electronics

MDPI

Article

Enhanced Multi-View Low-Rank Graph Optimization for Dimensionality Reduction

Haohao Li [1,*] and Huibing Wang [2]

[1] Department of Mathematics, Zhejiang Sci-Tech University, Hangzhou 310018, China
[2] College of Information and Science Technology, Dalian Maritime University, Dalian 116021, China; huibing.wang@dlmu.edu.cn
* Correspondence: hhl820@zstu.edu.cn

Abstract: In the last decade, graph embedding-based dimensionality reduction for multi-view data has been extensively studied. However, constructing a high-quality graph for dimensionality reduction is still a significant challenge. Herein, we propose a new algorithm, named multi-view low-rank graph optimization for dimensionality reduction (MvLRGO), which integrates graph optimization with dimensionality reduction into one objective function in order to simultaneously determine the optimal subspace and graph. The subspace learning of each view is conducted independently by the general graph embedding framework. For graph construction, we exploit low-rank representation (LRR) to obtain reconstruction relationships as the affinity weight of the graph. Subsequently, the learned graph of each view is further optimized throughout the learning process to obtain the ideal assignment of relations. Moreover, to integrate information from multiple views, MvLRGO regularizes each of the view-specific optimal graphs such that they align with one another. Benefiting from this term, MvLRGO can achieve flexible multi-view communication without constraining the subspaces of all views to be the same. Various experimental results obtained with different datasets show that the proposed method outperforms many state-of-the-art multi-view and single-view dimensionality reduction algorithms.

Keywords: multi-view learning; low-rank representation; graph embedding; graph optimization

check for updates

Citation: Li, H.; Wang, H. Enhanced Multi-View Low-Rank Graph Optimization for Dimensionality Reduction. *Electronics* **2024**, *13*, 2421. https://doi.org/10.3390/electronics13122421

Academic Editor: Daniel Gutiérrez Reina

Received: 14 May 2024
Revised: 3 June 2024
Accepted: 13 June 2024
Published: 20 June 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the development of technology, numerous algorithms and devices have been proposed to characterize samples from different views [1–3]. Multi-view features capture various attributes of the same sample from multiple views. For example, in an image analysis domain, one color image can be described by feature vectors extracted by multiple feature extractors such as local binary pattern (LBP) [4], histogram of oriented gradient (HOG) [5], and Gist [6]. While features extracted by different extractors may contain different types of attribute information, they all describe the same image, indicating some common intrinsic relationships among them. Unlike single-view features, multi-view features contain more information. Therefore, fully exploiting compatible and complementary information across multiple views is crucial and challenging for improving the performance of dimensionality reduction algorithms [7–10].

In most research fields, such as image retrieval, image classification, and textual classification, most of the features utilized within these fields are typically high-dimensional, and dealing with these high-dimensional features leads to significant consumption of computational resources. Therefore, it is crucial for researchers to construct meaningful dimensionality reduction algorithms to solve the optimal subspace while keeping as much of the important information from the original data as possible. To address this issue, researchers have proposed numerous efficient dimensionality reduction methods. Some linear dimensionality reduction algorithms assume that the underlying structure of the

original data can be discovered through linear transformations, and they seek to learn an optimal projection matrix to project the original high-dimensional data into the low-dimensional subspace, such as principal component analysis (PCA) [11], linear discriminant analysis (LDA) [12], locality-preserving projections (LPPs) [13], neighborhood-preserving embedding (NPE) [14], and sparsity-preserving projection [15]. Some manifold hypothesis-based nonlinear algorithms consider that high-dimensional data are distributed around a submanifold of the original space. These methods aim to discover the manifold structure concealed in the data for dimensionality reduction. Locally linear embedding (LLE) [16], Laplacian Eigenmaps (LEs) [17], and Isomap [18] are representative nonlinear algorithms. One common limitation of these traditional dimensionality reduction methods is that the construction processes of the graph utilized for dimensionality reduction are irrelevant to the given task. They construct the graph based on original data that contain noise and redundant information, which usually leads to the constructed graphs being suboptimal. To address this issue, some algorithms [19–22] incorporate graph optimization and dimensionality reduction into one objective function, which learns the optimal graph and the low-dimensional representation simultaneously. Although these algorithms are available and suitable for dealing with a variety of high-dimensional data, most of them can only utilize a feature of the sample from a single view and cannot make full use of multi-view data to improve performance.

In the past decade, multi-view learning has undergone remarkable development in various fields [3,23], such as clustering [24,25], metric learning [26–28], and dimensionality reduction [7,29,30]. Kumar et al. [31] extended the traditional spectral embedding technique into the multi-view setting using a co-training framework, and they proposed co-regularized multi-view spectral clustering. Multi-view canonical correlation analysis (MCCA) [32] is an extension of canonical correlation analysis (CCA) [33], which constructs a common subspace for all views simultaneously. Multi-view spectral embedding (MSE) [34] constructs Laplace matrices for different view features and computes a shared spectral embedding, aiming to preserve information from each viewpoint as effectively as possible. Multi-view dimensionality co-reduction (McDR) [35] constructs the dimensionality reduction objective for each view and maximizes the dependencies across the low-dimensional representation of multiple views via the Hilbert–Schmidt independence criterion (HSIC). Co-regularized multi-view sparse reconstruction embedding (CMSRE) [29] explores the sparse reconstruction correlation between samples in multiple views based on sparse representation and the co-regularization technique to learn the sparse structure of samples on manifolds across all views and achieve low-dimensional embedding of high-dimensional samples.

Most existing multi-view dimensionality reduction algorithms have achieved good results when dealing with high-dimensional multi-view data. However, these proposed algorithms usually suffer from the following limitations. First, they utilize the graphs constructed with original high-dimensional data that are redundant and always contain noise, which results in an unreliable and inaccurate graph. A suboptimal graph may have negative effects on performance in terms of dimensionality reduction. Second, most methods project different view features into one common low-dimensional space or minimize the difference between low-dimensional representations of different views, which is inappropriate if the dimensionalities of features from different views are unbalanced. Moreover, since different view features are usually located in different spaces and have different statistical properties, directly projecting all view features into one common subspace is unreasonable.

To address the aforementioned problems, we propose a novel multi-view dimensionality reduction algorithm named multi-view low-rank graph optimization for dimensionality reduction (MvLRGO) in this paper. Based on the graph embedding framework, MvLRGO constructs an affinity graph for each view via low-rank representation (LRR), which facilitates the exploration of meaningful structures in the data [36] and further optimizes the low-rank graph within the whole learning process. We consider that, although there are differences between the representations of different views, they all describe one sample.

