adjacency matrix as input. More-

[*]Equal contribution [1]College of Computer Science, Sichuan University, Chengdu 610065, China [2]Engineering Research Center of Machine Learning and Industry Intelligence, Ministry of Education, Chengdu 610065, China. Correspondence to: Shudong Huang <huangsd@scu.edu.cn>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

over, the definition of the reward R_k, as the within-cluster square of sum (WSS) for each k, does not perform well in numerous established multi-view clustering approaches.

To overcome these challenges, we propose a novel multi-view clustering algorithm that utilizes a parsimonious reinforcement learning (RL) paradigm to infer k automatically and execute clustering heuristically during each iteration. Considering a supervised scenario, it would be clear whether the clustering result under a given k is distorted with the guidance of true labels. However, true labels are inaccessible in an unsupervised scenario. Hence, we propose to define inter-cluster connectivity to substitude for the role of labels, as shown in Fig. 1. Under the assumption that an optimal k should maintain little inter-cluster connectivity, whereas an incorrect k would significantly increase the inter-cluster connectivity, we can make the inference of k by finding the minimal connectivity it presents. Fig. 2 depicts the framework of our proposed method. We first construct graphs $\mathbf{G}^v$ for each view using mutual nearest neighbors (MNN), and obtain a consensus graph $\mathbf{G}$ with the most authentic information through view voting. After each iteration, $\mathbf{G}$ is used to compute inter-cluster connectivity with indicator matrix $\mathbf{Y}$ to get an immediate reward R with respect to different k. In this framework, the decision module for agnostic k and the clustering module for MVC are separated. Hence, the reward can be integrated into any existing MVC framework and independently make decisions. Besides, our proposed reward offers two main advantages. First, it constitutes a non-sparse reward, meaning it does not necessitate the use of experience replay for accelerating convergence in slow-learning RL. Second, we observe that, aside from initial oscillations during the first few iterations, this reward tends to stabilize as the cluster indicator matrix $\mathbf{Y}$ converges. Consequently, we only need to model a parsimonious reinforcement learning paradigm, i.e., a multi-armed bandit—to adequately address this decision-making problem.

2. Related Work

2.1. Determining the Number of Clusters

In traditional clustering, various methods have been developed to automate the determination of the number of clusters. (Kodinariya et al., 2013; Yuan & Yang, 2019) presents a review of methods for estimating cluster numbers based on unsupervised criteria. For example, the ELBOW calculates cluster distortion, a metric of within-cluster dispersion, for each potential cluster number and identifies the optimal number at the point of significant change in distortion. However, this distortion curve can sometimes be ambiguous. The jump statistic method (Tibshirani et al., 2001) mitigates this by applying rate distortion theory. Nonetheless, these estimation techniques can entail substantial computational overhead as they assess the efficacy of different cluster numbers after the algorithm has been executed. Furthermore, their design primarily suits k-means clustering, rendering them less effective for identifying non-spherical clusters (Rodriguez & Laio, 2014). To address the challenge of clustering data with arbitrary shapes and to obviate the requirement of manually predefining the number of clusters, a number of density-based approaches have been introduced (Rodriguez & Laio, 2014; Ester et al., 1996). These methods emphasize the utilization of data density variations to determine cluster formations.

Recently, several non-parametric methods focused on single-view data have emerged. For instance, (Ronen et al., 2022) proposes a deep clustering method capable of inferring the cluster count k during learning process. In graph clustering, (Liu et al., 2023) utilizes Deep Q-Networks (DQN) (Silver et al., 2017) for reasoning k, and (Zhao et al., 2024) proposes to infer k in a topological-hierarchical way. Although effective in single-view contexts, these methods fall short in multi-view settings as they do not address the heterogeneity issue across different data views. In multi-view setting, (Peng et al., 2019) proposes COMIC which constructs a connection graph and an interpretative representation maintaining the dimensionality of the original data space. This connection graph is then used to guide the clustering process, where connected samples are assigned to the same cluster. While COMIC operates as a non-parametric method, it has a tendency to produce components significantly larger than the true number of categories, resulting in suboptimal clustering result. Motivated by these observations, we propose a reward-driven learning process to accurately infer the number of clusters k in multi-view setting.

2.2. Multi-view Clustering

In recent years, Multi-view Clustering (MVC) (Huang et al., 2021; Wan et al., 2022; Huang et al., 2023; Cai et al., 2023; Tan et al., 2023) has garnered significant attention within the research community. The existing MVC algorithms can be broadly classified into two groups: traditional multi-view clustering methods and deep multi-view clustering methods. Traditional approaches mainly encompass four categories: Matrix factorization-based algorithms (Liu et al., 2013; Zhao et al., 2017; Huang et al., 2020): These techniques aim to uncover a shared latent factor to extract information from multi-view data using non-negative matrix factorization. (Cai et al., 2013) introduced a shared clustering indicator matrix within multi-view context. Kernel-based MVC (Huang et al., 2019): This category employs predefined kernels to handle diverse views, aiming to devise a unified kernel through linear or non-linear combinations of predefined kernels. Graph-based MVC (Li et al., 2021; Tang et al., 2020; Tian et al., 2020): These methods utilize multi-view data to construct graphs that preserve data

