

# MCFG with GUMAP: A Simple and Effective Clustering Framework on Grassmann Manifold

Benchao Li<sup>1</sup>, Yun Zou<sup>1</sup>, Ruisheng Ran<sup>2</sup>  
Chongqing Normal University  
Chongqing Normal University, Chongqing 401331, China  
rshran@cqnu.edu.cn

## Abstract

In this study, we propose a simple and efficient clustering framework on Grassmann manifold space. Initially, we address the transformation of distance metrics from Euclidean space to Grassmann manifold space, thereby extending existing clustering methods in Euclidean space to Grassmann manifold space. For convenience, we refer to this method as the Metric-Based Clustering Framework on Grassmann Manifold (MCFG). To further enhance the performance of the clustering framework, we introduce the Uniform Manifold Approximation and Projection on Grassmann Manifold (GUMAP). GUMAP is employed to extract key features from image-set data, which are subsequently utilized within the aforementioned clustering method. We designate this integrated approach as MCFG with GUMAP. This method is applicable to all clustering analyses, thereby presenting a straightforward and effective clustering framework. Experimental results on multiple image-set datasets demonstrate that MCFG with GUMAP outperforms both MCFG and existing clustering methods on Grassmann manifold. MCFG with GUMAP effectively transfers clustering methods from Euclidean space to Grassmann manifold, establishing itself as a potent tool for clustering tasks on Grassmann manifold.

*Keywords: Grassmann Manifold Image-Set Analysis Dimensionality Reduction Clustering UMAP*

## 1. Introduction

In the era of big data, the rapid advancements in computer technology and the drastic surge in data volume have necessitated heightened demands for data processing and analysis techniques. Cluster analysis[22], a pivotal technique within the realm of unsupervised learning, possesses the capability to partition data into distinct categories or

clusters based on their intrinsic features, solely relying on the similarities or differences among samples, without pre-existing labels. This approach facilitates the discovery of latent patterns and structures within vast datasets. Consequently, clustering analysis plays an indispensable role across numerous domains, including data mining and image processing. Classical clustering algorithms, such as K-means[4, 14] and Spectral Clustering[23], have been extensively applied to various data analysis tasks. These algorithms not only uncover potential patterns and anomalies within data but also provide support for data analysis and decision-making. Therefore, conducting in-depth research into clustering algorithms and their applications is of significant importance for advancing the field of data science, fostering knowledge discovery, and optimizing decision-making processes.

As data scales and complexity continue to escalate, traditional clustering methods struggle to capture the intrinsic structures of high-dimensional data, necessitating an increasingly urgent need for the optimization and extension of clustering algorithms. This is crucial for enhancing the efficiency and accuracy of data analysis. Numerous studies[13, 19] have demonstrated that preliminary dimensionality reduction can effectively aid the clustering process. By decreasing the dimensions of data, the computational burden of clustering algorithms is alleviated, thereby improving clustering efficiency. Furthermore, high-dimensional data often contain noise and redundant features, which can be eliminated through dimensionality reduction techniques. These techniques retain features that significantly influence clustering outcomes, thereby optimizing clustering performance. For instance, dimensionality reduction methods such as Principal Component Analysis (PCA)[1], t-Distributed Stochastic Neighbor Embedding (t-SNE)[16, 28], and Uniform Manifold Approximation and Projection (UMAP)[18, 30] have been integrated with clustering methods in numerous practical applications. This integration not only enables clustering algorithms to better handle larger datasets but also enhances clustering efficiency.

<sup>1</sup>Benchao Li and Yun Zou have contributed equally to this work.

<sup>2</sup>Corresponding author.

The Grassmann manifold[2, 12], as a specialized type of subspace structure, is extensively utilized in handling intricate high-dimensional data, particularly in the analysis of image-set data and video data. By representing image-sets or videos as low-dimensional subspaces embedded within a high-dimensional manifold space, it effectively addresses the challenges posed by traditional Euclidean space in processing complex high-dimensional video data. This representation not only preserves the geometric structure of the data but also enables efficient computations on the manifold through appropriate distance metrics, thereby finding widespread application in fields such as computer vision and pattern recognition. Currently, several clustering algorithms have been developed on Grassmann manifold, including Clustering on Grassmann Manifold via Kernel Embedding (CGMKE)[24], Grassmann K-Means [25, 27], and Grassmann LBG (GLBG)[25]. These methods leverage the geometric structure of the Grassmann manifold to extend traditional clustering algorithms into a framework of subspace clustering, providing effective solutions for processing high-dimensional image-set data. Furthermore, these algorithms can reasonably cluster the data without compromising its structural integrity, demonstrating the feasibility and superiority of clustering on Grassmann manifold.

Due to the fundamental differences in their principles, various clustering methods often find application in distinct domains. In this study, we attempt to extend multiple clustering methods, which are widely used in Euclidean space, to Grassmann manifold through a straightforward approach. We propose a novel technique, namely the Metric-Based Clustering Framework on Grassmann Manifold (MCFG), which migrates clustering methods from Euclidean space to Grassmann manifold through a predefined similarity matrix based on Grassmann metrics. To enhance the performance of these clustering framework, we introduce Uniform Manifold Approximation and Projection on Grassmann Manifold (GUMAP), a new nonlinear dimensionality reduction technique implemented on Grassmann manifold. By leveraging GUMAP to extract key features of samples in high-dimensional space and subsequently applying MCFG for clustering, we can further improve the performance of MCFG.

The structure of this paper is organized as follows: Section 2 provides a detailed overview of the pertinent work regarding Grassmann manifold, UMAP and clustering analysis; Section 3 elucidates the MCFG technique and the GUMAP approach; Section 4 presents the clustering performance of the MCFG framework, and Section 5 concludes the findings.

## 2. Related Works

### 2.1. Metric on Grassmann Manifold

A Grassmann manifold  $\mathbb{G}(m, D)$  is a collection of all  $m$ -dimensional linear subspaces within a  $D$ -dimensional Euclidean space.[2, 12] Each point on  $\mathbb{G}(m, D)$  represents a linear subspace, which is spanned by  $m$  orthonormal basis vectors in  $D$ -dimensional space. Approximating the geodesic distance on the Grassmann manifold based on linear subspaces is a widely used technique. The projection metric on the Grassmann manifold can be expressed as follows:

$$d_{proj}(\mathbf{X}_1, \mathbf{X}_2) = 2^{-1/2} \left\| \mathbf{X}_1 \mathbf{X}_1^T - \mathbf{X}_2 \mathbf{X}_2^T \right\|_F \quad (1)$$

In addition to projection metric, a series of metrics based on principal angles exist on the Grassmann manifold. The principal angles  $\{\theta_i\}_{i=1}^m$  between  $\mathbf{X}_1$  and  $\mathbf{X}_2$  can be obtained through the Singular Value Decomposition (SVD) of  $\mathbf{X}_1^T \mathbf{X}_2$ . The Procrustes metric based on principal angles between  $\mathbf{X}_1$  and  $\mathbf{X}_2$  can be expressed as follows:

$$d_{CF}(\mathbf{X}_1, \mathbf{X}_2) = 2 \left( \sum_{i=1}^m \sin^2(\theta_i/2) \right)^{1/2} \quad (2)$$

By embedding samples on Grassmann manifold into a Reproducing Kernel Hilbert Space (RKHS), the similarity between samples can be measured through kernel functions defined on Grassmann manifold. The projection kernel function between  $\mathbf{X}_1$  and  $\mathbf{X}_2$  can be represented as:

$$k_{proj}(\mathbf{X}_1, \mathbf{X}_2) = \left\| \mathbf{X}_1^T \mathbf{X}_2 \right\|_F^2 \quad (3)$$

### 2.2. UMAP

Uniform Manifold Approximation and Projection (UMAP)[18] is an advanced nonlinear dimensionality reduction and visualization technique, extensively utilized in high-dimensional data analysis fields such as bioinformatics and spectral image analysis. UMAP is constructed based on the theoretical framework of Riemannian geometry and algebraic topology. Its core idea involves creating a fuzzy topological representation between high-dimensional data and their low-dimensional embeddings, and optimizing the low-dimensional embeddings to ensure that their fuzzy topological representations resemble the high-dimensional data, thereby achieving the desired dimensionality reduction results.