MvLRGO adopts a graph-based reasonable hypothesis [37,38] where the affinity graphs of each view describing the similarity between samples should be relatively close to each other. Therefore, for graph learning, MvLRGO further regularizes the low-rank graphs of each view so that they are similar to each other based on this hypothesis, meaning the compatible and complementary information of multi-view features can be efficiently incorporated for graph learning. In general, by introducing graph optimization into the final objective function, MvLRGO not only utilizes the optimal graph to learn the subspace for each view but also effectively integrates compatible and complementary information for dimensionality reduction without forcibly projecting the high-dimensional feature of different views into one common subspace. Finally, the alternating direction minimization (ADM) algorithm is adopted to solve the objective of MvLRGO. We summarize the contributions of MvLRGO as follows:

- We propose a novel algorithm, termed multi-view low-rank graph optimization for dimensionality reduction, which can deal with high-dimensional multi-view data by effectively exploring the underlying structure of samples and simultaneously seeking the optimal subspace for each view.
- MvLRGO adopts low-rank representation to obtain the affinity graph for each view, thus effectively extracting the subspace structure. The constructed graph utilized for dimensionality reduction will be further optimized in the whole learning process, which means that MvLRGO utilizes low-dimensional representations that contain less noise and redundant information for graph construction. Therefore, MvLRGO can construct a low-dimensional representation for each view based on the optimal graph.
- We assume that the affinity relations obtained from different views should be the same, and the graph of each view is regularized so that they are close to each other in the objective function; this is reasonable because these graphs all reflect the affinity relationship of the samples. Through this regularization, MvLRGO can effectively integrate complementary and compatible information from multi-view data.

The rest of this paper is organized as follows: In Section 2, we review some related dimensionality reduction algorithms. In Section 3, we provide the details of the proposed algorithm, MvLRGO. Then, we demonstrate the experimental results for various datasets in Section 4. Section 5 provides our conclusions.

2. Related Work

In this section, we review the first single-view dimensionality reduction method based on graph optimization, named graph-optimized locality-preserving projections (GoLPP) [19]. GoLPP integrates graph construction with a dimensionality reduction process into a unified framework, which results in an optimal graph for dimensionality reduction rather than a predefined one. Then, we introduce a classical multi-view dimensionality reduction method, termed multi-view spectral embedding (MSE) [34], which fuses the graph of each view with different weights to simultaneously construct sufficiently smooth embedding over all views.

2.1. Graph-Optimized Locality-Preserving Projections

GoLPP is a graph-based dimensionality reduction method that introduces the graph optimization procedure into the objective of LPP [13] to jointly learn the optimal graph and projection matrix. Let $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N] \in \mathbb{R}^{D \times N}$ denotes N samples in D-dimensional

space. GoLPP aims to simultaneously learn the optimal graph and projection matrix using the following objective function:

$$\begin{aligned} &\min_{\boldsymbol{P},\boldsymbol{S}} \frac{\sum_{i,j=1}^{N} \left\|\boldsymbol{P}\boldsymbol{x}_i - \boldsymbol{P}\boldsymbol{x}_j\right\|^2 s_{ij}}{\sum_{i=1}^{N} \left\|\boldsymbol{P}\boldsymbol{x}_i\right\|^2} + \eta \sum_{i,j=1}^{N} s_{ij} ln s_{ij} \\ &s.t. \quad \sum_{j=1}^{N} s_{ij} = 1, \quad s_{ij} \geq 0, \end{aligned} \tag{1}$$

where $\boldsymbol{P} \in \mathbb{R}^{d \times D}$ is the projection matrix, $\boldsymbol{S} = \left(s_{ij}\right)_{n \times n}$ is the affinity weight matrix, and $\eta > 0$ is a trade-off parameter. The optimal solution of this objective function can be solved via an alternative iteration scheme, in which the affinity weight matrix can be updated during each iteration until the optimal one is reached. For GoLPP, it was previously verified that combining graph optimization with dimensionality reduction can achieve optimal affinity. Moreover, the authors argued that GoLPP returns superior results to traditional graph-based dimensionality reduction methods.

2.2. Multi-View Spectral Embedding

MSE [34] is a classical multi-view dimensionality reduction algorithm that has attracted significant widespread attention. MSE constructs a common low-dimensional embedding representation by fusing the graph affinity for the features of all views based on the complementary principle. Given a multi-view dataset with N samples and m views, $\left\{\boldsymbol{X}^v \in \mathbb{R}^{D^v \times N}\right\}_{v=1}^{m}$, MSE first constructs the affinity matrix for each view. Then, it obtains the low-dimensional embedding representation via graph embedding for each view. Finally, MSE unifies these low-dimensional embedding representations of different views using global coordinate alignment. The objective of MSE is formulated as follows:

$$\begin{aligned} &\max_{\boldsymbol{\alpha},\boldsymbol{Y}} \sum_{v=1}^{m} \alpha_v^r Tr\left\{(\boldsymbol{Y})^T \boldsymbol{L}^v \boldsymbol{Y}\right\} \\ &s.t. \quad (\boldsymbol{Y})^T \boldsymbol{Y} = \boldsymbol{I}, \quad \sum_{v=1}^{m} \alpha_v = 1, \alpha_v \geq 0 \end{aligned} \tag{2}$$

where $\boldsymbol{L}^v$ is a Laplacian matrix of $\boldsymbol{X}^v$. $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_m]$ is a set of non-negative weights that can be learned based on the importance of the views, and $r > 1$. Moreover, $\boldsymbol{Y}$ is the final spectral embedding representation. MSE effectively exploits the complementary information of multiple views to construct the common low-dimensional representation $\boldsymbol{Y}$ for multi-view features.

3. Multi-View Low-Rank Graph Optimization for Dimensionality Reduction

This section presents the details of our proposed MvLRGO, which integrates dimensionality reduction and graph learning into one unified objective function. Considering a multi-view dataset comprising N samples, each with m views $\left\{\boldsymbol{X}^v \in \mathbb{R}^{D^v \times N}\right\}_{v=1}^{m}$ (where $\boldsymbol{X}^v \in \mathbb{R}^{D^v \times N}$ is the data matrix containing features from the vth view of the sample with each feature represented as a column, and D^v is the dimensionality of the high-dimensional features), MvLRGO endeavors to simultaneously construct an optimal affinity graph and the projection matrices $\left\{\boldsymbol{P}^v \in \mathbb{R}^{D^v \times d^v}\right\}_{v=1}^{m}$ in order to project original high-dimensional features into the low-dimensional subspace for each view.

Leveraging the aforementioned statements, MvLRGO formulates an objective to simultaneously learn the projection matrix and the affinity graph for each view. The projection matrix learning for each view is based on a general graph embedding framework [39]. In the graph learning phase, the affinity graph of each view is optimized, utilizing both low-dimensional representations and the graphs of other views. This enables MvLRGO to

effectively utilize information from multi-view features when constructing an affinity graph for each view. Finally, an iterative optimization scheme is employed to find the solution.

3.1. The Construction of MvLRGO

This section introduces the details of the approach for jointly learning the projection matrix and the affinity graph for each view. MvLRGO employs an LRR scheme to identify the reconstruction relation as the affinity, which aims to construct a low-rank self-representation to capture the inherent correlations between features. We can formulate a single-view graph optimization objective for each single view's dimensionality reduction, as follows:

$$\begin{aligned} &\min_{\boldsymbol{P}^v, \boldsymbol{S}^v} \quad \sum_{i=1}^{N} \left\| (\boldsymbol{P}^v)^T (\boldsymbol{x}_i^v - \boldsymbol{X}^v \boldsymbol{s}_i^v) \right\|_2^2 + \eta \|\boldsymbol{S}^v\|_* \\ &s.t. \quad s_{ij}^v \geq 0, \quad \sum_{j=1}^{N} s_{ij}^v = 1, \quad (\boldsymbol{P}^v)^T \boldsymbol{X}^v (\boldsymbol{X}^v)^T \boldsymbol{P}^v = \boldsymbol{I}, \end{aligned} \tag{3}$$