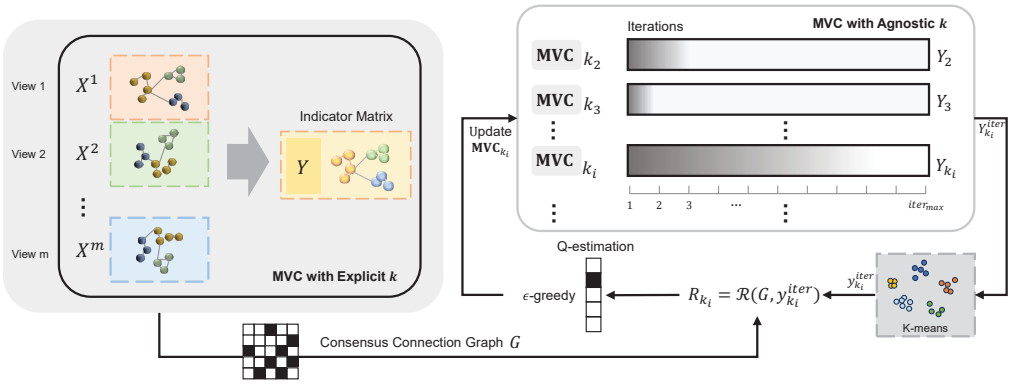


Figure 2. The framework of our proposed method. The left half of the figure depicts a multi-view clustering module under the condition of a known k. It attains an indicator matrix Y derived from executing spectral clustering across similarity graphs (Kang et al., 2020a). The right half of the figure depicts a k inferring module during the iterative process. Specifically, for each possible value of k, we instantiate k different replicas of the algorithm on the left side. During each time step, an epsilon-greedy strategy selects the k-th replica for the next update. An updated $\mathbf{Y}$, calculated from this replica, is then combined with the pre-constructed neighbor graph $\mathbf{G}$ to compute a reward. This reward is further used to update the estimated action value for each replica. As observed, the module initially updates replicas with incorrect k values for exploration. Over time, the optimal $k = k^*$ emerges gradually through reward accumulation, leading to more frequent updates of the $k*$-th replicas, as exemplified by $k = k_i$ in the figure.

structures. Subspace-based MVC: This approach focuses on achieving consistent subspace representation learning across multiple views. For example, (Cao et al., 2015) proposed a diversity-induced mechanism for multi-view subspace clustering. However, the efficacy of the majority of existing methods is significantly dependent on an apriori assumption regarding the true number of clusters, a condition that is seldom attainable in real-world scenarios.

2.3. Bandit Problem

The bandit problem (Bergemann & Valimaki, 2006) represents a fundamental challenge within decision theory and reinforcement learning, depicting a scenario wherein an agent, colloquially termed a "bandit", is confronted with the task of selecting from a multitude of actions or options, each characterized by an uncertain reward or outcome. Within this context, reward hypotheses serve as conjectures concerning the potential rewards associated with each action or option. Specifically, each action is presumed to possess an underlying real reward, which remains obscured from the agent's direct observation throughout the decision-making process. The primary objective of the agent is to optimize long-term rewards by strategically selecting actions. However, given the inherent uncertainty surrounding the rewards, the agent is compelled to estimate the reward values of individual actions through iterative engagement with the environment. We posit that these k classes can be analogized to k bandits. Our investigation reveals that the rewards offered by each class in every iteration correspond with the reward assumptions inherent to the bandit problem. Consequently, leveraging this assumption, we propose a methodology for addressing the clustering problem with an agnostic number of classes. Notably, our approach exhibits high interpretability and can be integrated into any MVC algorithm, as it merely requires original data and cluster indicator matrix to compute the reward.

3. The Proposed Method

In this section, we first outline the notations adopted in this paper for clarity and consistency. We then proceed to detail the formulation of our proposed method.

Notation: For a given multi-view dataset $\mathcal{X} = \{\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(v)}, \ldots, \mathbf{X}^{(V)}\} \in \mathbb{R}^{d_v \times n}$, let $\mathbf{X}^{(v)}$ denote the original data from the v-th view, and d_v, n, V is the feature dimension of $\mathbf{X}^{(v)}$, number of samples, and number of views respectively. $\mathcal{G} = \{\mathbf{G}^{(1)}, \ldots, \mathbf{G}^{(v)}, \ldots, \mathbf{G}^{(V)}\} \in \mathbb{R}^{n \times n}$ denote connection graph of each view constructed by mutual nearest neighbors (MNN). $\mathbf{G} \in \mathbb{R}^{n \times n}$ is consensus graph jointly decided by all views with m edges. $\mathbf{Y}$ denotes the indicator matrix learned by the MVC algorithm, and $\mathbf{y_i}$ denotes i-th column of $\mathbf{Y}$. k denotes number of clusters in descriptive contexts. When referring to mathematical expressions and formulas, we use k to represent the number of clusters.