Compared to earlier manifold learning methods such as Local Linear Embedding (LLE)[21] and t-SNE, UMAP is capable of accommodating larger-scale and non-uniformly dense datasets. Additionally, UMAP exhibits lower time complexity, making it more widely applicable in tasks involving the visualization of large-scale high-dimensional data.

### 2.3. Clustering Analysis

Clustering analysis[22] is a pivotal technique in the fields of data mining and machine learning. It involves grouping samples into clusters such that similar samples are assigned to the same cluster, thereby revealing the intrinsic relationships among the data. The objective of clustering analysis is to maximize the similarity within the same cluster while minimizing the similarity between different clusters.

Spectral Clustering (SPC)[23] is a graph-based clustering method whose fundamental idea involves partitioning the undirected connectivity graph of all samples such that the distances among samples within each subgraph are minimized while the distances between samples across different subgraphs are maximized, thereby achieving the purpose of clustering. Affinity Propagation (AFP)[10] represents another graph-based clustering approach that does not require a predefined number of clusters. It achieves robust clustering results by iteratively propagating messages of affinity and responsibility within dense regions of the sample space. Shared Nearest Neighbor-Based Density Peaks Clustering (SNN-DPC)[17] is an improvement upon the traditional density peaks clustering method, offering more accurate identification of cluster centers and demonstrating robustness against noise and outliers.

Density-Based Spatial Clustering of Applications with Noise (DBSCAN)[9] is a clustering method that relies on the density distribution of samples to delineate clusters, commonly utilized for the detection of outliers. Hierarchical DBSCAN (HDBSCAN)[5], as an extension of DBSCAN, is another density-based hierarchical clustering approach that can adaptively adjust the hyperparameters, demonstrating robustness against noise and outliers. Ordering Points To Identify the Clustering Structure (OPTICS)[3] is another density-based clustering method that does not require a predefined number of clusters in the data, capable of discovering complex clustering structures and boundaries within the dataset.

Agglomerative Clustering (AGC)[26] is a widely used hierarchical clustering method that begins with each sample as an independent cluster and iteratively merges the two closest clusters until a specific condition is met.

## 3. Methods

### 3.1. MCFG

The traditional Euclidean space provides an imperfect representation of image-set and video data, leading to suboptimal clustering results when applying conventional methods such as SNN-DPC and DBSCAN to image-set data. A series of clustering methods, including SNN-DPC and DBSCAN, rely on a similarity matrix of the data to perform clustering analysis. By constructing a similarity matrix for samples on Grassmann manifold using an appropriate

metric, these methods, such as SNN-DPC and DBSCAN, can be naturally extended to Grassmann manifold, enabling clustering analysis of image-set and video data.

In this study, the Riemannian metric presented in Table 1 is employed to enhance a range of sample-based clustering methods that rely on similarity matrices, including SNN-DPC, AFP, AGC, DBSCAN, HDBSCAN, OPTICS, and SPC. Consequently, we propose the following methods tailored for Grassmann manifold: Shared Nearest Neighbor DPC on Grassmann Manifold (SNN-DPCG), Affinity Propagation on Grassmann Manifold (GAFFP), Agglomerative Clustering on Grassmann Manifold (GAGC), DBSCAN on Grassmann Manifold (GDBSCAN), HDBSCAN on Grassmann Manifold (GHDBSCAN), OPTICS on Grassmann Manifold (GOPTICS), and Spectral Clustering on Grassmann Manifold (GSPC).

### 3.2. GUMAP

Numerous existing studies have demonstrated that pre-processing data through dimensionality reduction techniques can effectively eliminate noise information and extract pivotal features from the data. Compared to performing clustering analysis directly on the original high-dimensional data, conducting clustering analysis on the extracted pivotal features allows for a better identification of the underlying cluster structure within the data, achieving superior clustering results at a lower computational cost.

In order to discover the low-dimensional manifold structure  $\mathbb{G}(m, d)$ , ( $d < D$ ), embedded within the high-dimensional Grassmann manifold  $\mathbb{G}(m, D)$ , analyze the nonlinear relationship between the high-dimensional and low-dimensional manifold structures, and extract key features from samples on the high-dimensional Grassmann manifold to assist in clustering tasks on Grassmann manifold, this paper proposes the Uniform Manifold Approximation and Projection on Grassmann Manifold (GUMAP).

Firstly, a  $k$ -nearest neighbor graph is constructed on the high-dimensional manifold to represent the neighborhood structure among samples. For a point  $\mathbf{X}_i \in \mathbb{G}(m, D)$  in a given sample set  $\{\mathbf{X}_i\}_{i=1}^N$ , it is straightforward to obtain a set  $\{\mathbf{X}_i^1, \mathbf{X}_i^2, \dots, \mathbf{X}_i^k\}$  of its  $k$  neighbors. The similarity between the sample  $\mathbf{X}_i$  and its neighbor pairs can be expressed as:

$$P(\mathbf{X}_i, \mathbf{X}_i^j) = \exp\left(-\frac{d_{proj}(\mathbf{X}_i, \mathbf{X}_i^j) - \rho_i}{\sigma_i}\right) \quad (4)$$

where,  $\rho_i$  denotes the distance between sample point  $\mathbf{X}_i$  and its nearest neighbor, which can be specifically expressed as:

$$\rho_i = \min\left\{d_{proj}(\mathbf{X}_i, \mathbf{X}_i^j) \mid i \leq j \leq k\right\} \quad (5)$$

Table 1: The Metrics Employed in Various Euclidean Clustering Approaches and Grassmann Methods.

Euclidean Methods	Euclidean Metrics	Grassmann Methods	Grassmann Metrics
AFP	$-\ x - y\ _2^2$	GAFP	$-2 \left(\sum_{i=1}^m \sin^2(\theta_i/2)\right)^{1/2}$
AGC	$\ x - y\ _2$	GAGC	$2 \left(\sum_{i=1}^m \sin^2(\theta_i/2)\right)^{1/2}$
DBSCAN	$\ x - y\ _2$	GDBSCAN	$2 \left(\sum_{i=1}^m \sin^2(\theta_i/2)\right)^{1/2}$
HDBSCAN	$\ x - y\ _2$	GHDBSCAN	$2 \left(\sum_{i=1}^m \sin^2(\theta_i/2)\right)^{1/2}$
OPTICS	$\ x - y\ _2$	GOPTICS	$2 \left(\sum_{i=1}^m \sin^2(\theta_i/2)\right)^{1/2}$
SPC	$\exp\left(-\frac{\ x-y\ _2^2}{2\sigma^2}\right)$	GSPC	$\ X_1^T X_2\ _F^2$
SNN-DPC	$\ x - y\ _2$	SNN-DPCG	$2 \left(\sum_{i=1}^m \sin^2(\theta_i/2)\right)^{1/2}$

In Eq. (4),  $\sigma_i$  represents the normalization factor of sample point  $X_i$ , and it can be calculated by performing the binary search on Eq. (6).

$$\sum_{j=1}^k P(X_i, X_j^j) = \log_2 k \quad (6)$$

Accordingly, the similarity between the projected points  $Y_i$  and  $Y_i^j$  of  $X_i$  and  $X_i^j$  onto the low-dimensional Grassmann manifold can be defined as:

$$Q(Y_i, Y_i^j) = \left[1 + ad_{proj}^{2b}(Y_i, Y_i^j)\right]^{-1} \quad (7)$$

where,  $a$  and  $b$  are 2 hyperparameters, which can be solved by fitting Eq. (7) using Eq. (8).

$$\Omega(X, Y) = \begin{cases} 1 & \text{if } \|X - Y\|_F \leq mindist \\ \exp(-(\|X - Y\|_F - mindist)) & \text{otherwise} \end{cases} \quad (8)$$