where $\boldsymbol{x}_i^v \in \mathbb{R}^{D^v \times 1}$ is the vth view feature of ith sample, $\boldsymbol{s}_i^v = [s_{i1}, s_{i2}, \ldots, s_{iN}]^T$ is an N-dimensional vector in which the element s_{ij} reflects the contribution of $\boldsymbol{x}_j^v$ to reconstructing $\boldsymbol{x}_i^v$, $\boldsymbol{S}^v = \left[\boldsymbol{s}_1^v, \boldsymbol{s}_2^v, \ldots, \boldsymbol{s}_N^v\right]^T$ is the low-rank reconstructive weight matrix of the vth view, and $\boldsymbol{P}^v \in \mathbb{R}^{D^v \times d^v}$ is the projection matrix which projects the original D^v dimensional features into a d^v dimensional subspace ($d^v < D^v$). The second term $\|\boldsymbol{S}^v\|_*$ is the nuclear norm which is computed as $\|\boldsymbol{S}^v\|_* = \sum_i^N \lambda_i^v$, where λ_i^v is the ith singular value of $\boldsymbol{S}^v$, and $\eta > 0$ is a trade-off parameter. The constraint $(\boldsymbol{P}^v)^T \boldsymbol{X}^v (\boldsymbol{X}^v)^T \boldsymbol{P}^v = \boldsymbol{I}$ prevents the trivial solution in which $\boldsymbol{I}$ denotes the identity matrix.

Furthermore, we add a non-negative constraint on the self-representation reconstruction weights in Equation (3), which has two beneficial properties: (1) each data point is contained within the convex hull of its neighbors; (2) the learned non-negative reconstruction weights offer a direct measure of the similarity between samples. Additionally, many researchers have shown that applying a non-negative constraint to affinity weights improves the interpretability of data representation and graph construction [40–42].

To jointly perform dimensionality reduction for all features of all views, we aggregate the objective function Equation (3) of each view, as follows:

$$\begin{aligned} &\min_{\boldsymbol{P}^v, \boldsymbol{S}^v} \quad \sum_{v=1}^{m} \sum_{i=1}^{N} \left\| (\boldsymbol{P}^v)^T (\boldsymbol{x}_i^v - \boldsymbol{X}^v \boldsymbol{s}_i^v) \right\|_2^2 + \sum_{v=1}^{m} \eta \|\boldsymbol{S}^v\|_* \\ &s.t. \quad s_{ij}^v \geq 0, \quad \sum_{j=1}^{N} s_{ij}^v = 1, \quad (\boldsymbol{P}^v)^T \boldsymbol{X}^v (\boldsymbol{X}^v)^T \boldsymbol{P}^v = \boldsymbol{I}, \quad v = 1, 2, \ldots, m. \end{aligned} \tag{4}$$

Equation (4) simply incorporates the dimensionality reduction of multiple views. Although the process of Equation (4) can simultaneously learn the low-rank reconstruction relationships and the low-dimensional subspace that maintains the relationships within each view, it does not take into account the feature information of samples across multiple views simultaneously.

To jointly exploit the information in multi-view features for dimensionality reduction, MvLRGO aligns the learned graph of each view with the others. Since the features of samples from different views reflect the various characteristics of the samples, the representations of the features differ across views. Therefore, minimizing the difference between the low-dimensional representations from different views is not a reasonable objective. Nevertheless, since the features from different views describe one common sample, the affinity relationships derived from the features should be relatively similar. Based on this consideration, we propose a hypothesis: the optimal low-rank self-representation matrices of features from each view are similar to each other. According to this hypothesis, we can achieve multi-view dimensionality reduction without projecting the original features from

different views into the same subspace [29,43]. We adopt the Frobenius norm to measure the disagreement between the affinity weight matrices of two views $\boldsymbol{S}^v$ and $\boldsymbol{S}^u$:

$$\mathcal{D}(\boldsymbol{S}^v, \boldsymbol{S}^u) = \|\boldsymbol{S}^v - \boldsymbol{S}^u\|_F^2. \tag{5}$$

Minimizing the difference $\mathcal{D}(\boldsymbol{S}^v, \boldsymbol{S}^u)$ can regularize two affinity weight matrices $\boldsymbol{S}^v$ and $\boldsymbol{S}^u$, making them similar, which is consistent with our proposed hypotheses. Accordingly, we introduce the pairwise disagreement Equation (5) of the learned self-representation matrix across all views into the dimensionality reduction objective Equation (4) to obtain the final objective:

$$\begin{aligned} \min_{\boldsymbol{P}^v, \boldsymbol{S}^v} \quad & \sum_{v=1}^{m} \sum_{i=1}^{N} \left\| (\boldsymbol{P}^v)^T (\boldsymbol{x}_i^v - \boldsymbol{X}^v \boldsymbol{s}_i^v) \right\|_2^2 + \sum_{v=1}^{m} \eta \|\boldsymbol{S}^v\|_* + \sum_{v,u=1}^{m} \gamma D(\boldsymbol{S}^v, \boldsymbol{S}^u) \\ s.t. \quad & s_{ij}^v \geq 0, \quad \sum_{j=1}^{N} s_{ij}^v = 1, \quad (\boldsymbol{P}^v)^T \boldsymbol{X}^v (\boldsymbol{X}^v)^T \boldsymbol{P}^v = \boldsymbol{I}, \quad v, u = 1, 2, \ldots, m, \end{aligned} \tag{6}$$

where $\eta, \gamma > 0$ are two trade-off parameters.

In the projection learning stage, the first term within each view is independently utilized to obtain the optimal subspace in which the learned low-rank reconstruction relationship can be preserved as much as possible. In the graph-learning stage, all the terms in Equation (6) are exploited. The first two terms aim to construct the optimal reconstruction relationship with low-rank constraint via high-quality low-dimensional representation for each independent view, and the third term aims to minimize the pairwise difference in the graph across all views. Therefore, the optimal graph of each view is learned by both the self-representation of the high-quality low-dimensional representation of features and the minimization of the disagreement between other views. Through the proposed objective function Equation (6), in the process of learning, the information from all other views can be effectively utilized, meaning multi-view data can be handled effectively and dimensionality reduction of sample features can be achieved across multiple views. The following section details the process of solving the MvLRGO algorithm.

3.2. The Optimization of MvLRGO

In this section, we introduce the details of the optimization of the proposed MvLRGO. It is clear that the final objective function, Equation (6), is non-convex for all variables; thus, directly solving it is difficult. To facilitate this problem, the augmented Lagrangian multiplier with alternating direction minimizing (ALM-ADM) strategy is adopted. The low-dimensional subspace and graph for each view are alternately optimized within the whole learning procedure. First, the low-rank graph $\boldsymbol{S}^v$ of each view is initialized via the LRR algorithm [36]. Then, the projection matrix $\boldsymbol{P}^v$ of each view can be obtained via a generalized eigenvalue problem based on the graph.