3.1. Inter-cluster Connectivity Reward

In this subsection, we introduce our reward setting for reinforcement learning. In the context of multi-view clustering tasks, our aim is to derive a clustering indicator matrix $\mathbf{Y}$. This can be achieved through various methods, but generally $\mathbf{Y}$ is calculated from the smallest k eigenvectors of a positive

semi-definite matrix. Besides, we use $\mathbf{X}^{(\mathbf{v})}$ to construct a mutual neighbor graph $\mathbf{G}^{(\mathbf{v})}$, where $\mathbf{G}^{(\mathbf{v})}_{\mathbf{ij}} = 1$ indicates that data point *i* and *j* is mutual neighbor, and vice versa. After that, $\mathbf{G}^{(\mathbf{v})}$ are used to vote for the $\mathbf{G}$ agreed by multiple views. Finally, our reward is calculated by the following equation:

$$\mathbf{R}(k) = -\alpha \sum_{ij}^{n} \mathbf{G}_{ij} \cap (\mathbf{y_i}(k, iter) \neq \mathbf{y_j}(k, iter)), \quad (1)$$

where $\mathbf{G}_{ij} \cap (\mathbf{y_i}(k, iter) \neq \mathbf{y_j}(k, iter))$ denotes the inter-cluster connectivity. The hyperparameter α serves to moderate the unavoidable increase in inter-cluster connectivity with the increase of *k*. Now, our goal is to maximize total rewards and decide which *k*-th replica should be updated as frequently as possible.

3.2. Bandit Clustering

In this subsection, we thoroughly explain how the problem of inferring the number of clusters *k* can be framed as a multi-armed bandit problem and how to address it.

The multi-armed bandit problem can be seen as a simplified version of a reinforcement learning problem, where there are no complex state transitions, only a single state. In this simplified scenario, we can focus on the exploration-exploitation trade-off without the need to consider state transitions and long-term strategies. We will now illustrate the simplicity and appropriateness of our RL module from the perspectives of reward and state.

Reward: Previous methods predominantly concentrate on extracting intrinsic information from the data distribution, such as calculating the within-cluster sum of squares (WSS) as a measure of distortion. To this end, (Liu et al., 2023) proposes a reward mechanism based on the WSS calculated in each iteration. However, the calculation of WSS relies on learning a representation space. Otherwise, the WSS calculated directly from $\mathbf{Y}_k$ would not have a connection with k^*. As discussed, $\mathbf{Y}_k$ can be easily segmented into *k* distinct, well-structured components following the application of *k*-means clustering for any given *k*. Although $\mathbf{Y}_k$ exhibits a clear structure conducive to clustering, this does not guarantee that the clustering results faithfully revert to the structure of the original data. In fact, an incorrect *k* in the learned $\mathbf{Y}_k$ can lead to a distortion of the original data structure, as illustrated in Fig. 1. The reward we have introduced in subsection 3.1 effectively addresses this problem and can be integrated into a simpler reinforcement learning framework. The multi-armed bandit problem requires the assumption of a stationary reward distribution, where the reward distribution for each arm remains stable over time. This stability implies that past reward data can effectively predict future rewards. Our experiments, discussed in subsequent sections, validate that our proposed reward meets this assumption.

State: Previous methods treat all representations of each *k*-th replica throughout the iterative process as states. We believe that such a state space exhibits significant redundancy because the representations encoded under the same *k* should have minimal differences throughout the iterations. Since the reward based on inter-cluster connectivity remains stable throughout the iterations, we can treat the *k* replicas as *k* separate states. As a result, the action value $Q(s_t, k)$ simplifies to $Q(k)$. $Q(k)$ and the ϵ-greedy strategy jointly determine the state transitions. In practice, most MVC algorithms requires fewer iterations to converge compared to the RL module. A lack of exploration could lead to premature convergence on an incorrect *k*-th replica, increasing the risk that the RL module identifies *k* too late. Hence, besides employing the ϵ-greedy strategy, we introduce optimistic initial values to further promote early exploration.

3.3. Discussion

In this subsection, we discuss the detail of proposed reward and showcase the relationship between inter-cluster connectivity and k^*. In a supervised setting, we can categorize the clustering results into True Positives (TP), False Negatives (FN), False Positives (FP) and True Negatives (TN). Let a, b, c and d denote the quantities of TP, FN, FP and TN respectively, where $a + b + c + d = \frac{n(n-1)}{2}$. Ideally, when a clustering algorithm perfectly allocates n samples into k^* clusters, we have $b_{k^*} = 0$ and $c_{k^*} = 0$. Naturally, the clustering results are compromised when $k_i \neq k^*$, leading to $a_{k_i} \leq a_{k^*}$, $b_{k_i} \geq b_{k^*}$, $c_{k_i} \geq c_{k^*}$, and $d_{k_i} \leq d_{k^*}$. Furthermore, it can be concluded that if there is a subset $\mathbf{X}_{sub}$ of $\mathbf{X}$ that contains at least one sample from each cluster, this property will also hold for $\mathbf{X}_{sub}$. Therefore, for a subset $\mathbf{X}_{sub}$ containing m samples, it also holds that $a^{sub}_{k_i} \leq a^{sub}_{k^*}$, $b^{sub}_{k_i} \geq b^{sub}_{k^*}$, $c^{sub}_{k_i} \geq c^{sub}_{k^*}$, and $d^{sub}_{k_i} \leq d^{sub}_{k^*}$, where $a^{sub}_k + b^{sub}_k + c^{sub}_k + d^{sub}_k = m$. Thus, we can infer k^* using any of the following expressions:

$$\begin{aligned} k^* &= \operatorname*{argmax}_k (a^{sub}_k), \\ k^* &= \operatorname*{argmin}_k (b^{sub}_k), \\ k^* &= \operatorname*{argmin}_k (c^{sub}_k), \\ k^* &= \operatorname*{argmax}_k (d^{sub}_k). \end{aligned} \quad (2)$$