GUMAP can minimize the discrepancy between  $P$  and  $Q$  through the cross-entropy function, thereby approximating the distribution of samples on the low-dimensional Grassmann manifold  $\mathbb{G}(m, d)$  to that on the high-dimensional Grassmann manifold  $\mathbb{G}(m, D)$ . When the similarity between  $X_i$  and  $X_j$  on the high-dimensional manifold is strong, correspondingly, the similarity between  $Y_i$  and  $Y_j$  on the low-dimensional manifold is also strong. By minimizing the cross-entropy function, the attractive force  $\mathcal{F}$  and repulsive force  $\mathcal{H}$  between samples on the low-dimensional manifold can be respectively expressed as:

$$\mathcal{F} = \frac{\partial P \log Q}{\partial Y_i} \quad (9)$$

$$= \frac{-2abd_{proj}^{2(b-1)}(Y_i, Y_j)}{[1 + ad_{proj}^{2b}(Y_i, Y_j)]} P(X_i, X_j) (Y_i Y_i^T - Y_j Y_j^T) Y_i \quad (10)$$

$$\mathcal{H} = \frac{\partial(1 - P) \log(1 - Q)}{\partial Y_i} \quad (11)$$

$$= \frac{2b(1 - P(X_i, X_j)) (Y_i Y_i^T - Y_j Y_j^T) Y_i}{[1 + ad_{proj}^{2b}(Y_i, Y_j)] [\varepsilon + d_{proj}^2(Y_i, Y_j)]} \quad (12)$$

In order to avoid the potential issue arising from the Eq. (12) of  $d_{proj}(Y_i, Y_j) = 0$ , a small positive number  $\varepsilon$  is introduced into the denominator as a regularization term. During the course of experimentation, the value assigned to  $\varepsilon$  is conventionally established at 0.001.

In order to provide a more intuitive representation of the training process of GUMAP, the pseudocode of the GUMAP method is presented in Algorithm 1.

### 3.3. MCFG with GUMAP

Within the context of this research, we undertake feature extraction for samples residing on Grassmann manifold, leveraging the methodology detailed in Section 3.2, namely the GUMAP. Subsequently, a similarity matrix is computed for the extracted pivotal features employing the Grassmann metric, thereby facilitating the attainment of enhanced clustering outcomes.

The transition from (a) to (b) in Fig. 1 illustrates the modeling of the original image-sets and videos onto Grassmann manifold. Specifically, a descriptor  $\mathbb{G}(m, D)$  on Grassmann manifold in (b) is employed to represent the collective of multiple images shown in (a). The sequence of steps from (a) to (b) and subsequently to (e) encapsulates the MCFG introduced in Section 3.1. Within this framework, clustering of the samples on Grassmann manifold in (b) is directly accomplished using a clustering algorithm that leverages the Grassmann metric.

The transition from (c) to (d) and subsequently to (e) in Fig. 1 elucidates the dimensionality reduction of high-dimensional Grassmann samples through the application of GUMAP, as introduced in Section 3.2. This process entails refining the topological configuration of the projected points in (d) by optimizing the interplay of attractive

---

**Algorithm 1** Uniform Manifold Approximation and Projection on Grassmann Manifold
 

---

**Require:** Image-set samples  $\mathcal{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$ ,  $\mathbf{X}_i \in \mathbb{G}(m, D)$ , number of neighbors  $k$ , order of Grassmann manifold  $m$ , target dimension  $d$ , neighborhood radius  $mindist$ , number of iterations  $epochs$ .

**Ensure:** Optimal projection points  $\mathcal{Y} = \{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N\}$ ,  $\mathbf{Y}_i \in \mathbb{G}(m, d)$ .

```

1: function GUMAP( $\mathcal{X}, k, d, m, mindist, epochs$ ):
2:   Construct the pairwise distance matrix using Eq. (1);
3:   Construct the  $k$  nearest neighbor graph  $G(V, E, \mathbf{P})$  using Eqs. (4) to (6);
4:   Randomly initialize the projection points  $\mathcal{Y}$  on the low-dimensional manifold;
5:   Determine the hyperparameters  $a$  and  $b$  through Eq. (8) and  $mindist$ ;
6:    $\alpha = 1.0$ ;
7:   for  $e \in [1, 2, \dots, epochs]$  do:
8:     for  $\mathbf{X}_a, \mathbf{X}_b \in \mathcal{X}$  do:
9:        $\mathbf{Y}_a = \mathbf{Y}_a + \nabla(\log(\mathbf{Q}))(\mathbf{Y}_a, \mathbf{Y}_b)$ ;
10:      Sampling from the  $k$  nearest neighbors of  $\mathbf{Y}_a$  or  $\mathbf{Y}_b$  to  $\mathcal{Y}_{neg}$ ;
11:      for  $\mathbf{Y}_c \in \mathcal{Y}_{neg}$  do:
12:         $\mathbf{Y}_a = \mathbf{Y}_a + \nabla(\log(1 - \mathbf{Q}))(\mathbf{Y}_a, \mathbf{Y}_c)$ ;
13:      end for
14:    end for
15:     $\alpha = \alpha - e/epochs$ ;
16:  end for
17: end function
  
```

---

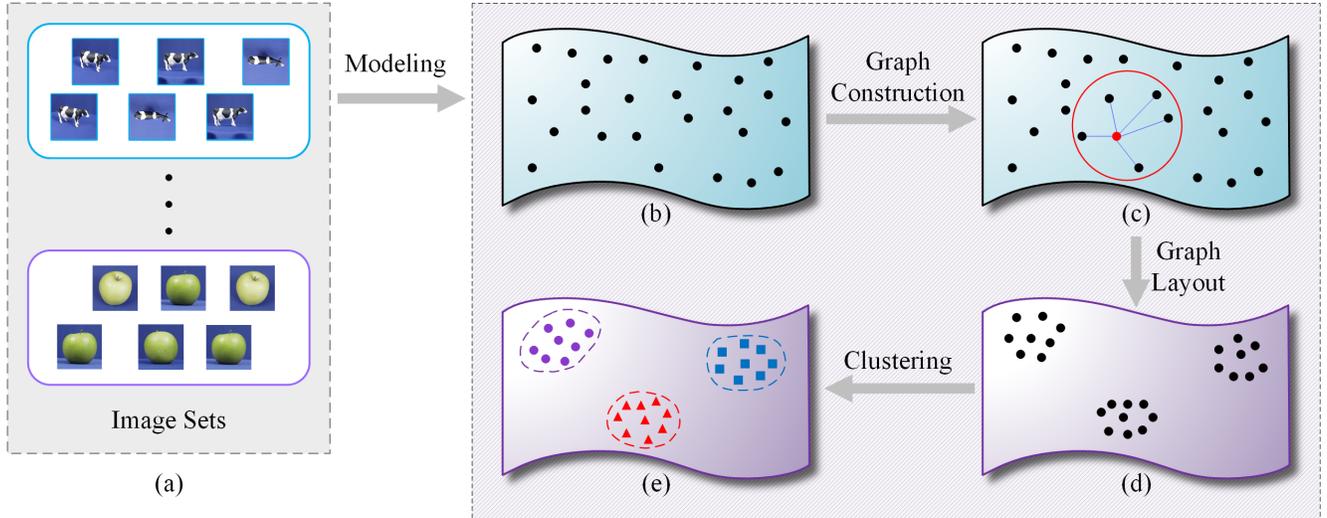


Figure 1: The Fundamental Concept of MCFG with GUMAP: (a) Represents the original image-sets; (b) Illustrates the descriptors of the image-sets on Grassmann manifold; (c) Depicts the topological structure constructed by GUMAP on the high-dimensional Grassmann manifold; (d) Shows the topological structure on the low-dimensional Grassmann manifold after optimization by GUMAP; (e) Displays the clustering results obtained by MCFG with GUMAP.

and repulsive forces among them, thereby aiming to mirror the topological organization of the samples on the high-dimensional Grassmann manifold depicted in (c). Within the context of GUMAP’s optimization, the distances between highly similar samples are progressively minimized, whereas those separating less similar samples are expanded.

In contrast to the clustering procedure spanning from (b)

to (e) within the standard MCFG, the clustering sequence from (d) to (e) in MCFG with GUMAP demonstrates enhanced performance. This superiority stems from the arrangement in (d), where samples exhibiting high similarity are positioned in closer proximity, while those with low similarity are spaced further apart, thereby offering a directive influence on the clustering process of (d)-(e).