To adopt the ADM strategy in solving our problem, we need to make our objective function separable. Therefore, we first introduce one additional variable $\boldsymbol{Z}^v$ for each view to make Equation (6) separable for optimization, as follows:

$$\begin{aligned} \min_{\boldsymbol{P}^v, \boldsymbol{S}^v} \quad & \sum_{v=1}^{m} \sum_{i=1}^{N} \left\| (\boldsymbol{P}^v)^T (\boldsymbol{x}_i^v - \boldsymbol{X}^v \boldsymbol{s}_i^v) \right\|_2^2 + \sum_{v=1}^{m} \eta \|\boldsymbol{Z}^v\|_* + \sum_{v,u=1}^{m} \gamma \mathcal{D}(\boldsymbol{S}^v, \boldsymbol{S}^u) \\ s.t. \quad & s_{ij}^v \geq 0, \quad \sum_{j=1}^{N} s_{ij}^v = 1, \\ & (\boldsymbol{P}^v)^T \boldsymbol{X}^v (\boldsymbol{X}^v)^T \boldsymbol{P}^v = \boldsymbol{I}, \quad \boldsymbol{S}^v = \boldsymbol{Z}^v, \quad v, u = 1, 2, \ldots, m. \end{aligned} \tag{7}$$

Then, we formulate the Lagrangian function for the objective function in Equation (7), as follows:

$$\begin{aligned}\mathcal{L}(\boldsymbol{P}^v, \boldsymbol{S}^v, \boldsymbol{Z}^v, \boldsymbol{J}^v) = &\sum_{v=1}^{m}\sum_{i=1}^{N}\left\|(\boldsymbol{P}^v)^T(\boldsymbol{x}_i^v - \boldsymbol{X}^v\boldsymbol{s}_i^v)\right\|_2^2 + \sum_{v=1}^{m}\eta\|\boldsymbol{Z}^v\|_* \\ &+ \sum_{v,u=1}^{m}\gamma\mathcal{D}(\boldsymbol{S}^v, \boldsymbol{S}^u) + \sum_{v=1}^{m}\mathcal{G}(\boldsymbol{J}^v, \boldsymbol{S}^v - \boldsymbol{Z}^v) \\ s.t.\quad & s_{ij}^v \geq 0, \quad \sum_{j=1}^{N} s_{ij}^v = 1, \\ & (\boldsymbol{P}^v)^T\boldsymbol{X}^v(\boldsymbol{X}^v)^T\boldsymbol{P}^v = \boldsymbol{I}, \quad v, u = 1, 2, \ldots, m,\end{aligned} \tag{8}$$

where $\mathcal{G}(\cdot,\cdot)$ is a matrix function, which is defined as $\boldsymbol{\mathcal{G}}(\boldsymbol{A}, \boldsymbol{B}) = \mu/2\|\boldsymbol{B}\|_F^2 + Tr(\boldsymbol{AB})$, $\mu > 0$, and $\boldsymbol{J}^v$ is the Lagrangian multiplier for the vth view. We exploit the alternating minimization strategy for solving problem Equation (8).

Solving $\boldsymbol{P}^v$: We fix $\boldsymbol{S}^l$ and $\boldsymbol{P}^{(u)}, u \neq v$ to update $\boldsymbol{P}^v$ for the v-th view, and the objective function can be transformed as follows:

$$\begin{aligned}\min_{\boldsymbol{P}^v}\quad & \sum_{i=1}^{N}\left\|(\boldsymbol{P}^v)^T(\boldsymbol{x}_i^v - \boldsymbol{X}^v\boldsymbol{s}_i^v)\right\|_2^2 \\ s.t.\quad & (\boldsymbol{P}^v)^T\boldsymbol{X}^v(\boldsymbol{X}^v)^T\boldsymbol{P}^v = \boldsymbol{I}.\end{aligned} \tag{9}$$

In order to facilitate the solution, we first transform Equation (9) into the following form by means of algebraic operations.

$$\begin{aligned}\min_{\boldsymbol{P}^v}\quad & Tr\left\{(\boldsymbol{P}^v)^T\boldsymbol{X}^v\boldsymbol{M}^v(\boldsymbol{X}^v)^T\boldsymbol{P}^v\right\} \\ s.t.\quad & (\boldsymbol{P}^v)^T\boldsymbol{X}^v(\boldsymbol{X}^v)^T\boldsymbol{P}^v = \boldsymbol{I},\end{aligned} \tag{10}$$

where $\boldsymbol{M}^v = \boldsymbol{I} - \boldsymbol{S}^v - (\boldsymbol{S}^v)^T + (\boldsymbol{S}^v)^T\boldsymbol{S}^v$. Problem Equation (10) can be solved by following generalized eigenvalue problems:

$$\boldsymbol{X}^v\boldsymbol{M}^v(\boldsymbol{X}^v)^T\boldsymbol{p}^v = \alpha\boldsymbol{X}^v(\boldsymbol{X}^v)^T\boldsymbol{p}^v. \tag{11}$$

Since the smallest eigenvalue of $\boldsymbol{M}^v$ is close to 0, we select the eigenvectors corresponding to the smallest 2 to $d^v + 1$ eigenvalue to construct the projection matrix $\boldsymbol{P}^v = \left[\boldsymbol{p}_2^v, \boldsymbol{p}_3^v, \ldots, \boldsymbol{p}_{d^v+1}^v\right]$.

Solving $\boldsymbol{S}^v$: We fix $\boldsymbol{P}^l$ and $\boldsymbol{S}^{(u)}, u \neq v$ to update $\boldsymbol{S}^v$ for the vth view, and the objective function can be transformed as follows:

$$\begin{aligned}\min_{\boldsymbol{S}^v}\quad & \sum_{i=1}^{N}\left\|(\boldsymbol{P}^v)^T(\boldsymbol{x}_i^v - \boldsymbol{X}^v\boldsymbol{s}_i^v)\right\|_2^2 + \gamma\sum_{u=1,u\neq v}^{m}\|\boldsymbol{S}^v - \boldsymbol{S}^u\|_F^2 \\ s.t.\quad & s_{ij}^v \geq 0, \quad \sum_j s_{ij}^v = 1.\end{aligned} \tag{12}$$

We find that the optimization of problem Equation (12) is independent of each $\boldsymbol{s}_i^v$, which is the row vector of $\boldsymbol{S}^v$. To simplify Equation (12), we utilize $\boldsymbol{Y}^v = (\boldsymbol{P}^v)^T\boldsymbol{X}^v$ to represent the low-dimensional projection of the original feature $\boldsymbol{X}^v$. Then, we can reformulate Equation (12) as follows:

$$\begin{aligned}\min_{\boldsymbol{s}_i^v}\quad & \sum_{i=1}^{N}\|\boldsymbol{y}_i^v - \boldsymbol{Y}^v\boldsymbol{s}_i^v\|_2^2 + \gamma\sum_{u=1,u\neq v}^{m}\|\boldsymbol{s}_i^v - \boldsymbol{s}_i^u\|_F^2 \\ s.t.\quad & s_{ij}^v \geq 0, \quad \sum_j s_{ij}^v = 1.\end{aligned} \tag{13}$$

It is evident that the above problem Equation (13) is a typical quadratic programming problem, which is equivalent to the following standard form without considering constant scalar:

$$\begin{aligned} \min_{s_i^v} \quad & \frac{1}{2}(s_i^v)^T \boldsymbol{H}^v s_i^v - (s_i^v)^T \boldsymbol{h}_i \\ s.t. \quad & s_{ij}^v \geq 0, \quad \sum_j s_{ij}^v = 1, \end{aligned} \tag{14}$$

where $\boldsymbol{H}^v = (\boldsymbol{Y}^v)^v \boldsymbol{Y}^v + (m-1)\gamma \boldsymbol{I}$ and $\boldsymbol{h}_i = \boldsymbol{Y}^v \boldsymbol{y}_i^v + \gamma\left(\sum_{u=1,u\neq v}^m \boldsymbol{s}_i^u\right)$. We can obtain the optimal solution of the above problem Equation (14) by solving the complementary slackness KKT condition.

