

TPCH: Tensor-interacted Projection and Cooperative Hashing for Multi-view Clustering

Zhongwen Wang¹, Xingfeng Li^{1*}, Yinghui Sun^{2†}, Quansen Sun¹,
Yuan Sun³, Han Ling¹, Jian Dai⁴, Zhenwen Ren⁵

¹School of Computer Science and Engineering, Nanjing University of Science and Technology

²School of Computer Science and Engineering, Southeast University

³College of Computer Science, Sichuan University

⁴Southwest Automation Research Institute, China South Industries Group Corporation

⁵School of National Defence Science and Technology, Southwest University of Science and Technology

jankinwang@njust.edu.cn, lixingfeng@njust.edu.cn, sunyh@seu.edu.cn, sunquansen@njust.edu.cn, sunyuan_work@163.com, 321106010190@njust.edu.cn, daijian1000@163.com, rzw@njust.edu.cn

Abstract

In recent years, anchor and hash-based multi-view clustering methods have gained attention for their efficiency and simplicity in handling large-scale data. However, existing methods often overlook the interactions among multi-view data and higher-order cooperative relationships during projection, negatively impacting the quality of hash representation in low-dimensional spaces, clustering performance, and sensitivity to noise. To address this issue, we propose a novel approach named Tensor-Interacted Projection and Cooperative Hashing for Multi-View Clustering (TPCH). TPCH stacks multiple projection matrices into a tensor, taking into account the synergies and communications during the projection process. By capturing higher-order multi-view information through dual projection and Hamming space, TPCH employs an enhanced tensor nuclear norm to learn more compact and distinguishable hash representations, promoting communication within and between views. Experimental results demonstrate that this refined method significantly outperforms state-of-the-art methods in clustering on five large-scale multi-view datasets. Moreover, in terms of CPU time, TPCH achieves substantial acceleration compared to the most advanced current methods.

Code — <https://github.com/jankin-wang/TPCH>

Introduction

In recent years, advanced information-gathering technologies allow us to obtain multi-view data from the same object, enabling a more in-depth and comprehensive description of the object (Lu et al. 2023a; Qin, Pu, and Wu 2024; Sun et al. 2023b; Wu et al. 2023; Chen et al. 2023; Qin et al. 2024; Sun et al. 2023a, 2024). For example, in the field of action recognition, different views such as RGB, optical flow, and skeletal data can be used to describe actions (Hu et al. 2023; Wang et al. 2022b). With the widespread application of multi-view data, a series of multi-view learning tasks have emerged,

such as multi-view classification (Liu et al. 2024a; Han et al. 2022), multi-view feature aggregation (Hu et al. 2024b; Yang et al. 2020; Hu et al. 2024a), incomplete multi-view clustering (Wen et al. 2023a,b) and so on. Among them, Multi-View Clustering (MVC) is an unsupervised method in multi-view learning (Li et al. 2023, 2024c,b). Compared to supervised methods, multi-view clustering does not require data to be labeled in advance, making it more efficient in explosive data growth. In practical applications, multi-view clustering has been widely used in computer vision, natural language processing, medical image analysis, and other fields (Fang et al. 2023; Wen et al. 2023c; Zhang et al. 2021). In the context of large-scale data, multi-view clustering faces challenges in computation and storage, known as the problem of Large-scale Multi-view Clustering (IMVC).

To solve the problem of large-scale multi-view clustering, early attempts involve reducing the number of matrix multiplications by seeking approximations of multi-view data matrices, thus reducing computational overhead (Wang et al. 2011; Zhang and Lu 2016). In recent years, the method of selecting or generating a representative set of data from multi-view data, known as "anchors," has been widely adopted (Nie et al. 2021; Liu et al. 2024b; Wang et al. 2022a; Zhang, Nie, and Li 2023). This approach significantly reduces computational complexity and memory requirements by selecting a small subset of anchor points to represent the entire dataset. Moreover, anchor-based methods establish bridges between different views through these anchor points, enhancing the exploration of consistent information between views. In large-scale clustering tasks, traditional multi-view learning methods that operate directly on the raw data incur high computational and storage costs. The emergence of Binary Multi-View Clustering (BMVC) (Zhang et al. 2018b) methods has alleviated this issue. BMVC maps data to a low-dimensional binary Hamming space and performs clustering analysis in this lower-dimensional space, accelerating computational and storage efficiency in handling large-scale multi-view data. Furthermore, existing hash-based methods typically project multi-view data from the original space to a shared Hamming representation space to enhance the ex-

*Corresponding authors.

†Corresponding authors.

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

pressive power of the Hamming representation space.

However, existing BMVC methods typically utilize multiple independent projection matrices to project multi-view data into low-dimensional space and then learn hash codes for clustering partitioning in the low-dimensional space. The independence of projection matrices leads to insufficient communication between multi-view data and a lack of high-order collaborative relationships, hindering the learning of compact hash representations in low-dimensional space and reducing clustering performance.

To solve the problems above, we proposed a novel method called Tensor-Interacted Projection and Cooperative Hashing for Multi-View Clustering (TPCH). Specifically, multiple independent projection matrices project nonlinear multi-view data from high-dimensional space to multiple low-dimensional spaces, resulting in multiple distinctive hash representations. To enhance the high-order collaboration of projection matrices during the projection process, we first superimpose multiple projection matrices into a high-order tensor. Simultaneously, we utilize t-SVD decomposition-based enhanced tensor low-rank constraints to enhance the compactness of hash representations. In the Hamming space, multiple hash representations are stacked into high-order tensors and enhanced with high-order tensor low-rank constraints to strengthen the collaborative capabilities within and between views, further improving the compactness of hash representations across multiple views. Finally, the average of multiple compact hash representations is taken to obtain clustering results. Our work has the following main contributions:

- For the first time, the higher-order interactions between projection matrices are considered, and a tensor-interacted projection and cooperative hashing framework is constructed to capture the high-order multi-view information of dual projections and Hamming space.
- An enhanced tensor nuclear norm on the core tensor is employed to improve the compactness and distinguishability of hash codes, reducing the impact of redundant information and noise to enhance robustness.
- Experimental results on five common large-scale multi-view datasets demonstrate that TPCH significantly improves clustering performance under five mainstream clustering evaluation metrics compared to current state-of-the-art methods.