In order to provide a more intuitive representation of the training process of MCFG with GUMAP, the pseudocode of the MCFG with GUMAP is presented in Algorithm 2.

## 4. Experiments

### 4.1. Datasets

The pivotal aspect of MCFG resides in leveraging GUMAP to extract pivotal features from the data, thereby augmenting the efficacy of clustering methods. This study examines the impact of GUMAP in the context of image-set clustering tasks, employing 4 datasets: ETH-80[8], Extended Yale B (EYB)[15], UTD-MHAD[7], and Weizmann[11]. Each video or image-set within these datasets is mapped onto the Grassmann manifold  $\mathbb{G}(10, 400)$  and subsequently undergoes dimensionality reduction to  $\mathbb{G}(10, 20)$  through the application of GUMAP. The essential parameter configurations for these datasets are recorded in Table 2.

Fig. 2 illustrates the process of modeling image-sets onto the Grassmann manifold. For each image within the image-set, it is individually resized to  $m \times n$  and subsequently flattened into a vector denoted as  $\mathbf{y}_i$ . The ensemble of all images from the image-set constitutes a matrix  $\mathbf{Y}$ . By performing Singular Value Decomposition on the matrix  $\mathbf{Y}\mathbf{Y}^T$ , the first  $p$  eigenvectors of  $\mathbf{U}$  elucidate a descriptor on the Grassmann manifold  $\mathbb{G}(p, mn)$ .

### 4.2. Evaluations

In this study, the clustering performance of the MCFG is evaluated through 3 metrics: the Adjusted Rand Index (ARI)[6], the Adjusted Mutual Information Score (AMI)[29], and the Homogeneity Score (HMS)[20].

The ARI provides a measure of the consistency between the true labels of the data and the clustering results, offering stronger reliability and interpretability for the clustering outcomes. Its value ranges from  $[-1, 1]$ . When the ARI approaches 0, it indicates that the clustering results are randomly assigned. Conversely, when the ARI approaches 1, it signifies a perfect alignment between the clustering results and the true labels.

The AMI score reflects the correlation between the true labels of the data and the clustering results, with a range of  $[-1, 1]$ . A higher AMI score indicates a better clustering performance.

The HMS measures the homogeneity between the true labels of the data and the clustering results, assessing whether the samples within each cluster belong to the same category. Its value ranges from  $[0, 1]$ , and an HMS closer to 1 indicates a superior clustering effect.

### 4.3. Clustering Performance of MCFG with GUMAP

We conducted clustering analyses on high-dimensional Grassmann manifold samples using both the MCFG and MCFG with GUMAP. The clustering performance was evaluated using the AMI, ARI, and HMS. The AMI, ARI, and HMS results for the clustering outcomes of MCFG and MCFG with GUMAP are visualized in Figs. 3 to 5, respectively. The time costs required by each algorithm are recorded in Table 3.

The application of the MCFG with GUMAP to various clustering methods, such as SNN-DPC, AFP, and AGC, has demonstrated superior performance compared to existing clustering methods on Grassmann manifolds, including CGMKE, Grassmann K-Means, and GLBG. This proves the effectiveness of MCFG with GUMAP in image-set clustering tasks.

The clustering performance of MCFG with GUMAP surpasses that of MCFG alone, a conclusion that is intuitive and effectively demonstrates the auxiliary role of GUMAP in enhancing clustering methods on Grassmann manifold. By extracting key features from high-dimensional manifolds, GUMAP reduces the distance between highly similar samples while increasing the separation between those with low similarity, thereby significantly improving the performance of clustering algorithms on Grassmann manifold.

Based on indices such as AMI, ARI, and HMS for GDBSCAN, GHDBSCAN, and GOPTICS, simply extending these methods to Grassmann manifold through predefined distance matrices is not an optimal choice. However, after applying GUMAP to reduce the dimensions of high-dimensional image-set samples, GDBSCAN, GHDBSCAN, and GOPTICS can achieve satisfactory clustering results. This occurs because DBSCAN, HDBSCAN, and OPTICS all focus on the density distribution of the samples, yet the distribution of image-sets on Grassmann manifold is uniform and dispersed, a characteristic that limits the performance of such algorithms. GUMAP, during the dimensionality reduction process, brings similar samples closer together while increasing the distance between those with low similarity, explaining its more significant enhancement for GDBSCAN, GHDBSCAN, and GOPTICS.

From the results presented in Table 3, it is evident that the MCFG exhibits lower computational time costs for processing key features compared to those associated with corresponding high-dimensional samples. The GUMAP approach significantly reduces the computational complexity and conserves both time and space costs for MCFG operations by extracting  $\mathbb{G}(10, 20)$  embedded within  $\mathbb{G}(10, 400)$ , thereby eliminating redundant information from the data.

### 4.4. Parameters Analysis

In GUMAP, there are 3 critical hyperparameters: the number of neighbors in the graph construction process, the

---

**Algorithm 2** Metric-Based Clustering Framework on Grassmann Manifold with GUMAP

---

**Require:** Image-set samples  $\mathcal{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N\}$ ,  $\mathbf{X}_i \in \mathbb{G}(m, D)$ .

**Ensure:** Clustering labels  $\{t_i\}_{i=1}^N$

- 1: **function** MCFG( $\mathcal{X}$ ):
  - 2:    $\mathcal{Y} = \text{GUMAP}(\mathcal{X}, k = 10, d = 20, m = 10, \text{mindist} = 0.1, \text{epochs} = 500)$ ;
  - 3:   Construct the pairwise distance matrix  $\mathbf{D}$  using Table 1 for  $\mathcal{Y}$ ;
  - 4:   Clustering base on the pairwise distance matrix  $\mathbf{D}$ ;
  - 5: **end function**
- 

Table 2: Summary of Dataset Information.

Names	Samples	Classes	$\mathbb{G}(m, D)$	$\mathbb{G}(m, d)$
<b>ETH-80</b> [8]	80	8	$\mathbb{G}(10, 400)$	$\mathbb{G}(10, 20)$
<b>EYB</b> [15]	252	28	$\mathbb{G}(10, 400)$	$\mathbb{G}(10, 20)$
<b>UTD-MHAD</b> [7]	861	27	$\mathbb{G}(10, 400)$	$\mathbb{G}(10, 20)$
<b>Weizmann</b> [11]	93	10	$\mathbb{G}(10, 400)$	$\mathbb{G}(10, 20)$

subspace orders of the samples, and the target dimensions. In this section, we evaluate the impact of these hyperparameters on model performance using Spectral Clustering on Grassmann Manifold (GSPC) as an example, conducting experiments on datasets such as ETH-80, EYB, UTD-MHAD, and Weizmann.

We evaluated the performance of GSPC across different values of the number of neighbors  $k(k \in [5, 15])$ . GSPC was run 5 times for each value of the number of neighbors, and the mean and standard deviation of the experimental results are shown in Fig. 6. The variation in clustering performance of GSPC with respect to the number of neighbors is minimal. In this study, the number of neighbors for GUMAP was uniformly set to 10.

We evaluated the performance of GSPC across different values of the target dimensions  $d(d \in [10, 30])$ . GSPC was run 5 times for each target dimension, and the mean and standard deviation of the results are shown in Fig. 7. As the target dimensions of GUMAP increases, the clustering performance of GSPC improves. When the target dimensions of GUMAP exceeds 16, the clustering metrics ARI, AMI, and HMS tend to stabilize. In this study, the target dimensions of GUMAP was uniformly set to 20 across all datasets.

We evaluated the performance of GSPC across different values of the subspace orders  $p(p \in [5, 15])$ . GSPC was run 5 times for each subspace order, and the mean and standard deviation of the results are shown in Fig. 8. The effect of the subspace order  $p$  on the clustering performance of GSPC varies across different datasets. Overall, when the subspace order  $p \in [8, 10]$ , the clustering metrics such as ARI, AMI, and HMS for GSPC remain stable. In this study, the subspace order was uniformly set to 10 across all datasets.