Solving $\boldsymbol{Z}^v$: Since $\boldsymbol{Z}^v$ and $\boldsymbol{Z}^u, u \neq v$ are independent of each other, we can focus on $\boldsymbol{Z}^v$ without considering others. We fix $\boldsymbol{S}^l$ and $\boldsymbol{P}^{(l)}$ to update $\boldsymbol{Z}^v$, and the objective function can be transformed as follows:

$$\min_{\boldsymbol{Z}^v} \quad \eta \|\boldsymbol{Z}^v\|_* + \frac{\mu}{2}\|\boldsymbol{S}^v - \boldsymbol{Z}^v\|_F^2 + Tr((\boldsymbol{S}^v - \boldsymbol{Z}^v)\boldsymbol{J}^v). \tag{15}$$

Without considering the constant scalar, the problem Equation (15) can be reformulated as the following equivalent form:

$$\min_{\boldsymbol{Z}^v} \quad \eta \|\boldsymbol{Z}^v\|_* + \frac{\mu}{2}\left\|\boldsymbol{Z}^v - \boldsymbol{S}^v - \frac{1}{\mu}\boldsymbol{J}^v\right\|_F^2. \tag{16}$$

The optimal solution to problem Equation (16) can be obtained as follows:

$$\boldsymbol{Z}^v = \mathcal{D}_{\frac{\eta}{\mu}}\left(\boldsymbol{S}^v + \frac{1}{\mu}\boldsymbol{J}^v\right) \tag{17}$$

in which $\mathcal{D}_{\frac{\eta}{\mu}}$ is a singular value-thresholding shrinkage operation [36].

Updating multipliers $\boldsymbol{J}^v$: Finally, we can update the Lagrangian multipliers $\boldsymbol{J}^v$ as follows:

$$\boldsymbol{J}^v = \boldsymbol{J}^v + \mu(\boldsymbol{S}^v - \boldsymbol{Z}^v). \tag{18}$$

We iteratively alternate the four steps until the objective function Equation (8) satisfies the convergence condition. We summarize the details of optimizing MvLRGO in Algorithm 1.

Algorithm 1: The optimization procedure of MvLRGO.

Input:
A set of data matrices: $\{\boldsymbol{X}^v\}_{v=1}^m$ and the regularization parameters η and γ.
Initialization:
$\boldsymbol{S}^v$ is initialized by LRR, $\boldsymbol{P}^v = \boldsymbol{Z}^v = \boldsymbol{J}^v = \boldsymbol{0}$, $\rho = 1.2$, $\max_\mu = 10^4$.
The optimization procedure of MvLRGO:
1. Do
2. For v = 1:m
Update projection matrix $\boldsymbol{P}^v$ for the vth view according to Equation (11).
Update graph matrix $\boldsymbol{S}^v$ for the vth view by solving KKT conditions of problem Equation (14).
Update $\boldsymbol{Z}^v$ for the vth view according to Equation (17).
Update the multiplier $\boldsymbol{J}^v$ for the vth view according to Equation (18).
Update the parameter μ by $\mu = \min(\rho\mu, \max_\mu)$.
End
3. Until convergence.
Output:
The projection matrices $\boldsymbol{P}^v$, $1 \leq v \leq m$ for all views.

3.3. Complexity

The optimization procedure of the proposed MvLRGO, as listed in Algorithm 1, includes four main steps. The updating Lagrangian multiplier $\boldsymbol{J}^v$ for each view only involves element-wise addition, which means that the computational complexities of this step are minimal enough to be ignored. Therefore, we only analyze the steps involved in updating $\boldsymbol{P}^v$, $\boldsymbol{S}^v$, and $\boldsymbol{Z}^v$ without considering the addition of a matrix. We adopt N, m, D, and d to denote the number of samples, the number of views, the maximum dimensionality of the original features of all views, and the maximum dimensionality of the optimal subspaces of all views. To solve the projection $\boldsymbol{P}^v$ for all views, the complexity is $\boldsymbol{O}(mD^3)$. The complexity of solving the graph $\boldsymbol{S}^v$ for all views is $\boldsymbol{O}(m(dN^2 + d^2N))$. For a matrix with size $N \times N$, the complexity of solving the singular value-thresholding shrinkage operation is $\boldsymbol{O}(rN^2)$, where $r < N$ denotes the rank of the objective matrix. Then, we can determine the complexity of updating the additional variable $\boldsymbol{Z}^v$ for all views, as follows: $\boldsymbol{O}(mrN^2)$. Therefore, overall, the complexity of our proposed algorithm is $\boldsymbol{O}(\tau m(D^3 + dN^2 + d^2N + rN^2))$, where τ is the number of iterations.

4. Experiments

To validate the effectiveness of the proposed MvLRGO algorithm, we conducted multiple experiments on two textual datasets and four image datasets. The textual datasets were standard multi-view datasets. For the image datasets, we used multiple image feature operators to extract features from the datasets. The extracted features served as the multi-view features of the images.

BBCSport (http://mlg.ucd.ie/datasets/bbc.html, accessed on 12 June 2024): This dataset is a multi-view textual dataset, which includes 544 samples of sports news from five topical areas, which correspond to five classes (athletics, cricket, football, rugby, and tennis). Each sample contains two view features.

3Sources (http://mlg.ucd.ie/datasets/3sources.html, accessed on 12 June 2024): This dataset includes material from three online news sources, including the BBC, Reuters, and Guardian websites, and each source was treated as a view and collated into the dataset. All the samples contain three view features.

Yale (http://www.cad.zju.edu.cn/home/dengcai/Data/FaceData.html, accessed on 12 June 2024): This face dataset contains 165 grayscale images of 15 individuals, and each subject is featured in 11 images. We extracted the features of each image from three views (grayscale intensity, LBP, and Gabor) to construct the multi-view feature.

ORL (http://www.cad.zju.edu.cn/home/dengcai/Data/FaceData.html, accessed on 12 June 2024): This dataset is also a face dataset, and it contains 400 images from 40 individuals with 10 face images per individual. For ORL datasets, three types of features are the same as those in Yale.

Caltech101-7 (https://data.caltech.edu/records/mzrjq-6wc02, accessed on 12 June 2024): This is a subset of Caltech-101, which contains seven classes. For the Caltech101-7 dataset, we extracted six types of features from each image: HOG, LBP, GIST, WM, Gabor, and CENTRIST as the multi-view feature.

MSRC-v1 (https://www.microsoft.com/en-us/research/project/image-understanding/downloads/, accessed on 12 June 2024): This set contains 240 images from eight classes. We exploited seven classes for our experiment and extracted six types (CENT, CMT, GIST, HOG, LBP, and SIFT) of view features from each image.