Related Work

Anchor-based Large-scale Multi-view Clustering

In the process of multi-view clustering, each view provides a unique feature description for objects (Lu et al. 2023b; Lu, Liu, and Wang 2020). Anchor-based multi-view clustering methods aim to achieve more precise and robust clustering results by facilitating the exchange between views through anchors. Selecting or generating a small number of anchors from the entire dataset can significantly reduce computational complexity and memory usage, which is particularly effective when dealing with large-scale datasets. (Kang et al. 2020) learns a smaller bipartite graph for each view and

then merges these bipartite graphs to obtain a unified representation. (Wang et al. 2022a; Li et al. 2024a) reduce complexity by learning consistent anchor points. However, this approach assumes that anchor points are shared among all views, emphasizing consistency between views while overlooking complementarity. To address this, (Liu et al. 2024b) uses view-independent anchor points to represent data. This method allows different anchor points to be chosen for different views, providing high flexibility to adapt to the characteristics and requirements of different views. In recent years, efforts have been made to find a balance between consistency and complementarity among views, leading to the emergence of tensor anchor-based methods (Long et al. 2023; Xia et al. 2023; Chen, Wang, and Lai 2022). These methods stack each bipartite graph as a tensor and participate as a whole in subsequent clustering, achieving the effect of seeking consistency between views while retaining complementarity among views.

Binary Multi-view Clustering

Binary multi-view clustering methods analyze data by mapping it into a low-dimensional binary space, significantly reducing the dimensionality of the data, alleviating the curse of dimensionality, and improving the efficiency of clustering and subsequent processing. The low-dimensional binary hash codes enable faster similarity calculations between data points, especially when using Hamming distance, greatly enhancing clustering computation speed. Reduced-dimensional data occupies less storage space, making binary hash-based multi-view clustering methods more suitable for large-scale datasets. Binary-based clustering methods have been around since the early days (Gong et al. 2015). However, these methods separate binary representation and clustering into two separate steps, requiring customized representation methods for different datasets, thus limiting the algorithm’s generalization ability. To address this issue, in recent years, researchers have adopted a joint learning approach to simultaneously learn multi-view binary representations and clustering structures, achieving good clustering results with low computational and storage costs (Shen et al. 2017; Zhao et al. 2023). To obtain better binary representations, (Zhang et al. 2018b) places images into a compact common binary space and uses a binary matrix decomposition model for clustering binary representations. (Zhang et al. 2024) employs novel tensor low-rank constraints to stack binary representations from various views into tensors to capture high-order interactions.

Methodology

To derive bipartite graphs from raw multi-view data, non-linear function methods effectively capture the inherent structure and semantic relationships within the data. These methods bring similar data closer together in high-dimensional space while pushing dissimilar data farther apart. Inspired by (Zhang et al. 2024), we employ a non-linear Radial Basis Function (RBF) kernel mapping to construct bipartite graphs for each view. The corresponding for-

mula is

$$\phi(X_p) = \left[\exp\left(-\frac{\|X_p^1 - S_p^1\|^2}{\delta}\right), \dots, \exp\left(-\frac{\|X_p^n - S_p^m\|^2}{\delta}\right) \right]^T \quad (1)$$

where $\phi(\cdot)$ denotes a mapping used to obtain the bipartite graph for each view. δ represents the width of the kernel. We randomly select m anchor point samples S_p from the p -th view to ensure that the original data structure remains intact after nonlinear mapping.

Problem Formulation

Based on the kernelized bipartite graphs, we propose to explore the high-order information in dual projection and Hamming space, facilitating intra-view and inter-view communication and synergy of the projection matrices and hash codes. In this way, we could learn more compact and distinguishable hash codes. Concretely, the proposed **Tensor-interacted Projection and Cooperative Hashing for Multi-view Clustering (TPCH)** can be implemented as

$$\begin{aligned} \min_{\mathbf{B}_p, \mathbf{Q}_p, \mathbf{Q}, \mathbf{B}} \quad & \alpha \sum_{p=1}^v \|\mathbf{Q}_p^\top \phi(\mathbf{X}_p) - \mathbf{B}_p\|_F^2 + (\|\mathbf{Q}\|_{tnn} + \|\mathbf{B}\|_{tnn}) \\ \text{s.t.} \quad & \mathbf{B} = \Phi([\mathbf{B}_1; \dots; \mathbf{B}_v]), \mathbf{Q} = \Phi([\mathbf{Q}_1; \dots; \mathbf{Q}_v]), \\ & \mathbf{B} \in \{-1, 1\}^{l \times n} \end{aligned} \quad (2)$$

where kernelized bipartite graph $\phi(\mathbf{X}_p) \in \mathbf{R}^{m \times n}$ is the input data of the p -th view. m and n are the number of anchors and the number of samples. α is a hyperparameter used to adjust the proportion between the binary representation of multi-view data and the dual tensor. $\mathbf{Q}_p \in \mathbf{R}^{m \times l}$ is the projection matrix of the p -th view, used to project the bipartite graph into a consistent hamming space of dimensionality l . These projection matrices are stacked to form a tensor \mathbf{Q} . $\mathbf{B}_p \in \mathbf{R}^{l \times n}$ is the embedding feature matrix/hash matrix learned for the p -th view, and these embedding feature matrices are stacked to form a tensor \mathbf{B} . $\|\cdot\|_{tnn}$ denotes the Tensor Nuclear Norm (TNN), which could be solved by tensor Singular Value Decomposition (t -SVD) as mentioned in Fig. 1. In this way, TPCH captures the high-order semantic information among bipartite graphs and then embeds them into hash codes $\mathbf{B}_p \in \{-1, 1\}$ to enhance expression ability.

From Eq. 2, TPCH could stack projection matrices and hash matrices to construct tensors \mathbf{Q} and \mathbf{B} to capture higher-order information of multi-view data in both dual projection and Hamming spaces. However, the tensors \mathbf{Q} and \mathbf{B} not only contain the desired high-order information but also incorporate noise or redundant information. As showcased in Fig. 1, the core tensor \mathcal{S} from the t -SVD decomposition may not be low-rank, containing harmful noise and redundant information. Moreover, the core tensor with a high rank tremendously limits the higher-order synergistic between the projection matrices and hash codes. To reduce the impact of noise and redundant information, we develop the enhanced tensor nuclear norm by enforcing low-rank property on the core tensor as shown in Fig. 2.

By further ranking the core tensor in tensor Singular Value Decomposition (t -SVD) of the TNN, we further propose the

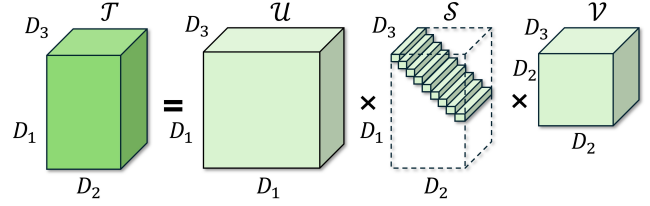


Figure 1: Basic process of t -SVD. $\mathcal{T} \in \mathbf{R}^{D_1 \times D_2 \times D_3}$ represents the tensor to be decomposed, $\mathcal{S} \in \mathbf{R}^{D_1 \times D_2 \times D_3}$ denotes the core tensor, $\mathcal{U} \in \mathbf{R}^{D_1 \times D_1 \times D_3}$, and $\mathcal{V} \in \mathbf{R}^{D_3 \times D_2 \times D_2}$ represent the left and right tensors, respectively, resulting from the t -SVD decomposition of \mathcal{T} .