## 5. Conclusion

In this paper, we propose the Metric-Based Clustering Framework on Grassmann Manifold (MCFG), which successfully transfers clustering methods from Euclidean space to Grassmann manifold through a predefined similarity matrix based on Grassmann metrics. Additionally, we introduce the Uniform Manifold Approximation and Projection on Grassmann Manifold (GUMAP), a nonlinear dimensionality reduction method on Grassmann manifold that effectively extracts low-dimensional Grassmann manifolds embedded within high-dimensional ones. Preprocessing high-dimensional Grassmann samples with GUMAP can significantly enhance the clustering performance of MCFG. Furthermore, MCFG with GUMAP outperforms existing clustering methods on Grassmann manifold across multiple image-set dataset benchmarks.

While our method has exhibited superior performance in image-set and video clustering tasks, it is not without limitations. For example, GUMAP is capable of iteratively learning the implicit nonlinear mapping relationship between high-dimensional and low-dimensional Grassmann manifolds. As a result, GUMAP is unable to directly project new, unknown samples onto the low-dimensional Grassmann manifold, thereby giving rise to the Out-of-Sample Problem. Developing appropriate methods to address this issue will be a key focus of our future research. Concurrently, we will also concentrate on extending a wider range of clustering methods from Euclidean space to the Grassmann manifold in a more direct and generalized fashion.

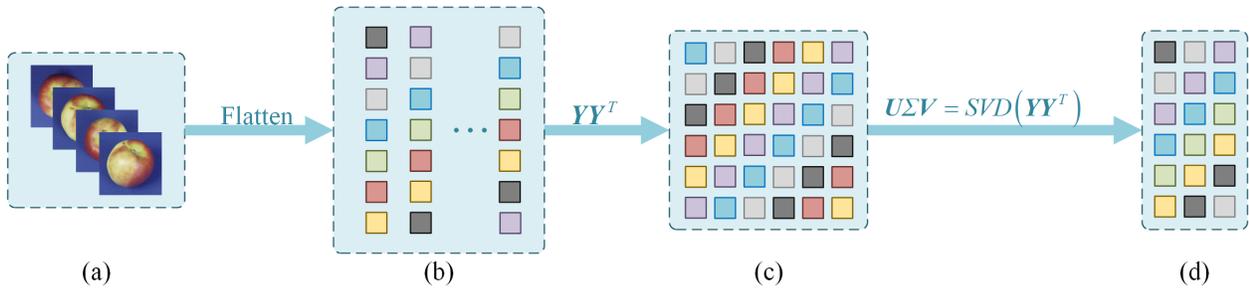


Figure 2: The process of modeling image-sets onto the Grassmann manifold: (a) Several images within an image-set; (b) Flatten each image into a vector  $y_i$ ; (c) Computing the covariance matrix  $YY^T$ ; (d) Modeling through Singular Value Decomposition.

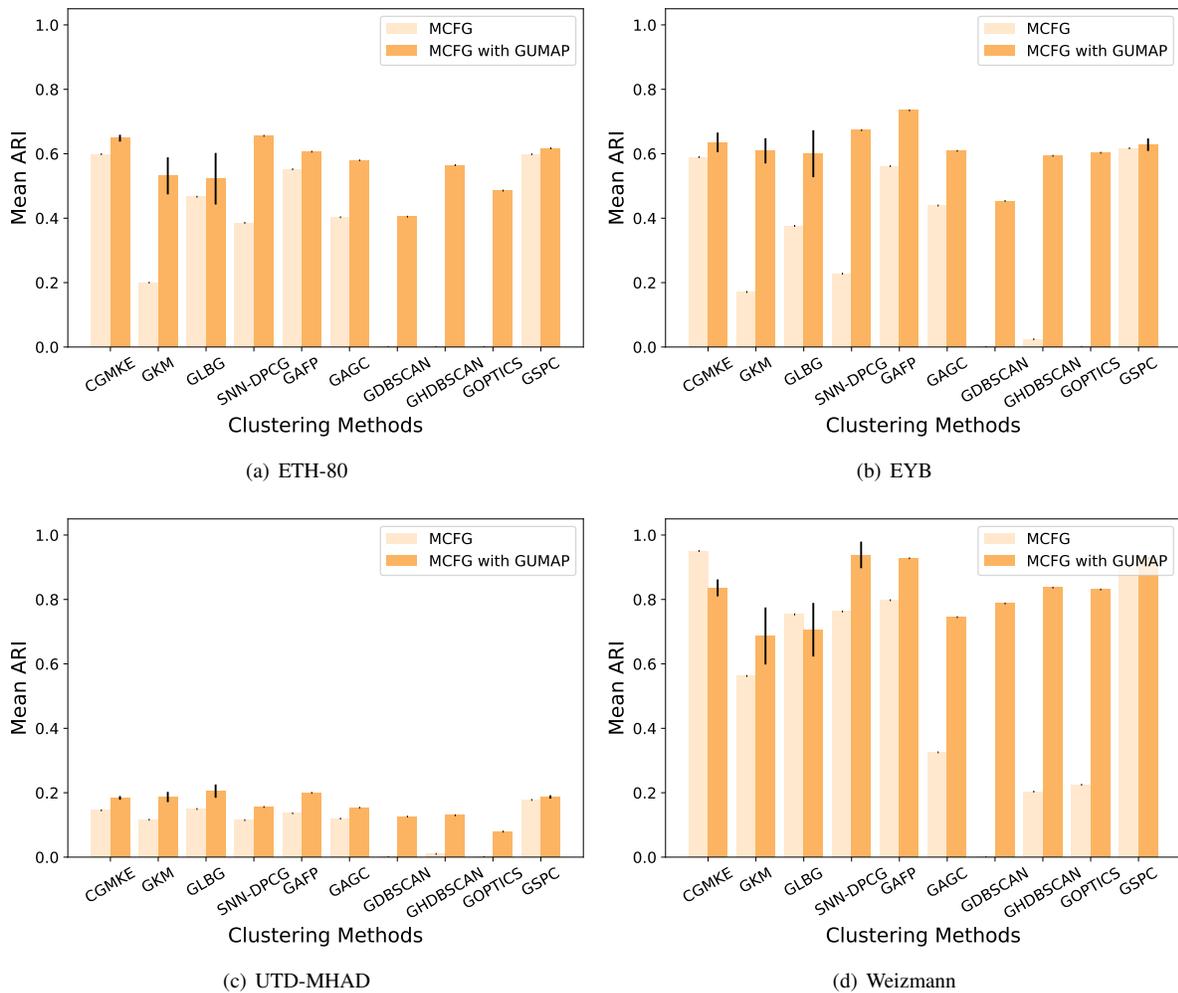
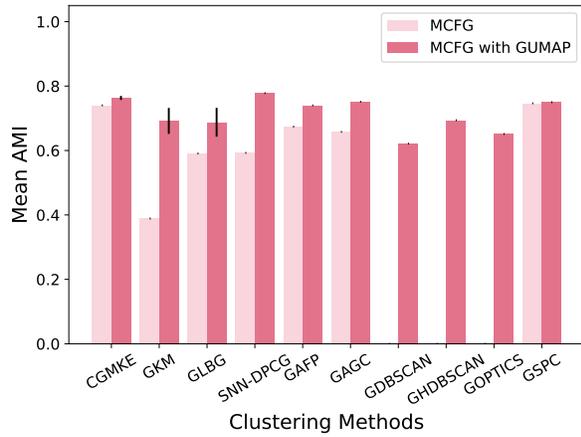
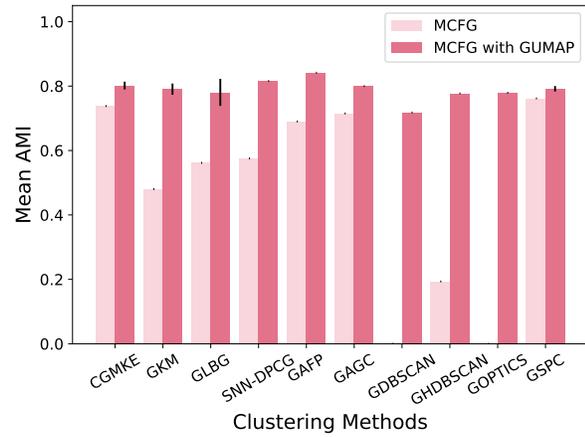


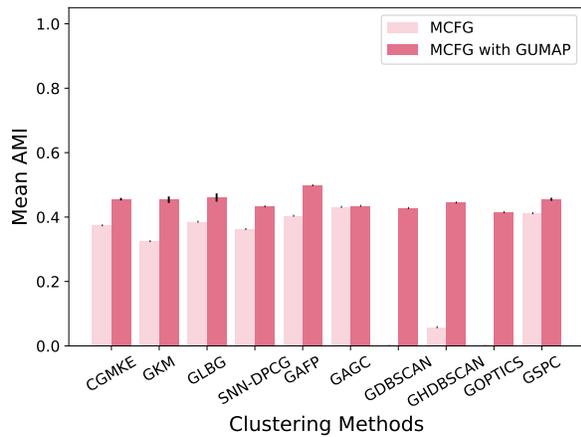
Figure 3: Comparison of Average ARI and Standard Deviation in Clustering Performance across Various Datasets between MCFG and MCFG with GUMAP: Light-colored bars represent the ARI of MCFG, while dark-colored bars indicate the ARI of MCFG with GUMAP. The error bars above the bars depict the standard deviation of multiple clustering outcomes.