We compared our method with the following baselines: To verify the performance of the MvLRGO algorithm, eight algorithms were chosen for the comparison: (1) locality-preserving projections (LPP) [13], (2) neighborhood-preserving embedding (NPE) [14], (3) graph-optimized locality-preserving projections (GoLPP) [19], (4). multi-view canonical correlation analysis (MCCA) [32], (5) multi-view spectral embedding (MSE) [34], (6) multi-view dimensionality co-reduction (McDR) [35], (7) co-regularized multi-view sparse reconstruction embedding (CMSRE) [29], and (8) multi-view unsupervised dimensionality

reduction with probabilistic neighbors (MDPN) [44]. To select the optimal hyperparameters for such experiments, five-fold cross-validation was adopted for all methods.

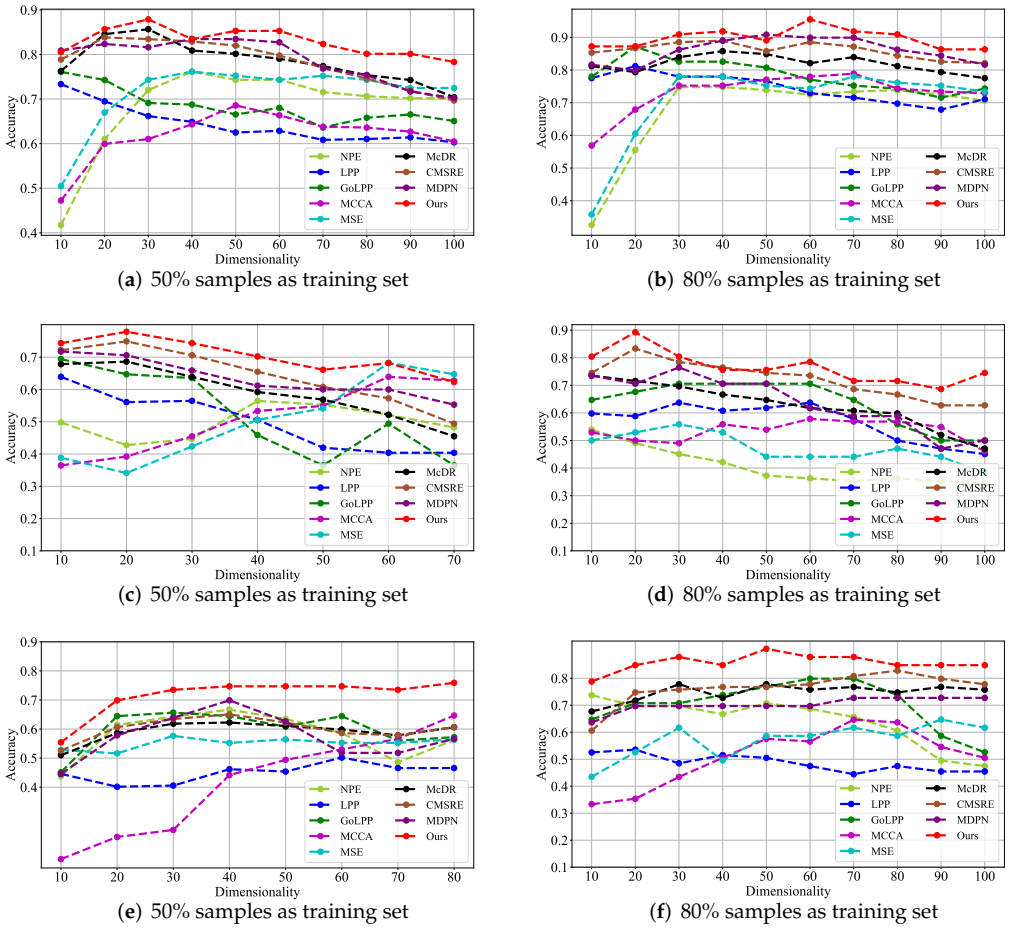
To validate the performance of the MvLRGO algorithm, in this section, we conduct classification and clustering experiments on six datasets. For classification experiments, we adopt four evaluation metrics: accuracy, precision, recall, and f1-score. For clustering, we also adopt four evaluation metrics: accuracy (ACC), normalized mutual information (NMI), adjusted rand index (AR), and F-score [45,46].

4.1. Classification Experiments

To validate the performance of the MvLRGO algorithm, this section presents classification experiments using six datasets.

For each dataset, we randomly selected 80% of the samples to construct the training set, and then all the algorithms were used to construct low-dimensional subspaces with 60 dimensions. A 3-NN classifier was then used to classify test samples and verify the performance of each algorithm. The experiment was run 20 times, and each time, a different sample was randomly selected as the training sample; the averages of the experimental results are shown in Tables 1–3.

Moreover, to evaluate the effect of dimensionality on classification performance, we set different dimensionalities for all methods to construct the low-dimensional subspace, and we also utilized a 3-NN classifier for classification. We demonstrate the average accuracies of all views with respect to varying dimensionality for all methods in Figures 1 and 2.

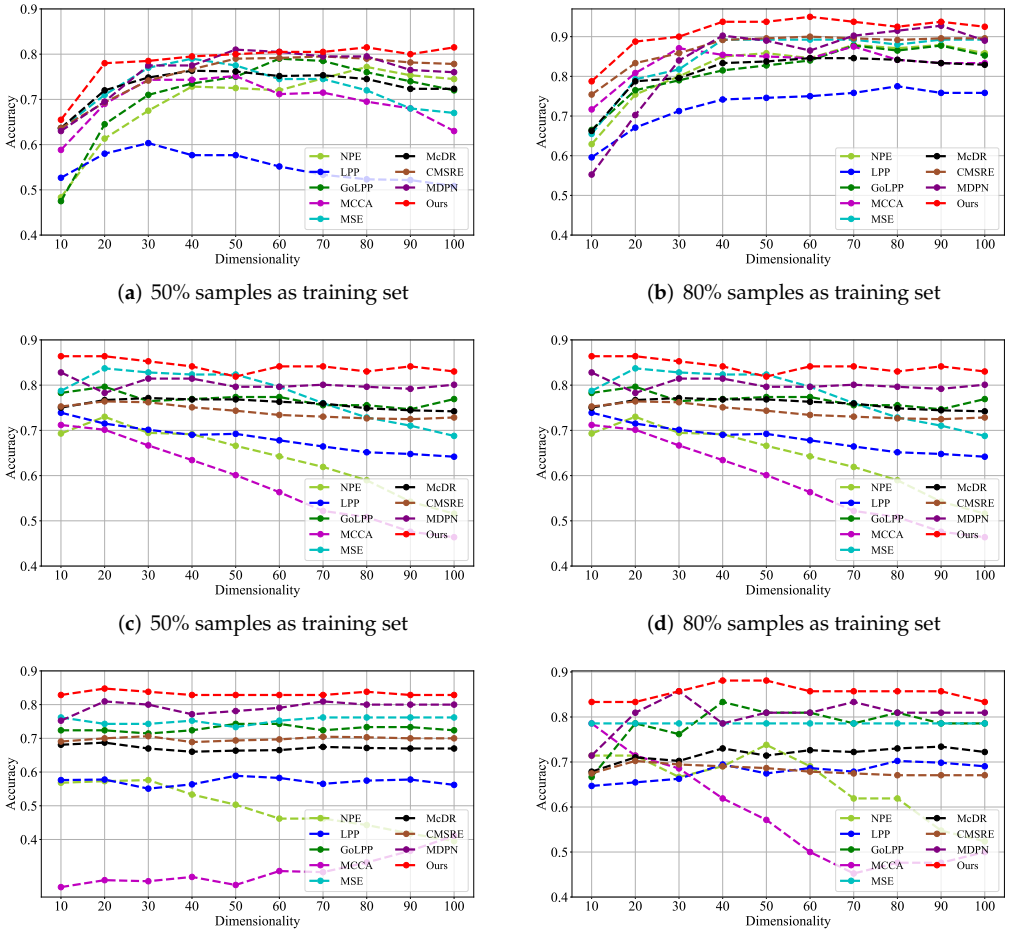


(**a**) 50% samples as training set (**b**) 80% samples as training set

(**c**) 50% samples as training set (**d**) 80% samples as training set

(**e**) 50% samples as training set (**f**) 80% samples as training set

Figure 1. The classification accuracies of different algorithms via the 3NN classifier with 50% and 80% samples per class as the training set. (**a**,**b**) BBCSport, (**c**,**d**) 3Sources, and (**e**,**f**) Yale.