While the labels are not revealed during the clustering process, we can estimate them using certain methods. Consider a graph $\mathbf{G}$ with n samples and m edges, where connected edges indicate that the sample pairs belong to the same class

(though unconnected edges do not necessarily imply that they belong to different classes). Using this graph, we can compute $\hat{a}_k^{sub}$, $\hat{b}_{k_i}^{sub}$, $\hat{c}_{k_i}^{sub}$, and $\hat{d}_{k_i}^{sub}$ that satisfy:

$$\begin{aligned} \hat{a}_k^{sub} &= a_k^{sub} + c_k^{sub}, \\ \hat{b}_k^{sub} &= b_k^{sub} + d_k^{sub}, \\ \hat{c}_k^{sub} &= 0, \\ \hat{d}_k^{sub} &= 0. \end{aligned} \tag{3}$$

We analyze the sources of error for the estimators $\hat{a}_k^{sub}$ and $\hat{b}_k^{sub}$: $\hat{a}_k^{sub}$ mistakenly identifies false positive pairs as true positive pairs, while $\hat{b}_k^{sub}$ mistakenly identifies true negative pairs as false negative pairs. To accurately reflect the label information, we need to minimize both errors, i.e., for a subset $\mathbf{X}_{sub}$ containing m samples, $a_k^{sub} \gg c_k^{sub}$ and $b_k^{sub} \gg d_k^{sub}$. In practice, most MVC algorithms construct a similarity graph based on nearest neighbors, in which case $d_k^{sub} \ll c_k^{sub}$. Therefore, by replacing b_k^{sub} with $\hat{b}_k^{sub}$, we have:

$$k^* = \underset{k}{\operatorname{argmin}}(\hat{b}_k^{sub}). \tag{4}$$

Eq. (4) is equivalent to $k^* = \operatorname{argmin}_k(b_k^{sub})$ if and only if $\Delta b^{sub} - \Delta d^{sub} \geq 0$, where $\Delta b^{sub} = b_{k_i}^{sub} - b_{k^*}^{sub}$ and $\Delta d^{sub} = d_{k^*}^{sub} - d_{k_i}^{sub}$ for any $k_i \neq k^*$. Such condition can be satisfied with suitable m.

4. Experiment

In this section, we compare our methodology with several state-of-the-art multi-view clustering methods on benchmark datasets.

4.1. Experiment Setup

Four multi-view datasets are adpoted in our experiment: **MSRC-v1**, **Handwritten-numerals**, **bbc-seg14of4**, and **ORL**:

MSRC-v1 [1] from Microsoft Research in Cambridge contains 240 images and 9 object classes with coarse pixel-wise labeled images.

Handwritten-numerals is an image dataset of handwritten digits, which consists of 2,000 samples from 0 to 9 digit classes. In experiments, we extract six different features to represent each image. We denote it as HW for simplicity.

bbc-seg14of4 [2] is a subset of BBC dataset, which consists of documents from the BBC news website corresponding to stories in five topical areas (business, entertainment, politics, sport, tech). We denote it as BBC14 for simplicity.

Table 1. The clustering results on MSRC-v1 dataset (%)

Method	NMI	Purity	ARI
Co-train	58.69 ± 0.73	71.90 ± 0.67	88.12 ± 0.19
Co-reg	61.20 ± 4.23	72.38 ± 6.06	87.58 ± 2.87
MLAN	45.35 ± 0.00	52.38 ± 0.00	76.52 ± 0.00
AWP	54.88 ± 0.00	63.33 ± 0.00	83.54 ± 0.00
mPAC	63.10 ± 0.00	74.76 ± 0.00	89.46 ± 0.00
GMC	73.90 ± 0.00	79.05 ± 0.00	90.45 ± 0.00
SMVSC	70.18 ± 0.00	81.43 ± 0.00	91.22 ± 0.00
FPMVS	66.84 ± 0.00	78.57 ± 0.00	90.61 ± 0.00
COMVSC	67.01 ± 0.00	79.52 ± 0.00	90.88 ± 0.00
CSMSC	71.43 ± 0.00	80.48 ± 0.00	91.51 ± 0.00
PMSC	34.29 ± 7.18	76.81 ± 2.50	25.34 ± 6.31
CoMSC	73.90 ± 0.43	82.06 ± 0.81	$\mathbf{91.77 \pm 0.27}$
EOMSC	54.94 ± 0.00	67.14 ± 0.00	86.60 ± 0.00
LMVSC	65.26 ± 2.66	75.56 ± 1.62	89.85 ± 0.59
ours	$\mathbf{79.76 \pm 0.00}$	$\mathbf{88.57 \pm 0.00}$	74.51 ± 0.00

ORL (The Olivetti Research Laboratory) [3] face dataset consists of 400 face images in 40 different themes in total. For each subject, the images are described in three features: facial expressions, facial details, and lighting.