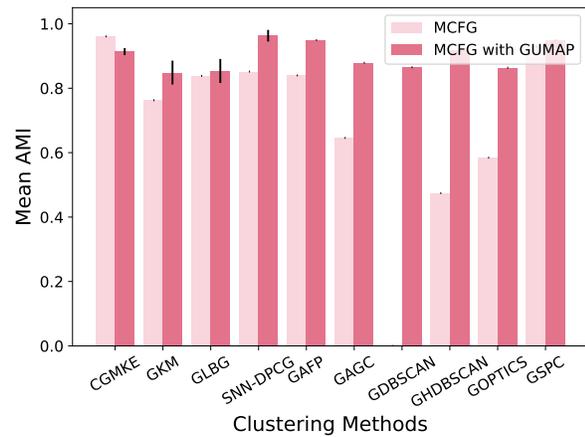
(a) ETH-80



(b) EYB

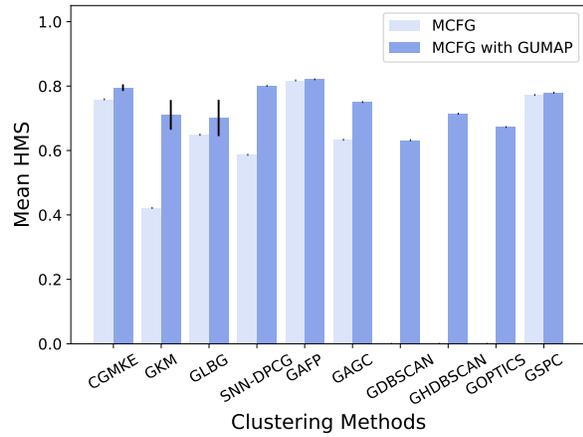


(c) UTD-MHAD

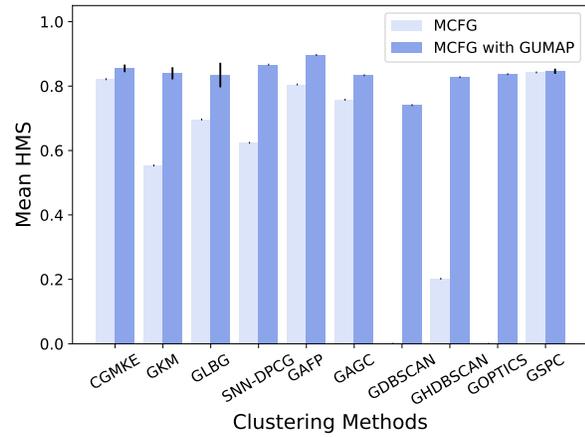


(d) Weizmann

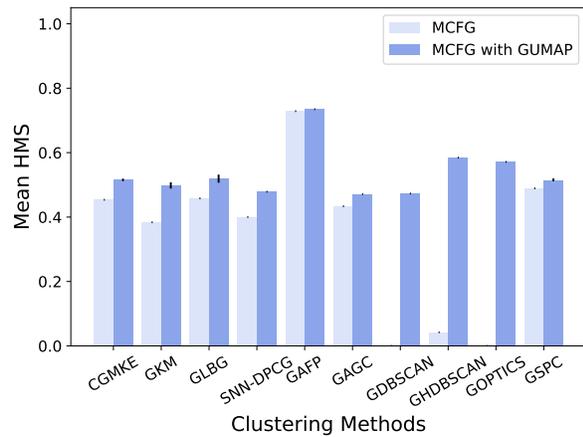
Figure 4: Comparison of Average AMI and Standard Deviation in Clustering Performance across Various Datasets between MCFG and MCFG with GUMAP: Light-colored bars represent the AMI of MCFG, while dark-colored bars indicate the AMI of MCFG with GUMAP. The error bars above the bars depict the standard deviation of multiple clustering outcomes.



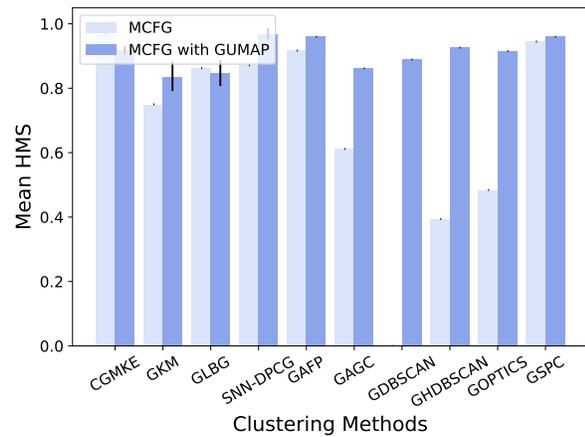
(a) ETH-80



(b) EYB



(c) UTD-MHAD

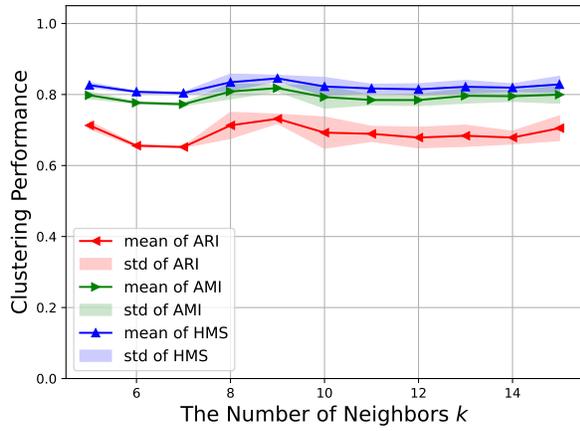


(d) Weizmann

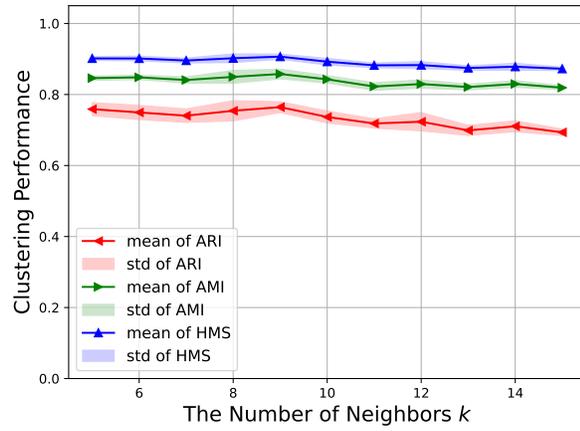
Figure 5: Comparison of Average HMS and Standard Deviation in Clustering Performance across Various Datasets between MCFG and MCFG with GUMAP: Light-colored bars represent the HMS of MCFG, while dark-colored bars indicate the HMS of MCFG with GUMAP. The error bars above the bars depict the standard deviation of multiple clustering outcomes.

Table 3: Comparison of Time Costs (s) for Clustering Across Various Datasets Between MCFG and MCFG with GUMAP. Specifically, GSPC and GUMAP-GSPC represent the respective applications of MCFG and MCFG with GUMAP, within the SPC method.