Enhanced Tensor Nuclear Norm (ETNN) as

$$\|\mathcal{T}\|_{etnn} = \|\bar{\mathcal{S}}\|_* + \zeta \left\| \mathcal{U} * \mathcal{B}^{-1} \left(\bar{\mathcal{S}} \right) * \mathcal{V} \right\|_* \quad (3)$$

where $\bar{\mathcal{S}} \in \mathbf{R}^{D \times D_3}$ is the low-rank approximation of the core tensor, \mathcal{B} and \mathcal{B}^{-1} are transformations between \mathcal{S} and \mathcal{S}^{-1} . ζ is a predefined parameter, $\|\cdot\|_*$ denotes the nuclear norm of the unfolded tensor. Fig. 2 illustrates the transformations between core tensors. Here, $D = \min(D_1, D_2)$, and \mathcal{B} can be defined using matrix multiplication. By incorporating Fig. 2 into Fig. 1, the proposed ETNN could first alleviate the noise and redundant information hidden in the core tensor. Further, TPCH with ETNN ranks the core tensor in dual projection and Hamming spaces through transformations \mathcal{B} and \mathcal{B}^{-1} , enhancing the robustness in the projection synergy and inter/intra-view communication. In this way, TPCH could generate more compact, distinguishable, and robust hash codes to improve the clustering performance.

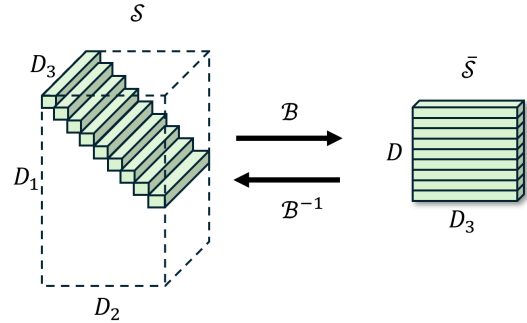


Figure 2: Transformation instructions for core tensors.

Optimization

For solving Eq. (2) with ETNN, an alternating direction minimizing strategy is proposed to iteratively optimize our CHBG. Two auxiliary variables \mathcal{A} and \mathcal{E} are employed to

get augmented Lagrangian function of Eq. (2) as

$$\begin{aligned} & \min_{\mathbf{B}_p, \mathbf{Q}_p, \mathbf{Q}, \mathcal{A}, \mathcal{B}, \mathcal{E}} \alpha \sum_{p=1}^v \|\mathbf{Q}_p^\top \phi(\mathbf{X}_p) - \mathbf{B}_p\|_F^2 + \frac{\mu}{2} \|\mathbf{Q} - \mathcal{A} + \frac{\mathcal{Y}}{\mu}\|_F^2 \\ & + \|\mathcal{A}\|_{etnn} + \|\mathcal{E}\|_{etnn} + \frac{\mu}{2} \|\mathcal{B} - \mathcal{E} + \frac{\mathcal{J}}{\mu}\|_F^2 \\ \text{s.t. } & \mathbf{B} \in \{-1, 1\}^{l \times n}, \mathcal{A} = \Phi([\mathbf{A}_1; \dots; \mathbf{A}_v]), \mathbf{Q} = \mathcal{A}, \\ & \mathcal{B} = \mathcal{E}, \mathcal{B} = \Phi([\mathbf{B}_1; \dots; \mathbf{B}_v]), \mathbf{Q} = \Phi([\mathbf{Q}_1; \dots; \mathbf{Q}_v]) \end{aligned} \quad (4)$$

In this way, all the variables can be iteratively optimized one by one as follows.

1- Updating the subproblem of Q: Fixing the other variables of subproblem Q, problem (4) *w.r.t.* Q simplifies as

$$\alpha \sum_{p=1}^v \|\mathbf{Q}_p^\top \phi(\mathbf{X}_p) - \mathbf{B}_p\|_F^2 + \frac{\mu}{2} \|\mathbf{Q}_p - \mathbf{A}_p + \frac{\mathbf{Y}_p}{\mu}\|_F^2 \quad (5)$$

where $\mathbf{Q}_p = [2\alpha\mathbf{B}_p\phi^\top(\mathbf{X}_p) + \mu(\mathbf{A}_p - \frac{\mathbf{Y}_p}{\mu})] / [2\alpha\phi(\mathbf{X}_p)\phi^\top(\mathbf{X}_p) + \mu\mathbf{I}]$. Q subproblem requires $\mathcal{O}(m^3l + nml + nm^2)$ complexity.

2- Updating the subproblem of B: Fixing the other variables of subproblem B, problem (6) *w.r.t.* B simplifies as

$$\begin{aligned} & \min_{\mathbf{B}_p} \alpha \sum_{p=1}^v \|\mathbf{Q}_p^\top \phi(\mathbf{X}_p) - \mathbf{B}_p\|_F^2 + \frac{\mu}{2} \|\mathbf{B}_p - \mathbf{E}_p + \frac{\mathbf{J}_p}{\mu}\|_F^2 \\ \text{s.t. } & \mathbf{B}_p \in \{-1, 1\}^{l \times n} \end{aligned} \quad (6)$$

where 'con' is the constant part of \mathbf{B}_p . The constant $\text{tr}(\mathbf{B}_p^\top \mathbf{B}_p) = \text{tr}(\mathbf{B}_p \mathbf{B}_p^\top) = nl$ could make problem (6) change to

$$\begin{aligned} & \max_{\mathbf{B}_p} \sum_{p=1}^v \text{tr}[\mathbf{B}_p^\top (\alpha \mathbf{Q}_p \phi(\mathbf{X}_p) + \frac{\mu}{2} (\mathbf{E}_p - \frac{\mathbf{J}_p}{\mu}))] \\ \text{s.t. } & \mathbf{B}_p \in \{-1, 1\}^{l \times n} \end{aligned} \quad (7)$$

which has a closed-form optimal solution:

$$\mathbf{B}_p = \text{sgn}(\alpha \mathbf{Q}_p \phi(\mathbf{X}_p) + \frac{\mu}{2} (\mathbf{E}_p - \frac{\mathbf{J}_p}{\mu})). \quad (8)$$

B subproblem requires $\mathcal{O}(mln + ncl)$ time complexity.

Update-3: Solving A can be written as

$$(9)$$

which is solved by following steps (Lu et al. 2019): (i) minimizing the core matrix, and (ii) minimizing TNN.

(i) Updating core matrix as

$$\min_{\mathfrak{S}} \|\mathfrak{S}\|_* + \frac{1}{2\lambda} \|\mathcal{F} - (\mathcal{Z} + \frac{\mathcal{Y}}{\mu})\|_F^2 \quad (10)$$

where regularization parameter $\lambda = 1/(\max(m, v)n)^{\frac{1}{2}}$. And the tensor \mathcal{S} is obtained from *t*-SVD on the temporary variable \mathcal{F} , *i.e.*, $\mathcal{F} = \mathcal{U} * \mathcal{S} * \mathcal{V}$.