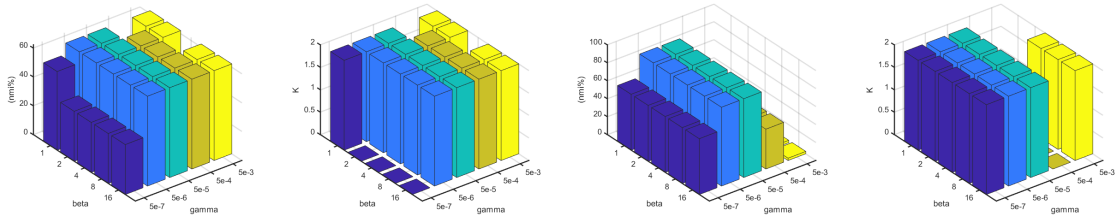
Fourteen competitive multi-view clustering methods are selected for comparison:

Multi-view Spectral Clustering with Co-train strategy (Co-train) (Kumar & Daumé, 2011), Multi-view Spectral clustering with Co-reg strategy (Co-reg) (Kumar et al., 2011), Multi-view clustering and Semi-supervised classification with Adaptive Neighbors (MLAN) (Nie et al., 2017), Multi-view Clustering via Adaptively Weighted Procrustes (AWP) (Nie et al., 2018), Multiple Partitions Aligned Clustering (mPAC) (Kang et al., 2019), Graph-based Multi-view Clustering(GMC) (Wang et al., 2019), Scalable Multi-view Subspace Clustering(SMVSC) (Sun et al., 2021), Fast Parameter-free Multi-view Subspace Clustering (FPMVS) (Wang et al., 2021), Consensus One-step Multi-view Subspace Clustering (COMVSC) (Zhang et al., 2020), Consistent and Specific Multi-view Subspace Clustering (CSMSC) (Luo et al., 2018), Consensus One-step Multi-view Subspace Clustering (CoMSC) (Zhang et al., 2020), Efficient One-pass Multi-view Subspace Clustering (EOMSC) (Liu et al., 2022), and Large-scale Multi-view Subspace Clustering (LMVSC) (Kang et al., 2020b).

[1] https://www.microsoft.com/en-us/research/project/image-understanding/

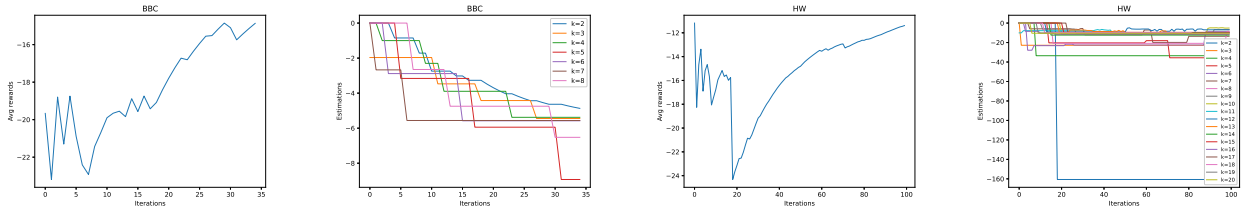
[2] http://mlg.ucd.ie/datasets/bbc.html

[3] https://www.kaggle.com/datasets/tavarez/the-orl-database-for-training-and-testing



(a) $\alpha = 5$, NMI (b) $\alpha = 5, k$ (c) $\alpha = 10$, NMI (d) $\alpha = 5, k$

Figure 3. Sensitivity analysis of hyperparameters α, β, γ.



(a) average reward, BBC14 (b) q estimation, BBC14 (c) average reward, HW (d) q estimation, HW

Figure 4. Convergence curves.

Table 2. The clustering results on HW dataset (%)

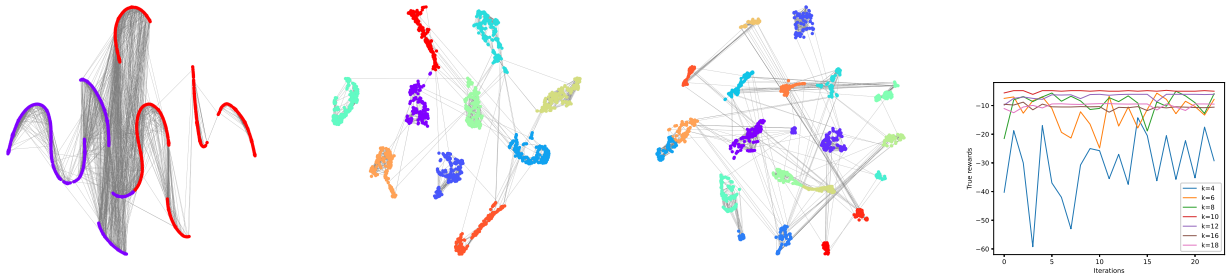
Method	NMI	Purity	ARI
Co-train	69.72 ± 0.85	77.82 ± 0.09	93.33 ± 0.17
Co-reg	62.54 ± 1.44	68.00 ± 2.55	91.21 ± 0.44
MLAN	90.40 ± 0.00	95.60 ± 0.00	$\mathbf{98.29 \pm 0.00}$
AWP	73.94 ± 0.00	74.85 ± 0.00	93.74 ± 0.00
mPAC	60.30 ± 0.00	61.60 ± 0.00	90.48 ± 0.00
GMC	90.24 ± 0.00	88.60 ± 0.00	97.35 ± 0.00
SMVSC	77.80 ± 0.00	84.80 ± 0.00	94.77 ± 0.00
FPMVS	78.14 ± 0.00	82.25 ± 0.00	94.83 ± 0.00
COMVSC	84.72 ± 0.00	92.00 ± 0.00	97.01 ± 0.00
CSMSC	84.15 ± 0.00	90.95 ± 0.00	96.67 ± 0.00
PMSC	61.34 ± 3.94	86.31 ± 1.29	48.74 ± 5.43
CoMSC	85.69 ± 2.74	90.95 ± 2.41	96.78 ± 0.76
EOMSC	77.89 ± 0.00	76.20 ± 0.00	93.69 ± 0.00
LMVSC	80.51 ± 0.17	85.42 ± 0.37	95.50 ± 0.08
ours	$\mathbf{94.38 \pm 0.00}$	$\mathbf{97.50 \pm 0.00}$	94.46 ± 0.00