(**a**) 50% samples as training set (**b**) 80% samples as training set

(**c**) 50% samples as training set (**d**) 80% samples as training set

(**e**) 50% samples as training set (**f**) 80% samples as training set

Figure 2. The classification accuracies of different algorithms via the 3NN classifier with 50% and 80% samples per class as the training set. (**a**,**b**) ORL, (**c**,**d**) Caltech101-7, and (**e**,**f**) MSRC-v1.

Table 1. The performance of all methods in the classification experiments on the 3Sources and BBCSport datasets. (The best results are bolded to show.)

Datasets	Methods	Accuracy	Precision	Recall	F1-Score
3Sources	NPE	0.461	0.274	0.259	0.210
	LPP	0.507	0.295	0.308	0.266
	GoLPP	0.500	0.403	0.313	0.299
	MCCA	0.446	0.271	0.258	0.231
	MSE	0.559	0.417	0.327	0.366
	McDR	0.680	0.514	0.506	0.473
	CMSRE	0.682	0.566	0.515	0.493
	MDPN	0.676	0.513	0.511	0.443
	Ours	**0.794**	**0.653**	**0.650**	**0.595**
BBCSport	NPE	0.788	0.862	0.752	0.783
	LPP	0.760	0.855	0.717	0.753
	GoLPP	0.771	0.828	0.767	0.777
	MCCA	0.707	0.686	0.725	0.689
	MSE	0.798	0.843	0.774	0.794
	McDR	0.853	0.888	0.829	0.848
	CMSRE	0.855	0.889	0.836	0.853
	MDPN	0.872	0.889	0.863	0.874
	Ours	**0.908**	**0.921**	**0.883**	**0.898**

Table 2. The performance of all methods in classification experiments on the Yale and ORL datasets.(The best results are bolded to show.)

Datasets	Methods	Accuracy	Precision	Recall	F1-Score
Yale	NPE	0.705	0.741	0.706	0.682
	LPP	0.516	0.561	0.524	0.489
	GoLPP	0.757	0.813	0.720	0.709
	MCCA	0.604	0.603	0.632	0.561
	MSE	0.667	0.651	0.581	0.579
	McDR	0.848	0.852	0.840	0.820
	CMSRE	0.878	0.886	0.850	0.847
	MDPN	0.909	0.917	0.931	0.914
	Ours	**0.939**	**0.946**	**0.952**	**0.942**
ORL	NPE	0.889	0.897	0.890	0.877
	LPP	0.782	0.798	0.778	0.759
	GoLPP	0.888	0.842	0.836	0.822
	MCCA	0.903	0.895	0.894	0.880
	MSE	0.900	0.873	0.879	0.858
	McDR	0.894	0.887	0.884	0.870
	CMSRE	0.892	0.888	0.886	0.871
	MDPN	0.913	0.931	0.906	0.903
	Ours	**0.950**	**0.944**	**0.936**	**0.931**

Table 3. The performance of all methods in classification experiments on Caltech101-7 and MSRC-v1 datasets. (The best results are bolded to show.)

Datasets	Methods	Accuracy	Precision	Recall	F1-Score
Caltech101-7	NPE	0.707	0.739	0.637	0.647
	LPP	0.737	0.763	0.670	0.675
	GoLPP	0.730	0.698	0.648	0.642
	MCCA	0.744	0.672	0.738	0.677
	MSE	0.719	0.783	0.729	0.701
	McDR	0.820	0.815	0.766	0.761
	CMSRE	0.773	0.774	0.722	0.723
	MDPN	0.831	0.825	0.791	0.787
	Ours	**0.853**	**0.873**	**0.834**	**0.828**
MSRC	NPE	0.581	0.623	0.585	0.555
	LPP	0.682	0.694	0.685	0.653
	GoLPP	0.690	0.752	0.682	0.628
	MCCA	0.583	0.593	0.626	0.558
	MSE	0.714	0.657	0.781	0.667
	McDR	0.729	0.736	0.731	0.705
	CMSRE	0.809	0.688	0.728	0.662
	MDPN	0.785	0.804	0.748	0.766
	Ours	**0.857**	**0.899**	**0.877**	**0.863**

As can be seen in Tables 1–3 and Figures 1 and 2, the proposed MvLRGO outperforms the other six algorithms in most cases. The performance of single-view algorithms is limited due to their reliance on one single-view feature. As shown in Tables 1–3, the McDR and CMSRE algorithms also perform well because they can utilize the complementary information of multiple features. MvLRGO utilizes a low-rank representation of the constructed correlation between samples, which means relationships between samples can be discovered effectively. Furthermore, MvLRGO employs a graph optimization framework to flexibly extract compatible and complementary information from multiple-view features to construct an optimal subspace for each view.

4.2. Clustering Experiments

The clustering experiments were conducted on six datasets. All the algorithms were used to reduce the dimensionality of original features from each view to 50 for all datasets. After dimensionality reduction, we performed k-means 20 times on the low-dimensional representation of each view for all methods and reported the average performance via the four evaluation metrics presented in Tables 4–6.

Table 4. The performance of all methods in clustering experiments on the 3Sources and BBCSport datasets.(The best results are bolded to show.)

Datasets	Methods	ACC	NMI	F-Score	AR
3Sources	NPE	0.345	0.105	0.014	0.192
	LPP	0.372	0.105	0.027	0.186
	GoLPP	0.390	0.163	0.083	0.259
	MCCA	0.368	0.134	0.052	0.212
	MSE	0.361	0.140	0.049	0.216
	McDR	0.437	0.157	0.031	0.221
	CMSRE	0.414	0.151	0.021	0.187
	MDPN	0.372	0.089	0.028	0.186
	Ours	**0.431**	**0.168**	**0.090**	**0.298**
BBCSport	NPE	0.396	0.097	0.021	0.201
	LPP	0.465	0.220	0.098	0.362
	GoLPP	0.520	0.320	0.211	0.402
	MCCA	0.374	0.043	0.013	0.157
	MSE	0.378	0.050	0.010	0.159
	McDR	0.422	0.153	0.036	0.223
	CMSRE	0.401	0.100	0.025	0.223
	MDPN	0.403	0.110	0.023	0.206
	Ours	**0.592**	**0.367**	**0.225**	**0.490**

Table 5. The performance of all methods in clustering experiments on the Yale and ORL datasets.(The best results are bolded to show.)