(ii) Updating A as

$$\min_{\mathcal{W}} \|\mathcal{W}\|_{etnn} + \frac{\mu}{2} \|\mathcal{W} - \mathcal{G}\|_F^2 \quad (11)$$

With the learned low-rank core matrix $\mathfrak{S}(\hat{\mathbf{T}})$, we can use *t*-product to reconstruct a tensor as $\mathcal{G} = \mathcal{U} * \mathfrak{S}^{-1}(\mathcal{T}) * \mathcal{V}$. The learned \mathcal{G} can further produce a closed-form solution via **Theorem 1** provided in supplementary materials. Its complexity costs $\mathcal{O}(ml + ml \log(m))$.

Update-4: Solving E is similar to the subproblem of A, which costs $\mathcal{O}(nl + nl \log(n))$ complexity. To update the multiplier variable Y, we employ the following procedure:

$$\begin{aligned} \mathcal{Y} &= \mathcal{Y} + \mu(\mathbf{Q} - \mathcal{A}), \eta = \min(\rho_2 \eta_2, \eta_{max}) \\ \mathcal{J} &= \mathcal{J} + \mu(\mathcal{B} - \mathcal{E}), \eta = \min(\rho_3 \eta_3, \eta_{max}) \end{aligned} \quad (12)$$

In this scenario, we set the parameter η to $1e^{-4}$ and the upper bound for μ (μ_{max}) to 10^{10} . $\rho_2 = 2$ is employed. The computational complexity of the algorithm is linear to n . The solution of problem (4) is provided in Algorithm 1. *obj* is the objective value.

After obtaining the multiple hash matrices $\{\mathbf{B}_p\}_{p=1}^v$, the final hashing matrix is further averaged as $\hat{\mathbf{B}} = \sum_{p=1}^v \mathbf{B}_p$. Then, the clustering indicator matrix could be achieved via the following

$$\begin{aligned} & \min_{\mathbf{C}, \mathbf{G}} \|\hat{\mathbf{B}} - \mathbf{C}\mathbf{G}\|_F^2 \\ \text{s.t. } & \mathbf{C}^\top \mathbf{1} = 0, \mathbf{C} \in \{-1, 1\}^{l \times k}, \mathbf{G} \in \{0, 1\}^{k \times n}, \sum_i g_{in} = 1. \end{aligned} \quad (13)$$

which can be solved according to the discrete proximal linearized minimization (DPLM) (Wang et al. 2023).

Optimizing G-Step is

$$g_{ij}^{t+1} = \begin{cases} 1, & j = \arg \min_* H(b_i, c^{t+1}), \\ 0, & \text{otherwise,} \end{cases} \quad (14)$$

where $H(\mathbf{b}_i, \mathbf{c}^n)$ represents the Hamming distance between the *i*-th binary code b_i and the *s*-th cluster centroid c^s , offering significantly faster computation compared to the Euclidean distance. The final results are obtained from G.

Time Complexity. The main complexity cost of Algorithm 1 cost $\mathcal{O}(m^3l + nml + nm^2 + mln + ncl + nl + nl \log(n) + ml + ml \log(m))$, which is linear to sample number n since $\log(n) \ll n$.

Algorithm 1: TPCH algorithm

Input: Bipartite graphs $\{\phi(\mathbf{X}_p)\}_{p=1}^v$ and Parameter α .

Output: Clustering results.

- 1: **repeat**
- 2: Update \mathbf{Q}_p via Eq. (5);
- 3: Update \mathbf{B}_p via Eq. (6);
- 4: Update \mathcal{A} via Eq. (9);
- 5: Update \mathcal{E} similar to \mathcal{A} ;
- 6: **until** Satisfy convergence.

Experiment

In this section, we introduce the datasets, comparisons, evaluation metrics, and experiment analysis. To assess the performance of TPCH, we compare two categories of competitors, *i.e.*, graph-based methods and hash-based methods. Additionally, we further evaluate parameter sensitivity analysis, convergence analysis, and ablation studies.

Methods	Graph-based Methods								Hash-based Methods											
	SC	Co-re-p	Co-re-c	AM-GL	Mul-NMF	ML-AN	mPAC	GMC	SH	DSH	SP	ITQ	SGH	RS-SH	RF-DH	HS-IC	BM-VC	AC-MVBC	GC-AE	TP-CH
	Acc								Acc											
SUNRGBD	0.11	0.18	0.18	0.10	0.14	0.19	0.13	0.22	0.12	0.16	0.19	0.19	0.18	0.16	0.17	0.16	0.14	0.19	0.24	0.22
Cifar-10	0.17	0.22	0.21	0.22	0.12	0.23	0.10	0.23	0.17	0.22	0.23	0.23	0.22	0.20	0.23	0.22	0.24	0.29	0.25	0.70
Caltech101	0.18	0.26	0.26	0.15	0.19	0.23	0.20	0.27	0.18	0.16	0.21	0.25	0.23	0.29	0.22	0.24	0.29	0.32	0.30	0.54
Caltech256	0.09	0.09	0.10	0.05	0.06	0.08	0.09	0.07	0.06	0.08	0.09	0.09	0.08	0.09	0.08	0.07	0.09	0.05	0.11	0.13
100 leaves	0.49	0.73	0.79	0.76	0.87	0.74	0.82	0.43	0.47	0.45	0.48	0.46	0.51	0.36	0.45	0.66	0.50	0.83	0.89	0.86
	NMI								NMI											
SUNRGBD	0.01	0.21	0.22	0.19	0.10	0.21	0.07	0.22	0.13	0.22	0.22	0.21	0.22	0.20	0.20	0.22	0.16	0.24	0.22	0.28
Cifar-10	0.077	0.09	0.09	0.09	0.02	0.09	0.01	0.10	0.03	0.09	0.10	0.10	0.10	0.07	0.10	0.09	0.10	0.13	0.10	0.67
Caltech101	0.18	0.26	0.26	0.15	0.19	0.23	0.20	0.27	0.33	0.36	0.40	0.44	0.44	0.49	0.44	0.45	0.49	0.51	0.47	0.80
Caltech256	0.28	0.30	0.28	0.11	0.15	0.28	0.22	0.14	0.26	0.29	0.30	0.27	0.30	0.31	0.29	0.24	0.32	0.22	0.29	0.37
100 leaves	0.77	0.88	0.93	0.91	0.93	0.89	0.93	0.70	0.72	0.73	0.72	0.74	0.76	0.62	0.72	0.83	0.73	0.93	0.94	0.97
	Purity								Purity											
SUNRGBD	0.11	0.33	0.34	0.11	0.16	0.33	0.14	0.23	0.24	0.33	0.34	0.34	0.34	0.33	0.33	0.35	0.28	0.34	0.34	0.37
Cifar-10	0.18	0.22	0.22	0.23	0.12	0.25	0.10	0.25	0.17	0.23	0.23	0.23	0.22	0.21	0.23	0.22	0.24	0.27	0.26	0.73
Caltech101	0.31	0.46	0.47	0.17	0.32	0.44	0.30	0.35	0.31	0.34	0.39	0.42	0.41	0.47	0.42	0.41	0.49	0.46	0.44	0.70
Caltech256	0.13	0.14	0.15	0.04	0.11	0.14	0.14	0.10	0.11	0.13	0.14	0.14	0.14	0.15	0.13	0.11	0.15	0.08	0.14	0.18
100 leaves	0.56	0.76	0.83	0.81	0.90	0.76	0.85	0.54	0.50	0.50	0.51	0.49	0.53	0.39	0.48	0.68	0.53	0.88	0.90	0.88