Table 3. The clustering results on BBC14 dataset (%)

Method	NMI	Purity	ARI
Co-train	66.92 ± 4.75	82.76 ± 4.88	85.38 ± 3.34
Co-reg	73.68 ± 0.00	89.66 ± 0.00	89.51 ± 0.00
MLAN	37.85 ± 2.53	62.64 ± 1.63	62.24 ± 1.92
AWP	43.13 ± 0.00	66.38 ± 0.00	69.78 ± 0.00
mPAC	56.45 ± 0.00	79.31 ± 0.00	84.00 ± 0.00
GMC	53.10 ± 0.00	68.10 ± 0.00	70.75 ± 0.00
SMVSC	14.12 ± 0.00	46.55 ± 0.00	68.13 ± 0.00
FPMVS	8.64 ± 0.00	40.52 ± 0.00	64.69 ± 0.00
COMVSC	43.90 ± 0.00	59.48 ± 0.00	71.75 ± 0.00
CSMSC	68.65 ± 0.00	76.72 ± 0.00	82.83 ± 0.00
PMSC	44.70 ± 0.00	96.55 ± 0.00	66.22 ± 0.00
CoMSC	74.08 ± 0.00	81.90 ± 0.00	89.51 ± 0.00
EOMSC	23.81 ± 0.00	53.45 ± 0.00	69.91 ± 0.00
LMVSC	32.42 ± 5.76	67.24 ± 2.11	64.28 ± 5.49
ours	$\mathbf{87.72 \pm 0.00}$	$\mathbf{99.31 \pm 0.00}$	$\mathbf{94.11 \pm 0.00}$

4.2. Results Analysis

Three criteria of clustering performance (Normalized Mutual Information (NMI), Purity, and Adjusted Rand Index (ARI)) are shown in Table 1–4. The comparison algorithms were repeatedly tested 10 times using the parameter settings recommended by corresponding papers. In most cases, our method consistently outperforms others, offering strong evidence of its effectiveness. In Table 3, on BBC14 dataset, our method outperforms other methods by about $13.64\%, 2.76\%, 4.60\%$ in terms of NMI, Purity, and ARI. On other datasets, our method also achieves comparable results. It should be noted that our method is non-parametric, capable of automatically inferring k during the algorithm's iterative process, in contrast to all the aforementioned approaches that require k to be input in advance. The reason our method achieves comparable results is that it correctly inferred k across all four datasets due to our multi-armed bandit hypothesis and meticulously designed rewards. Ad-



(a) $k = 2$ (b) $k = 10$ (c) $k = 19$ (d) true reward

Figure 5. t-SNE visualization of known parametric MVC on HW dataset with different k.

Table 4. The clustering results on ORL dataset (%)

Method	NMI	Purity	ARI
Co-train	79.94 ± 0.74	66.58 ± 0.59	97.69 ± 0.05
Co-reg	79.77 ± 0.17	66.50 ± 1.41	97.71 ± 0.04
MLAN	52.96 ± 0.53	40.83 ± 0.47	88.50 ± 0.37
AWP	89.45 ± 0.00	80.75 ± 0.00	98.60 ± 0.00
mPAC	83.68 ± 0.00	68.75 ± 0.00	97.94 ± 0.00
GMC	81.66 ± 0.00	72.50 ± 0.00	94.60 ± 0.00
SMVSC	76.23 ± 0.00	62.00 ± 0.00	96.94 ± 0.00
FPMVS	73.59 ± 0.00	59.00 ± 0.00	96.64 ± 0.00
COMVSC	88.34 ± 0.00	81.75 ± 0.00	98.45 ± 0.00
CSMSC	89.89 ± 0.00	81.00 ± 0.00	$\mathbf{98.71 \pm 0.00}$
PMSC	71.74 ± 2.28	70.60 ± 3.49	34.75 ± 3.55
CoMSC	86.06 ± 0.40	79.25 ± 0.94	98.18 ± 0.07
EOMSC	81.70 ± 0.00	$\mathbf{92.50 \pm 0.00}$	98.62 ± 0.00
LMVSC	78.84 ± 0.75	71.67 ± 3.60	96.84 ± 0.43
ours	$\mathbf{90.39 \pm 0.00}$	82.25 ± 0.00	69.18 ± 0.00

ditionally, the module we propose for inferring agnostic k is capable of being applied in any scenario as aforementioned. We believe that integrating our proposed paradigm with more advanced MVC algorithms could further enhance its performance.

4.3. Sensitivity Analysis

Figs. 3(a) and 3(c) present the Normalized Mutual Information (NMI) performance metrics of the clustering algorithm under various parameter settings. Figs. 3(b) and 3(d) report whether the algorithm identified the optimal k, with any incorrect k values denoted as $k = 0$. To study the influence of different parameter settings on the clustering results, we vary α, β, and γ in the ranges $[5, 10]$, $[0, 16]$, and $\left[5e^{-7}, 5e^{-3}\right]$. Noted that changes in MVC hyperparameters can affect the quality of the clustering indicator matrix $\mathbf{Y}$, which is subsequently used to infer k. Nevertheless, it can be observed that our method successfully identifies the optimal k in 37 out of 50 parameter settings, demonstrating the robustness of proposed RL module.