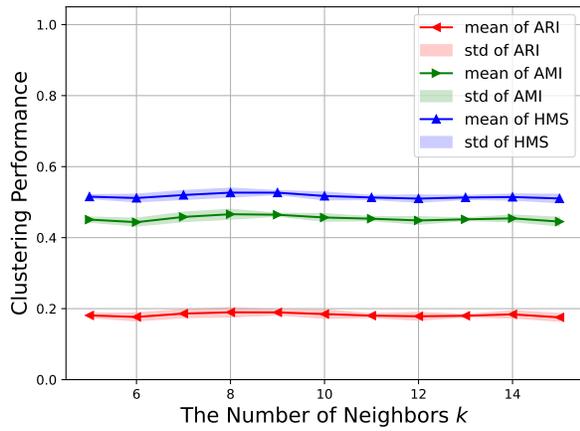
<b>Methods</b>	<b>ETH-80</b>	<b>EYB</b>	<b>UTD-MHAD</b>	<b>Weizmann</b>
<b>CGMKE</b> [24]	0.1186	0.4015	4.1684	0.1471
<b>GUMAP-CGMKE</b>	0.0933	0.1760	1.8067	0.1306
<b>GKM</b> [27]	6.1801	19.4436	88.3841	11.1410
<b>GUMAP-GKM</b>	1.6940	26.8549	35.3539	1.4856
<b>GLBG</b> [25, 27]	0.3218	4.0518	26.9033	0.9866
<b>GUMAP-GLBG</b>	0.1511	2.8923	14.7118	0.2126
<b>SNN-DPCG</b>	0.1810	1.7897	21.0272	0.2427
<b>GUMAP-SNN-DPCG</b>	0.1743	1.7179	20.0815	0.2331
<b>GAFP</b>	0.1606	1.5826	18.6691	0.2143
<b>GUMAP-GAFP</b>	0.1542	1.5215	17.7559	0.2060
<b>GAGC</b>	0.1579	1.5743	18.5356	0.2122
<b>GUMAP-GAGC</b>	0.1516	1.5005	17.5784	0.2037
<b>GDBSCAN</b>	0.1585	1.5701	18.5539	0.2122
<b>GUMAP-GDBSCAN</b>	0.1524	1.4992	17.5919	0.2043
<b>GHDBSCAN</b>	0.1586	1.5777	18.5585	0.2134
<b>GUMAP-GHDBSCAN</b>	0.1525	1.5064	17.6081	0.2050
<b>GOPTICS</b>	0.1891	1.6682	18.9754	0.2475
<b>GUMAP-GOPTICS</b>	0.1826	1.5989	18.0299	0.2395
<b>GSPC</b>	0.0466	0.4441	4.5966	0.0594
<b>GUMAP-GSPC</b>	0.0258	0.2489	2.2340	0.0330



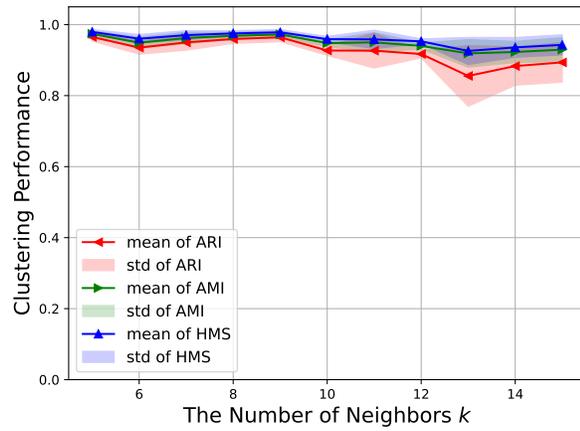
(a) ETH-80



(b) EYB

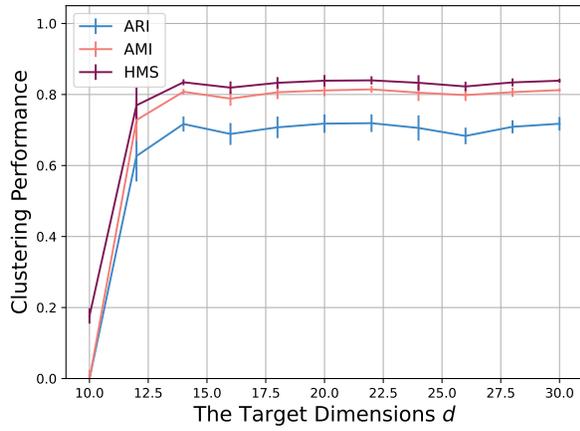


(c) UTD-MHAD

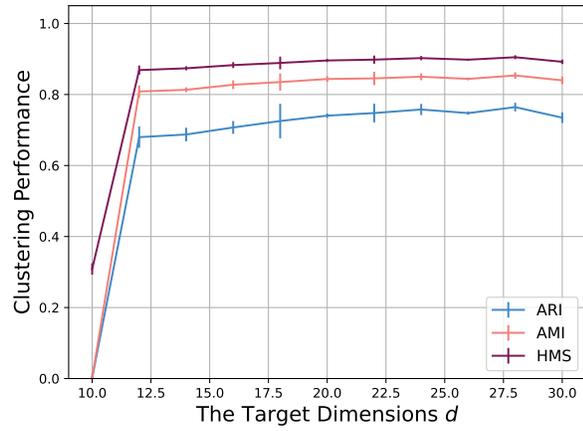


(d) Weizmann

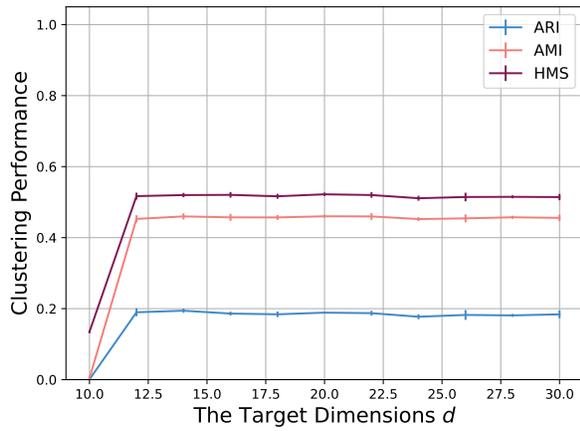
Figure 6: The Mean and Standard Deviation of ARI, AMI, and HMS for GSPC with Different Values of the Number of Neighbors  $k$ .