Table 1: Comparison results with graph-based and hash-based methods

Experimental Setting

Five benchmark multi-view datasets include:

- **SUNRGBD:** This dataset consists of colour images and depth images of indoor scenes to provide rich information. This dataset comprises 37 categories, totalling 10,335 images.
- **Cifar-10:** This dataset consists of colour images categorized into 10 different classes, each containing 6,000 32x32-pixel colour images.
- **Caltech101:** Created by the California Institute of Technology, this dataset contains 101 object categories with approximately 9,000 images, including animals, plants, vehicles, and various other objects.
- **Caltech256:** Similar to Caltech101, it comprises 256 categories, totaling approximately 30,607 images.
- **100 leaves:** This dataset includes 100 plant species with a total of 1,600 samples. Each sample consists of three views: shape, fine-scale margin, and texture histograms.

Comparison methods and setting: To evaluate the validity of our TPCH, two categories of competitors are employed: (1) Hash-based methods include SH (Weiss, Torralba, and Fergus 2008), DSH (Jin et al. 2013), SP (Xia et al. 2015), ITQ (Gong et al. 2012), SGH (Jiang and Li 2015), RSSH (Tian et al. 2020), RFDH (Wang, Wang, and Gao 2017), HSIC (Zhang et al. 2018a), BMVC (Zhang et al. 2018b), AC-MVBC (Zhang et al. 2022) and GCAE (Wang et al. 2023). (2) Graph based methods include SC (Ng, Jordan, and Weiss 2001), Co-regularize (Kumar, Rai, and Daume 2011), AMGL (Nie et al. 2016), Mul-NMF (Liu

et al. 2013), MLAN (Nie, Cai, and Li 2017), mPAC (Kang et al. 2019), GMC (Wang, Yang, and Liu 2019) and GCAE. For these methods, we conduct experiments using publicly available source codes and parameter settings of the competitors. By comparing against a large number of advanced methods, we can thoroughly verify the effectiveness of our TPCH. In the experiment, five mainstream metrics are used to evaluate clustering performance: Accuracy (ACC), Normalized Mutual Information (NMI), Purity, F-score, and Adjusted Rand Index (ARI). Due to space limitations, their definitions, along with the experimental results of F-score and ARI, are provided in the supplementary materials. The hardware for running our experiments includes CPU: i9 14900k 24-core processor, RAM: 32GB 7200MHz, Platform: Windows 11, Software: MATLAB 2023b.

Comparative Experimental Results

Table 1 reports the clustering performance of TPCH and various hash-based and graph-based clustering methods on five multi-view datasets, with the optimal results highlighted in bold. According to the results of the five clustering evaluation metrics, Table 1 demonstrates that:

- AC-MVBC, GCAE, and our TPCH outperform other methods significantly. The reasons may be that: 1) AC-MVBC and our TPCH could capture the high-order information among multi-view data, which enhances the discriminability of hash codes; 2) GCAE and our TPCH encode the dynamic graph semantic information into hash codes to enlarge their quality.
- Compared to most recently GCAE and AC-MVBC,

TPCH also outperforms them consistently in most cases. The main reasons are that 1) our TPCH first captures the high-order information of data in dual projection and Hamming spaces to learn the more compact and discriminable hash representations as shown in Fig. 5 and Fig. 6. This phenomenon indicates that our enhanced tensor nuclear norm on dual projection and Hamming spaces could simultaneously promote the projection synergy, as well as inter-view and intra-view communication.

- Conventional multi-view clustering methods use Euclidean distance to measure the distance between two samples, which leads to inefficiency and high computational complexity. In contrast, hash-based methods obtain clustering results in the Hamming space, thereby improving computational efficiency.

Robust analysis: Fig. 3 reports the comparison experiment results of TPCH with AC-MVBC on the Caltech101 and Caltech256 datasets with salt-and-pepper noise. Experiments conducted on the SUNRGBD, Cifar-10 and 100 leaves datasets are provided in the Supplementary Materials. Through Fig. 3, it can be observed that TPCH is completely superior to AC-MVBC on five main clustering indicators. Similar to TPCH, AC-MVBC also uses tensor nuclear norm to capture the latent higher-order correlations in multi-view data, but it ignores the low-rank property of the core tensor, making existing tensor nuclear norm more susceptible to noise. More importantly, core tensor with high rank limits the higher-order synergistic between the projection matrices. By further explore the low-rank property of core tensor on dual projection and Hamming spaces, the projection co-operation and inter-view/intra-view communication would negotiate with each other to achieve the compact hash representation. Thus, our TPCH could enjoy more robustness to noise and better generate the more compact hash codes for improve the clustering performance. Furthermore, experiments on noisy datasets also verify the robustness of TPCH to noise as mentioned in Fig. 4.

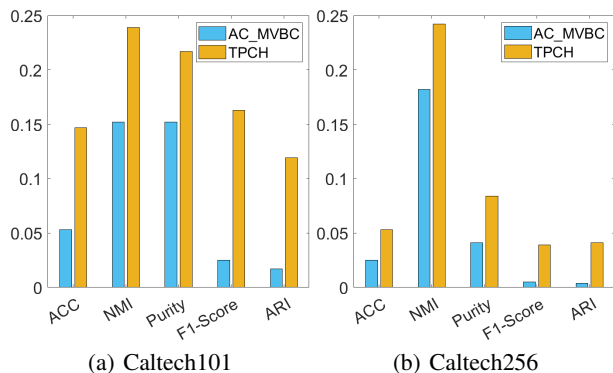


Figure 3: Clustering performance of AC_MVBC and TPCH on the two datasets with salt-and-pepper noise.

Table 2 demonstrates the runtime of TPCH compared to the state-of-the-art method, GCAE, across five large-scale datasets. It is evident from Table 2 that TPCH significantly outperforms GCAE in terms of computational speed despite

both methods being based on hashing. The time complexity of GCAE is $\mathcal{O}(n^3 + 3n^2l + n^2 + nlc)$, which implies cubic complexity with respect to the number of samples, whereas TPCH has a time complexity of $\mathcal{O}(m^3l + nml + nm^2 + mln + ncl + nl + nl \log(n) + ml + ml \log(m))$, indicating linear complexity for the number of samples. The lower time complexity grants TPCH higher computational efficiency.