4.4. Convergence Study

We demonstrate the convergence of the RL module by showcasing the average reward curve and the value estimation for each k-th replica. In Figs. 4(b) and 4(d), we observe that our reinforcement learning paradigm estimates the highest value for the correct k. In Figs. 4(a) and 4(c), the obtained average rewards are consistently increasing. Overall, the RL module effectively fulfills its role in accurately inferring k. Fluctuations in the learning process are caused by two factors. First, the exploration strategy based on optimistic initial values leads to frequent exploration in the early stages, and second, the ϵ-greedy strategy ensures ongoing exploration throughout the learning process. The former results in mysterious spikes in the early average reward curve, while the latter causes the curve to exhibit small oscillations continuously. In practice, most MVC algorithms requires fewer iterations to converge compared to the RL module. Hence, we emphasize the importance of incorporating early exploration. Additionally, the curve lacks a stable trend in later stages because the clustering process doesn't wait for the RL-learned rewards to fully converge before iteratively updating the optimal k-th replica.

4.5. Reward Analysis

In this subsection, we remove the RL module for inferring agnostic k and explicitly input varying k values on the HW dataset to observe the clustering result and the inter-cluster connectivity after convergence. From the t-SNE (Van der Maaten & Hinton, 2008) visualization, as shown in Figs. 5(a), 5(b) and 5(c), it is evident that $k = 10$ exhibits significantly less inter-cluster connectivity compared to the incorrect k values of $k = 2$ and $k = 19$. Both $k = 2$ and $k = 19$, served as extreme cases of an erroneous number of k, yield clear, discriminative representations, which corroborates our hypothesis and clarifies the reason why WSS is inappropriate for this scenario. In addition, we inde-

pendently calculate the actual reward for each k during the iteration process by Eq. (1) across a broader range of k values. As shown in Fig. 5(d), the designed reward is effective, with the reward at each iteration being essentially maximal when $k^* = 10$. As mentioned above, the expected reward remains relatively stable throughout the iteration process, which provides theoretical assurance for simplifying this reinforcement learning task.

5. Conclusion

In conclusion, this paper addresses the challenge of determining the optimal cluster number, k, in multi-view clustering, a critical issue in real-world scenarios where prior information about k is often unavailable. We introduce a novel graph-based multi-view clustering algorithm that leverages a graph consistency reward mechanism to infer k autonomously. By evaluating the cluster indicator matrix iteratively across varying k values and formulating the inference process as a reinforcement learning paradigm, our approach emphasizes inter-cluster connectivity as the reward signal, thereby achieving robust clustering results independent of predefined cluster numbers. Experimental evaluations on diverse benchmark datasets substantiate the superiority of our method over existing state-of-the-art approaches, underscoring its effectiveness in addressing the inherent challenges of multi-view clustering without relying on strong prior knowledge.

Acknowledgement

This work was partially supported by the National Science Foundation of China under Grant 62106164 and 62376175, the 111 Project under Grant B21044, and the Sichuan Science and Technology Program under Grants 2021ZDZX0011.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Bergemann, D. and Valimaki, J. Bandit problems. 2006.

Bickel, S. and Scheffer, T. Multi-view clustering. In *ICDM*, volume 4, pp. 19–26. Citeseer, 2004.

Cai, H., Tan, Y., Huang, S., and Lv, J. Lifelong multi-view spectral clustering. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pp. 3488–3496, 2023.

Cai, X., Nie, F., and Huang, H. Multi-view k-means clustering on big data. In *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.

Cao, X., Zhang, C., Fu, H., Liu, S., and Zhang, H. Diversity-induced multi-view subspace clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586–594, 2015.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD*, volume 96, pp. 226–231, 1996.

Huang, S., Kang, Z., Tsang, I. W., and Xu, Z. Auto-weighted multi-view clustering via kernelized graph learning. *Pattern Recognition*, 88:174–184, 2019.

Huang, S., Kang, Z., and Xu, Z. Auto-weighted multi-view clustering via deep matrix decomposition. *Pattern Recognition*, 97:107015, 2020.

Huang, S., Tsang, I. W., Xu, Z., and Lv, J. Measuring diversity in graph learning: A unified framework for structured multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 34(12):5869–5883, 2021.

Huang, S., Liu, Y., Cai, H., Tan, Y., Tang, C., and Lv, J. Smooth representation learning from multi-view data. *Information Fusion*, 100:101916, 2023.

Kang, Z., Guo, Z., Huang, S., Wang, S., Chen, W., Su, Y., and Xu, Z. Multiple partitions aligned clustering. In *IJCAI*, pp. 2701–2707. ijcai.org, 2019.

Kang, Z., Zhao, X., Peng, C., Zhu, H., Zhou, J. T., Peng, X., Chen, W., and Xu, Z. Partition level multiview subspace clustering. *Neural Networks*, 122:279–288, 2020a.

Kang, Z., Zhou, W., Zhao, Z., Shao, J., Han, M., and Xu, Z. Large-scale multi-view subspace clustering in linear time. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 4412–4419, 2020b.

Kodinariya, T. M., Makwana, P. R., et al. Review on determining number of cluster in k-means clustering. *International Journal*, 1(6):90–95, 2013.