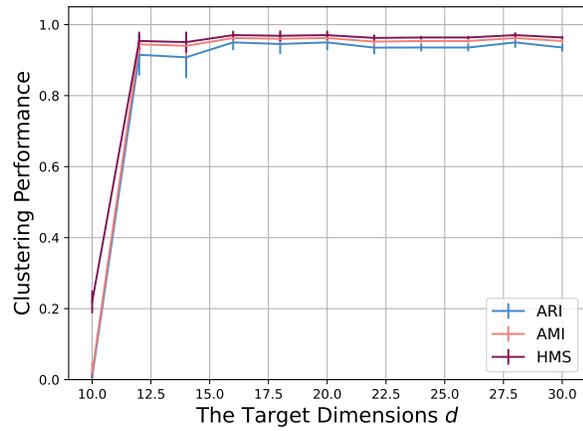
(a) ETH-80



(b) EYB

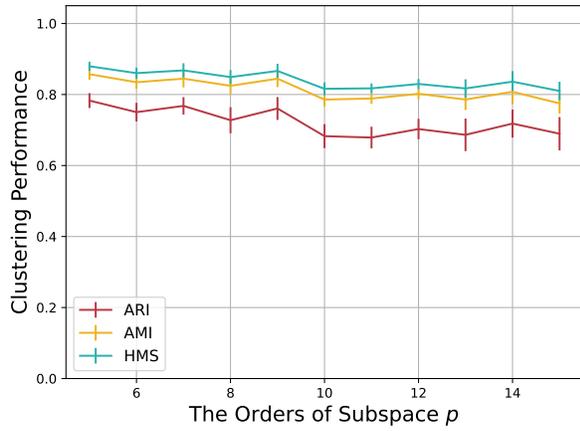


(c) UTD-MHAD

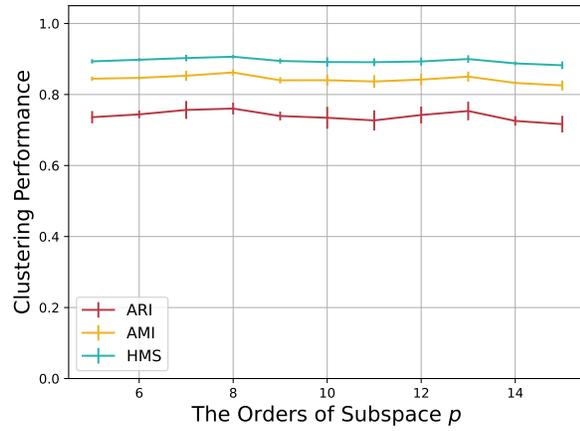


(d) Weizmann

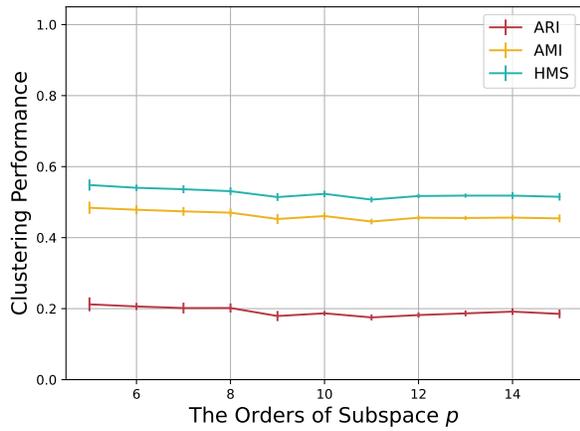
Figure 7: The Mean and Standard Deviation of ARI, AMI, and HMS for GSPC with Different Values of the Target Dimensions  $d$ .



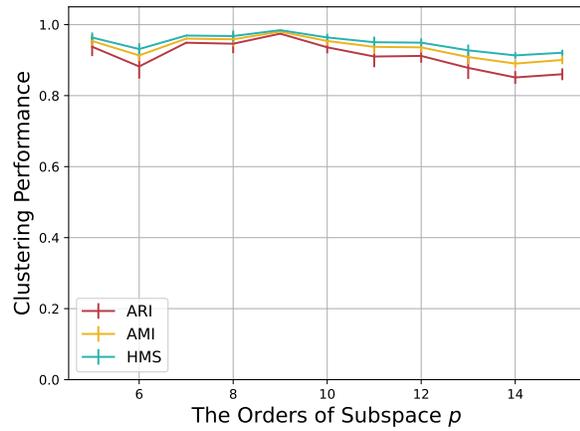
(a) ETH-80



(b) EYB



(c) UTD-MHAD



(d) Weizmann

Figure 8: The Mean and Standard Deviation of ARI, AMI, and HMS for GSPC with Different Values of the Subspace Orders  $p$ .

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgments

This work was supported by Chongqing Municipal Education Commission under grant KJZD-K202100505.

## References

- [1] H. Abdi and L. J. Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010. **1**
- [2] P.-A. Absil, R. Mahony, and R. Sepulchre. Riemannian geometry of grassmann manifolds with a view on algorithmic computation. *Acta Applicandae Mathematica*, 80:199–220, 2004. **2**
- [3] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander. Optics: Ordering points to identify the clustering structure. *ACM Sigmod record*, 28(2):49–60, 1999. **3**
- [4] M. Ay, L. Özbakır, S. Kulluk, B. Gülmez, G. Öztürk, and S. Özer. Fc-kmeans: Fixed-centered k-means algorithm. *Expert Systems with Applications*, 211:118656, 2023. **1**
- [5] R. J. Campello, D. Moulavi, and J. Sander. Density-based clustering based on hierarchical density estimates. In *Pacific-Asia conference on knowledge discovery and data mining*, pages 160–172. Springer, 2013. **3**
- [6] J. E. Chacón and A. I. Rastrojo. Minimum adjusted rand index for two clusterings of a given size. *Advances in Data Analysis and Classification*, 17(1):125–133, 2023. **6**
- [7] C. Chen, R. Jafari, and N. Kehtarnavaz. Utd-mhad: A multi-modal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor. In *2015 IEEE International conference on image processing (ICIP)*, pages 168–172. IEEE, 2015. **6, 7**
- [8] K.-X. Chen, J.-Y. Ren, X.-J. Wu, and J. Kittler. Covariance descriptors on a gaussian manifold and their application to image set classification. *Pattern Recognition*, page 107463, 2020. **6, 7**
- [9] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996. **3**
- [10] B. J. Frey and D. Dueck. Clustering by passing messages between data points. *science*, 315(5814):972–976, 2007. **3**
- [11] L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri. Actions as space-time shapes. *IEEE transactions on pattern analysis and machine intelligence*, 29(12):2247–2253, 2007. **6, 7**
- [12] J. Hamm and D. D. Lee. Grassmann discriminant analysis: a unifying view on subspace-based learning. In *Proceedings of the 25th international conference on Machine learning*, pages 376–383, 2008. **2**
- [13] Y. Hozumi, R. Wang, C. Yin, and G.-W. Wei. Umap-assisted k-means clustering of large-scale sars-cov-2 mutation datasets. *Computers in biology and medicine*, 131:104264, 2021. **1**
- [14] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622:178–210, 2023. **1**
- [15] K.-C. Lee, J. Ho, and D. J. Kriegman. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on pattern analysis and machine intelligence*, 27(5):684–698, 2005. **6, 7**
- [16] H. Liu, J. Yang, M. Ye, S. C. James, Z. Tang, J. Dong, and T. Xing. Using t-distributed stochastic neighbor embedding (t-sne) for cluster analysis and spatial zone delineation of groundwater geochemistry data. *Journal of Hydrology*, 597:126146, 2021. **1**
- [17] R. Liu, H. Wang, and X. Yu. Shared-nearest-neighbor-based clustering by fast search and find of density peaks. *information sciences*, 450:200–226, 2018. **3**
- [18] L. McInnes, J. Healy, and J. Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018. **1, 2**
- [19] A. Pletl, M. Fernandes, N. Thomas, A. P. Rossi, and B. Elser. Spectral clustering of crism datasets in jezero crater using umap and k-means. *Remote Sensing*, 15(4):939, 2023. **1**
- [20] A. Rosenberg and J. Hirschberg. V-measure: A conditional entropy-based external cluster evaluation measure. In *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, pages 410–420, 2007. **6**
- [21] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326, 2000. **2**
- [22] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O. P. Patel, A. Tiwari, M. J. Er, W. Ding, and C.-T. Lin. A review of clustering techniques and developments. *Neurocomputing*, 267:664–681, 2017. **1, 3**
- [23] Shi. Multiclass spectral clustering. In *Proceedings ninth IEEE international conference on computer vision*, pages 313–319. IEEE, 2003. **1, 3**
- [24] S. Shirazi, M. T. Harandi, C. Sanderson, A. Alavi, and B. C. Lovell. Clustering on grassmann manifolds via kernel embedding with application to action analysis. In *2012 19th IEEE international conference on image processing*, pages 781–784. IEEE, 2012. **2, 11**
- [25] S. J. Stiverson. An adaptation of k-means-type algorithms to the grassmann manifold. Master’s thesis, Colorado State University, 2019. **2, 11**
- [26] E. K. Tokuda, C. H. Comin, and L. d. F. Costa. Revisiting agglomerative clustering. *Physica A: Statistical mechanics and its applications*, 585:126433, 2022. **3**

- [27] P. Turaga, A. Veeraraghavan, A. Srivastava, and R. Chellappa. Statistical computations on grassmann and stiefel manifolds for image and video-based recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(11):2273–2286, 2011. [2](#), [11](#)
- [28] L. Van der Maaten and G. Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008. [1](#)
- [29] N. X. Vinh, J. Epps, and J. Bailey. Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *Journal of Machine Learning Research*, 11(2):2837–2854, 2010. [6](#)
- [30] D. Yang, V. Wei, Z. Jin, Z. Yang, and X. Chen. A umap-based clustering method for multi-scale damage analysis of laminates. *Applied Mathematical Modelling*, 111:78–93, 2022. [1](#)