In summary, TPCH not only excels in various metrics for executing multi-view clustering tasks but also significantly surpasses existing algorithms in computational efficiency, making it well-suited for rapidly processing large-scale multi-view clustering tasks.

Datasets	GCAE	TPCH	Speedup
CPU-Time(s)			
SUNRGBD	1188.33	157.62	7.54 ×
Cifar-10	1120.15	157.25	7.12 ×
Caltech101	1158.71	114.11	10.15 ×
Caltech256	47618.00	5406.76	8.81 ×
100 leaves	14.16	1.75	8.09 ×

Table 2: Running time of GCAE and TPCH.

Ablation Studies

The TPCH model mainly includes two tensors: \mathcal{Q} and \mathcal{B} . To further investigate why TPCH performs well, we remove \mathcal{Q} and \mathcal{B} respectively, naming these methods as without (w/o) \mathcal{B} and without (w/o) \mathcal{Q} , while conducting experiments with datasets processed with salt-and-pepper noise. The results of the ablation study are shown in Fig.4. From Fig. 4, we observe that the removal of \mathcal{Q} and \mathcal{B} leads to a significant decline in clustering performance, verifying the robustness of proposed enhanced tensor nuclear norm on projection matrices and hash codes.

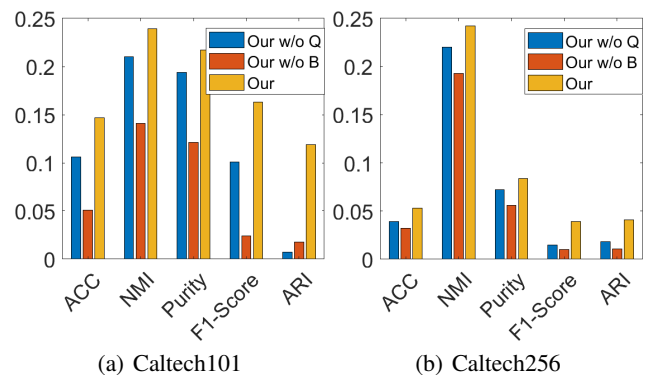


Figure 4: Ablation Studies of TPCH on the Caltech101 and Caltech256 datasets with salt-and-pepper noise.

Visualization of Clustering Results

The visualization of clustering results for AC-MVBC and TPCH on the Synthetic_4clu datasets is presented in Fig.5. Different colors represent different clusters learned on the

hash codes. We also visualized the clustering results on the Synthetic_3d dataset. Due to space constraints, we have included it in the supplementary materials.

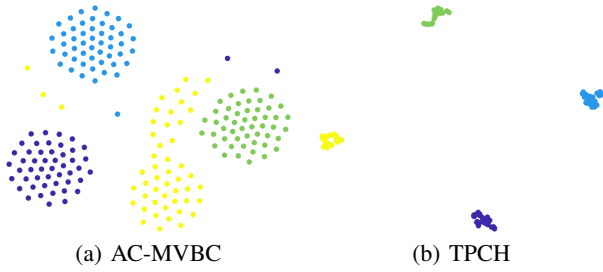


Figure 5: Visualization of clustering results for TPC and AC-MVBC on the Synthetic_4clu dataset.

Fig. 5(a) and Fig. 5(b) reveals that the clustering performance of TPC on the Synthetic_4clu dataset is significantly better than AC-MVBC. Fig. 5(b) exhibits larger inter-cluster distances and smaller intra-cluster distances. Additionally, the clusters displayed in Fig. 5(b) are more compact, directly demonstrating the effectiveness of TPC in improving hash code compactness.

Fig. 6 displays the performance of TPC at the beginning and after convergence of clustering on the Synthetic_4clu and Synthetic_3d datasets. It can be seen that in both datasets, TPC exhibits a clear clustering effect after convergence. It is noteworthy that, despite the more complex data distribution in Synthetic_3d compared to Synthetic_4clu, TPC can still effectively partition the clusters. The stable clustering performance of TPC under these two different initial data distributions once again proves the superiority of TPC.

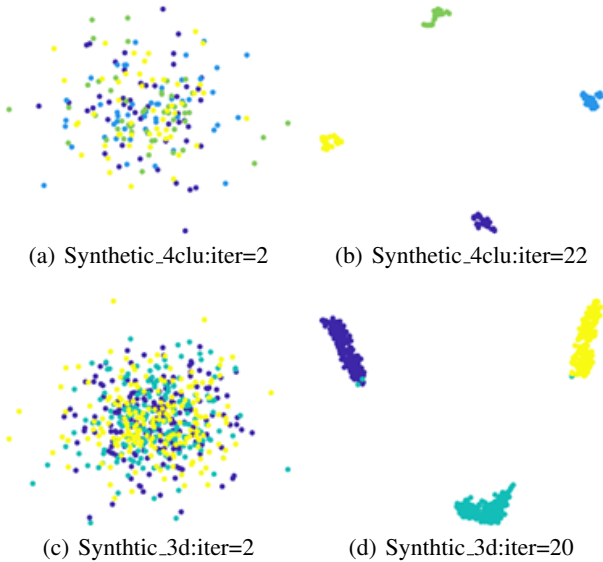


Figure 6: Clustering performance of TPC at the beginning and after convergence of clustering on the Synthetic_4clu and Synthetic_3d datasets.

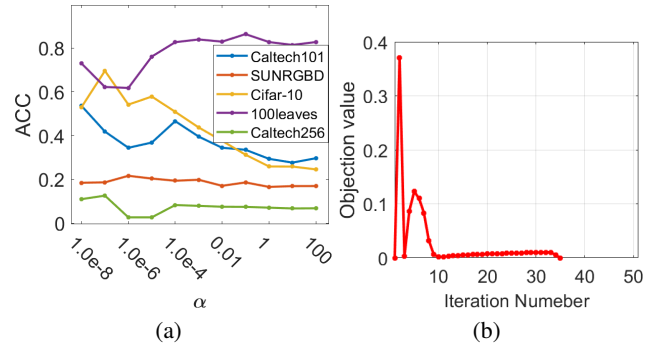


Figure 7: Sensitive analysis *w.r.t.* α for five large-scale datasets and convergence curve for the Caltech256 dataset.

Sensitivity and Convergence Analysis

In TPC, we set the dimensionality of the hash codes and the representation space to be 64. TPC has only one hyperparameter, α , which adjusts the balance between the binary representation of multi-view data and tensor factorization. We employ a logarithmic search approach, varying α between 1.0e-8 and 100 at each order of magnitude. Fig. 7 illustrates the sensitivity of different datasets to α across the clustering metrics ACC. The 100leaves dataset maintains high performance across all α . As shown in Fig. 7 (a), TPC achieves satisfactory results over a wide range of α values, indicating insensitivity to parameters. For the sensitivity analysis of more metrics, please refer to supplementary materials. Fig. 7 (b) illustrates the convergence curves of TPC on the Caltech256 datasets, which demonstrates that TPC converges very quickly. For the convergence of more metrics, please refer to supplementary materials.