Kumar, A. and Daumé, H. A co-training approach for multi-view spectral clustering. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pp. 393–400, 2011.

Kumar, A., Rai, P., and Daume, H. Co-regularized multi-view spectral clustering. *Advances in Neural Information Processing Systems*, 24, 2011.

Langley, P. Crafting papers on machine learning. In *ICML*, pp. 1207–1216, 2000.

Li, Z., Tang, C., Liu, X., Zheng, X., Zhang, W., and Zhu, E. Consensus graph learning for multi-view clustering. *IEEE Transactions on Multimedia*, 24:2461–2472, 2021.

Liu, J., Wang, C., Gao, J., and Han, J. Multi-view clustering via joint nonnegative matrix factorization. In *Proceedings of the 2013 SIAM International Conference on Data Mining*, pp. 252–260. SIAM, 2013.

Liu, J., Liu, X., Yang, Y., Guo, X., Kloft, M., and He, L. Multiview subspace clustering via co-training robust data representation. *IEEE Transactions on Neural Networks and Learning Systems*, 33(10):5177–5189, 2021.

Liu, S., Wang, S., Zhang, P., Xu, K., Liu, X., Zhang, C., and Gao, F. Efficient one-pass multi-view subspace clustering with consensus anchors. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 7576–7584, 2022.

Liu, Y., Liang, K., Xia, J., Yang, X., Zhou, S., Liu, M., Liu, X., and Li, S. Z. Reinforcement graph clustering with unknown cluster number. In *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 3528–3537, 2023.

Luo, S., Zhang, C., Zhang, W., and Cao, X. Consistent and specific multi-view subspace clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Nie, F., Cai, G., and Li, X. Multi-view clustering and semi-supervised classification with adaptive neighbours. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

Nie, F., Tian, L., and Li, X. Multiview clustering via adaptively weighted procrustes. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2022–2030, 2018.

Onumanyi, A. J., Molokomme, D. N., Isaac, S. J., and Abu-Mahfouz, A. M. Autoelbow: An automatic elbow detection method for estimating the number of clusters in a dataset. *Applied Sciences*, 12(15):7515, 2022.

Peng, X., Huang, Z., Lv, J., Zhu, H., and Zhou, J. T. Comic: Multi-view clustering without parameter selection. In *International Conference on Machine Learning*, pp. 5092–5101. PMLR, 2019.

Rodriguez, A. and Laio, A. Clustering by fast search and find of density peaks. *Science*, 344(6191):1492–1496, 2014.

Ronen, M., Finder, S. E., and Freifeld, O. Deepdpm: Deep clustering with an unknown number of clusters. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9861–9870, 2022.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.

Sun, M., Zhang, P., Wang, S., Zhou, S., Tu, W., Liu, X., Zhu, E., and Wang, C. Scalable multi-view subspace clustering with unified anchors. In *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 3528–3536, 2021.

Tan, Y., Liu, Y., Huang, S., Feng, W., and Lv, J. Sample-level multi-view graph clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 23966–23975, 2023.

Tang, C., Liu, X., Zhu, X., Zhu, E., Luo, Z., Wang, L., and Gao, W. Cgd: Multi-view clustering via cross-view graph diffusion. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 5924–5931, 2020.

Tian, Y., Krishnan, D., and Isola, P. Contrastive multiview coding. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XI 16*, pp. 776–794. Springer, 2020.

Tibshirani, R., Walther, G., and Hastie, T. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423, 2001.

Umargono, E., Suseno, J. E., and Gunawan, S. V. K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula. In *The 2nd International Seminar on Science and Technology (ISSTEC 2019)*, pp. 121–129. Atlantis Press, 2020.

Van der Maaten, L. and Hinton, G. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(11), 2008.

Wan, X., Liu, J., Liang, W., Liu, X., Wen, Y., and Zhu, E. Continual multi-view clustering. In *Proceedings of the 30th ACM International Conference on Multimedia*, pp. 3676–3684, 2022.

Wang, C.-D., Lai, J.-H., and Philip, S. Y. Multi-view clustering based on belief propagation. *IEEE Transactions on Knowledge and Data Engineering*, 28(4):1007–1021, 2015.

Wang, H., Yang, Y., and Liu, B. Gmc: Graph-based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 32(6):1116–1129, 2019.

Wang, S., Liu, X., Zhu, X., Zhang, P., Zhang, Y., Gao, F., and Zhu, E. Fast parameter-free multi-view subspace

clustering with consensus anchor guidance. *IEEE Transactions on Image Processing*, 31:556–568, 2021.

Yuan, C. and Yang, H. Research on k-value selection method of k-means clustering algorithm. *J*, 2(2):226–235, 2019.

Zhang, P., Liu, X., Xiong, J., Zhou, S., Zhao, W., Zhu, E., and Cai, Z. Consensus one-step multi-view subspace clustering. *IEEE Transactions on Knowledge and Data Engineering*, 34(10):4676–4689, 2020.

Zhao, H., Ding, Z., and Fu, Y. Multi-view clustering via deep matrix factorization. In *Proceedings of the AAAI Conference on Artificial intelligence*, volume 31, 2017.

Zhao, H., Yang, X., and Deng, C. Parameter-agnostic deep graph clustering. *ACM Transactions on Knowledge Discovery from Data*, 18(3):1–20, 2024.