Conclusion

In conclusion, this work introduces a novel Tensor-interacted Projection and Cooperative Hashing for Multi-view Clustering(TPC) approach for multi-view clustering, which effectively addresses the limitations of existing methods by considering the synergistic interactions and communications during the projection process. This innovative approach significantly enhances the compactness and distinguishability of hash representations, thereby improving both the quality of clustering and the robustness against noise. Experimental results validate the effectiveness of TPC in multi-view clustering scenarios.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (project no. 62406069), the National Natural Science Foundation of China (Grant No. 62372235), the China Postdoctoral Science Foundation (project no. 2024M750425), the Mianyang Science and Technology Program (Grant nos. 2023ZYDF003, 2023ZYDF091), the Sichuan Science and Technology Program (Grant no. 2023YFG0135), and the Sichuan Science and Technology Program (Grant No. MZGC20240057).

References

- Chen, M.-S.; Wang, C.-D.; and Lai, J.-H. 2022. Low-rank tensor based proximity learning for multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 35(5): 5076–5090.
- Chen, Z.; Wu, Z.; Wang, S.; and Guo, W. 2023. Dual low-rank graph autoencoder for semantic and topological networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 4191–4198.
- Fang, U.; Li, M.; Li, J.; Gao, L.; Jia, T.; and Zhang, Y. 2023. A Comprehensive Survey on Multi-View Clustering. *IEEE Transactions on Knowledge and Data Engineering*, 35(12): 12350–12368.
- Gong, Y.; Lazebnik, S.; Gordo, A.; and Perronnin, F. 2012. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *IEEE transactions on pattern analysis and machine intelligence*, 35(12): 2916–2929.
- Gong, Y.; Pawlowski, M.; Yang, F.; Brandy, L.; Bourdev, L.; and Fergus, R. 2015. Web scale photo hash clustering on a single machine. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 19–27.
- Han, Z.; Zhang, C.; Fu, H.; and Zhou, J. T. 2022. Trusted multi-view classification with dynamic evidential fusion. *IEEE transactions on pattern analysis and machine intelligence*, 45(2): 2551–2566.
- Hu, K.; Li, Y.; Zhang, S.; Wu, J.; Gong, S.; Jiang, S.; and Weng, L. 2024a. FedMMD: A Federated weighting algorithm considering Non-IID and Local Model Deviation. *Expert Systems with Applications*, 237: 121463.
- Hu, K.; Shen, C.; Wang, T.; Shen, S.; Cai, C.; Huang, H.; and Xia, M. 2023. Action Recognition Based on Multi-Level Topological Channel Attention of Human Skeleton. *Sensors*, 23(24): 9738.
- Hu, K.; Zhang, E.; Xia, M.; Wang, H.; Ye, X.; and Lin, H. 2024b. Cross-dimensional feature attention aggregation network for cloud and snow recognition of high satellite images. *Neural Computing and Applications*, 36(14): 7779–7798.
- Jiang, Q.-Y.; and Li, W.-J. 2015. Scalable graph hashing with feature transformation. In *International Joint Conference on Artificial Intelligence*, volume 15, 2248–2254.
- Jin, Z.; Li, C.; Lin, Y.; and Cai, D. 2013. Density sensitive hashing. *IEEE transactions on cybernetics*, 44(8): 1362–1371.
- Kang, Z.; Guo, Z.; Huang, S.; Wang, S.; Chen, W.; Su, Y.; and Xu, Z. 2019. Multiple partitions aligned clustering. *arXiv preprint arXiv:1909.06008*.
- Kang, Z.; Zhou, W.; Zhao, Z.; Shao, J.; Han, M.; and Xu, Z. 2020. Large-Scale Multi-View Subspace Clustering in Linear Time. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(04): 4412–4419.
- Kumar, A.; Rai, P.; and Daume, H. 2011. Co-regularized multi-view spectral clustering. *Advances in neural information processing systems*, 24.
- Li, M.; Wang, S.; Liu, X.; and Liu, S. 2024a. Parameter-Free and Scalable Incomplete Multiview Clustering With Prototype Graph. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1): 300–310.
- Li, X.; Pan, Y. P.; Sun, Y.; Sun, Q.; Sun, Y.; W. Tsang, I.; and Ren, Z. 2024b. Incomplete Multi-view Clustering with Paired and Balanced Dynamic Anchor Learning. *IEEE Transactions on Multimedia*.
- Li, X.; Pan, Y. P.; Sun, Y.; Sun, Q. S.; Tsang, I. W.; and Ren, Z. 2024c. Fast Unpaired Multi-view Clustering.
- Li, X.; Ren, Z.; Sun, Q.; and Xu, Z. 2023. Auto-weighted tensor Schatten p-norm for robust multi-view graph clustering. *Pattern Recognition*, 134: 109083.
- Liu, C.; Wen, J.; Liu, Y.; Huang, C.; Wu, Z.; Luo, X.; and Xu, Y. 2024a. Masked two-channel decoupling framework for incomplete multi-view weak multi-label learning. *Advances in Neural Information Processing Systems*, 36: 32387–32400.
- Liu, J.; Wang, C.; Gao, J.; and Han, J. 2013. Multi-view clustering via joint nonnegative matrix factorization. In *Proceedings of the 2013 SIAM international conference on data mining*, 252–260.
- Liu, S.; Liu, X.; Wang, S.; Niu, X.; and Zhu, E. 2024b. Fast incomplete multi-view clustering with view-independent anchors. *IEEE Transactions on Neural Networks and Learning Systems*, 1–12.
- Long, Z.; Zhu, C.; Chen, J.; Li, Z.; Ren, Y.; and Liu, Y. 2023. Multi-view MERA subspace clustering. *IEEE Transactions on Multimedia*.
- Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; and Yan, S. 2019. Tensor robust principal component analysis with a new tensor nuclear norm. *IEEE transactions on pattern analysis and machine intelligence*, 42(4): 925–938.
- Lu, Z.; Liu, G.; and Wang, S. 2020. Sparse neighbor constrained co-clustering via category consistency learning. *Knowledge-Based Systems*, 201: 105987.
- Lu, Z.; Nie, F.; Wang, R.; and Li, X. 2023a. A differentiable perspective for multi-view spectral clustering with flexible extension. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(6): 7087–7098.
- Lu, Z.; Wang, S.; Liu, G.; and Nie, F. 2023b. Robust weighted co-clustering with global and local discrimination. *Pattern Recognition*, 138: 109405.
- Ng, A.; Jordan, M.; and Weiss, Y. 2001. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 14.
- Nie, F.; Cai, G.; and Li, X. 2017. Multi-view clustering and semi-supervised classification with adaptive neighbours. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31.
- Nie, F.; Li, J.; Li, X.; et al. 2016. Parameter-free auto-weighted multiple graph learning: a framework for multi-view clustering and semi-supervised classification. In *International Joint Conference on Artificial Intelligence*, volume 9, 1881–1887.

- Nie, F.; Liu, C.; Wang, R.; Wang, Z.; and Li, X. 2021. Fast fuzzy clustering based on anchor graph. *IEEE Transactions on Fuzzy Systems*, 30(7): 2375–2387.
- Qin, Y.; Pu, N.; and Wu, H. 2024. EDMC: Efficient Multi-View Clustering via Cluster and Instance Space Learning. *IEEE Transactions on Multimedia*, 26: 5273–5283.
- Qin, Y.; Qin, C.; Zhang, X.; and Feng, G. 2024. Dual Consensus Anchor Learning for Fast Multi-View Clustering. *IEEE Transactions on Image Processing*, 33: 5298–5311.
- Shen, X.; Liu, W.; Tsang, I.; Shen, F.; and Sun, Q.-S. 2017. Compressed k-means for large-scale clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 31.
- Sun, Y.; Qin, Y.; Peng, D.; Peng, X.; and Hu, P. 2024. Robust Multi-View Clustering with Noisy Correspondence. *IEEE Transactions on Knowledge and Data Engineering*, 1–14.
- Sun, Y.; Ren, Z.; Hu, P.; Peng, D.; and Wang, X. 2023a. Hierarchical consensus hashing for cross-modal retrieval. *IEEE Transactions on Multimedia*.
- Sun, Y.; Wang, X.; Peng, D.; Ren, Z.; and Shen, X. 2023b. Hierarchical hashing learning for image set classification. *IEEE Transactions on Image Processing*, 32: 1732–1744.
- Tian, Z.; Zhang, H.; Chen, Y.; and Zhang, D. 2020. Unsupervised hashing based on the recovery of subspace structures. *Pattern Recognition*, 103: 107261.
- Wang, D.; Wang, Q.; and Gao, X. 2017. Robust and flexible discrete hashing for cross-modal similarity search. *IEEE transactions on circuits and systems for video technology*, 28(10): 2703–2715.
- Wang, H.; Nie, F.; Huang, H.; and Makedon, F. 2011. Fast nonnegative matrix tri-factorization for large-scale data co-clustering. In *Twenty-Second International Joint Conference on Artificial Intelligence*. Citeseer.
- Wang, H.; Yang, Y.; and Liu, B. 2019. GMC: Graph-based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 32(6): 1116–1129.
- Wang, H.; Yao, M.; Jiang, G.; Mi, Z.; and Fu, X. 2023. Graph-collaborated auto-encoder hashing for multiview binary clustering. *IEEE Transactions on Neural Networks and Learning Systems*.
- Wang, S.; Liu, X.; Liu, L.; Tu, W.; Zhu, X.; Liu, J.; Zhou, S.; and Zhu, E. 2022a. Highly-efficient incomplete large-scale multi-view clustering with consensus bipartite graph. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 9776–9785.
- Wang, Z.; Lu, H.; Jin, J.; and Hu, K. 2022b. Human Action Recognition Based on Improved Two-Stream Convolution Network. *Applied Sciences*, 12(12): 5784.
- Weiss, Y.; Torralba, A.; and Fergus, R. 2008. Spectral hashing. *Advances in neural information processing systems*, 21.
- Wen, J.; Liu, C.; Deng, S.; Liu, Y.; Fei, L.; Yan, K.; and Xu, Y. 2023a. Deep double incomplete multi-view multi-label learning with incomplete labels and missing views. *IEEE transactions on neural networks and learning systems*.
- Wen, J.; Xu, G.; Tang, Z.; Wang, W.; Fei, L.; and Xu, Y. 2023b. Graph regularized and feature aware matrix factorization for robust incomplete multi-view clustering. *IEEE Transactions on Circuits and Systems for Video Technology*.
- Wen, J.; Zhang, Z.; Fei, L.; Zhang, B.; Xu, Y.; Zhang, Z.; and Li, J. 2023c. A Survey on Incomplete Multiview Clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 53(2): 1136–1149.
- Wu, Z.; Lin, X.; Lin, Z.; Chen, Z.; Bai, Y.; and Wang, S. 2023. Interpretable graph convolutional network for multi-view semi-supervised learning. *IEEE Transactions on Multimedia*.
- Xia, W.; Gao, Q.; Wang, Q.; Gao, X.; Ding, C.; and Tao, D. 2023. Tensorized bipartite graph learning for multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4): 5187–5202.
- Xia, Y.; He, K.; Kohli, P.; and Sun, J. 2015. Sparse projections for high-dimensional binary codes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3332–3339.
- Yang, B.; Wang, S.; Markham, A.; and Trigoni, N. 2020. Robust Attentional Aggregation of Deep Feature Sets for Multi-view 3D Reconstruction. *International Journal of Computer Vision*, 128(1): 53–73.
- Zhang, H.; Nie, F.; and Li, X. 2023. Large-Scale Clustering With Structured Optimal Bipartite Graph. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8): 9950–9963.
- Zhang, R.; and Lu, Z. 2016. Large Scale Sparse Clustering. In *International Joint Conference on Artificial Intelligence*, 2336–2342.
- Zhang, X.; Yang, Y.; Li, T.; Zhang, Y.; Wang, H.; and Fujita, H. 2021. CMC: A consensus multi-view clustering model for predicting Alzheimer’s disease progression. *Computer Methods and Programs in Biomedicine*, 199: 105895.
- Zhang, Y.; Xie, Y.; Li, C.; Wu, Z.; and Qu, Y. 2022. Learning all-in collaborative multiview binary representation for clustering. *IEEE Transactions on Neural Networks and Learning Systems*.
- Zhang, Y.; Xie, Y.; Li, C.; Wu, Z.; and Qu, Y. 2024. Learning All-In Collaborative Multiview Binary Representation for Clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 35(3): 4260–4273.
- Zhang, Z.; Liu, L.; Qin, J.; Zhu, F.; Shen, F.; Xu, Y.; Shao, L.; and Shen, H. T. 2018a. Highly-economized multi-view binary compression for scalable image clustering. In *Proceedings of the European conference on computer vision (ECCV)*, 717–732.
- Zhang, Z.; Liu, L.; Shen, F.; Shen, H. T.; and Shao, L. 2018b. Binary multi-view clustering. *IEEE transactions on pattern analysis and machine intelligence*, 41(7): 1774–1782.
- Zhao, J.; Kang, F.; Zou, Q.; and Wang, X. 2023. Multi-view clustering with orthogonal mapping and binary graph. *Expert Systems with Applications*, 213: 118911.