Datasets	Methods	ACC	NMI	F-Score	AR
Yale	NPE	0.501	0.567	0.242	0.533
	LPP	0.545	0.594	0.338	0.550
	GoLPP	0.573	0.604	0.355	0.580
	MCCA	0.576	0.617	0.336	0.577
	MSE	0.581	0.647	0.415	0.572
	McDR	0.606	0.679	**0.473**	0.597
	CMSRE	0.579	0.642	0.404	0.589
	MDPN	0.587	0.679	0.449	0.565
	Ours	**0.654**	**0.689**	0.455	**0.675**
ORL	NPE	0.541	0.763	0.394	0.524
	LPP	0.573	0.767	0.438	0.570
	GoLPP	0.616	0.797	0.475	0.610
	MCCA	0.640	0.805	0.504	0.644
	MSE	0.600	0.785	0.477	0.596
	McDR	0.643	0.812	0.517	0.632
	CMSRE	0.657	0.820	0.528	0.659
	MDPN	0.618	0.800	0.486	0.625
	Ours	**0.673**	**0.838**	**0.557**	**0.665**

Table 6. The performance of all methods in clustering experiments on Caltech 101-7 and MSRC-v1.(The best results are bolded to show.)

Datasets	Methods	ACC	NMI	F-Score	AR
Caltech101-7	NPE	0.541	0.492	0.334	0.440
	LPP	0.561	0.513	0.391	0.514
	GoLPP	0.616	0.554	0.411	0.555
	MCCA	0.612	0.564	0.407	0.559
	MSE	0.601	0.535	0.361	0.602
	McDR	0.609	0.540	0.403	0.549
	CMSRE	0.566	0.500	0.320	0.523
	MDPN	0.557	0.491	0.304	0.516
	Ours	**0.646**	**0.588**	**0.443**	**0.587**
MSRC	NPE	0.581	0.523	0.380	0.564
	LPP	0.593	0.506	0.381	0.594
	GoLPP	0.581	0.517	0.375	0.586
	MCCA	0.752	0.633	0.546	0.751
	MSE	0.733	0.628	0.533	0.684
	McDR	0.614	0.546	0.424	0.619
	CMSRE	0.642	0.531	0.423	0.651
	MDPN	0.614	0.546	0.424	0.618
	Ours	**0.800**	**0.800**	**0.707**	**0.758**

From Tables 4–6, it is clear that MvLRGO outperforms the other methods. In addition, although GoLPP can construct low-dimensional subspaces using superior affinity graphs, it is unable to explore the intrinsic correlation between each view in order to improve the performance. It is clear from these results that MvLRGO can fully learn compatible and complementary information about the underlying structures of multiple views and, therefore, obtain better low-dimensional subspaces.

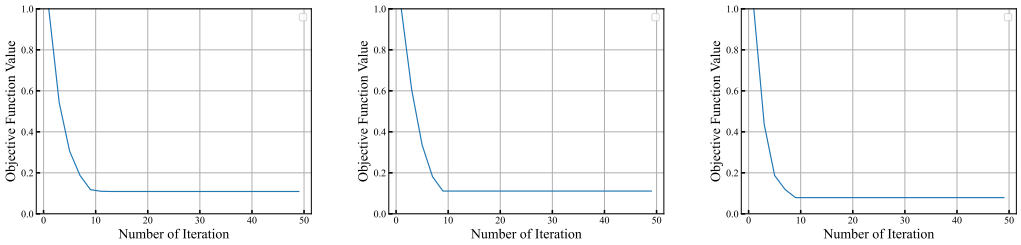
4.3. Discussion

This section discusses the convergence and training time of the proposed MvLRGO algorithm and the effect of hyperparameters on its performance. Since the MvLRGO algorithm uses an iterative optimization approach for low-dimensional representations, it is necessary to discuss its convergence. Herein, we evaluate the training process for the BBCSport, 3Sources, and Yale datasets in clustering experiments and summarize the objective function value with the increase in iterations. Figure 3 demonstrates the trend of the objective function value alongside the number of iterations, which shows the convergence of the MvLRGO algorithm. As can be seen in Figure 3, the objective function value of the MvLRGO tends to stabilize when the number of iterations is around 10.

Moreover, the training times of all methods in classification experiments are presented in Table 7. We can see from Table 7 that MDPN and our proposed MvLRGO algorithm require longer times for training. This is because MDPN and MvLRGO both incorporate graph learning and subspace learning into one framework, and the graph learning step is usually more complex than subspace learning.

Table 7. The training time(s) of all methods in classification experiments on six datasets.

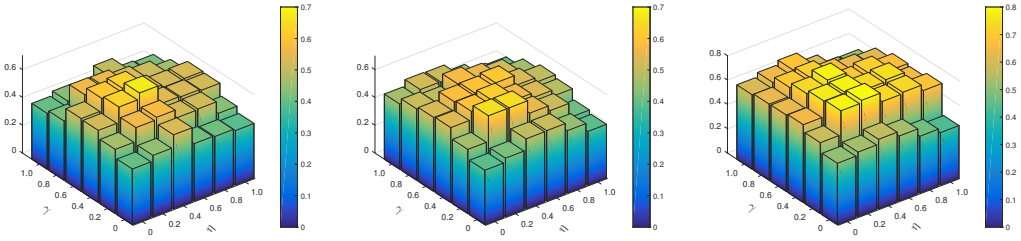
Methods	3Sources	BBCSport	Yale	ORL	Caltech101-7	MSRC
NPE	20.1 s	43.5 s	21.2 s	50.0 s	108.0 s	50.0 s
LPP	0.3 s	0.5 s	0.5 s	1.1 s	0.7 s	0.4 s
GoLPP	4.5 s	8.0 s	7.9 s	17.0 s	10.0 s	7.0 s
MCCA	1.7 s	1.2 s	1.6 s	2.0 s	3.4 s	3.2 s
MSE	5.3 s	4.0 s	5.0 s	5.0 s	7.3 s	10.0 s
McDR	0.5 s	0.6 s	0.5 s	1.1 s	0.9 s	0.6 s
CMSRE	20.2 s	140.0 s	20.8 s	115.6 s	237.3 s	60.1 s
MDPN	217.6 s	111.5 s	189.3 s	107.5 s	657.7 s	445.6 s
Ours	240.0 s	97.2 s	201.0 s	140.3 s	432.0 s	532.2 s



(**a**) BBCSport (**b**) 3Sources (**c**) Yale

Figure 3. The values of objective functions compared with the number of iterations on the BBCSport, 3Sources, and Yale databases.

In the final objective function, there are two balance hyperparameters, γ and η. In order to discuss the effect of different values of hyperparameters on the performance of the algorithm, a grid search strategy is used to search for the appropriate γ and η. We vary one of the hyperparameters in the range of $\{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ when another is fixed at a certain value. Then, we demonstrate the trend in ACC with different values of γ and η via clustering experiments on the ORL, Caltech101-7, and MSRC-v1 datasets, as shown in Figure 4. As can be seen in Figure 4, the clustering performance ACC usually shows promising results with these three datasets when the values of γ are in the range of 0.2 to 0.4 and η is in the range of 0.2 to 0.8.



(**a**) ORL (**b**) Caltech101-7 (**c**) MSRC-v1

Figure 4. The clustering accuracy of the proposed MvLRGO with different values of parameters γ and η with ORL, Caltech101-7, and MSRC-v1 datasets for clustering.

5. Conclusions

We propose a novel dimensionality reduction algorithm for multi-view data via low-rank representation and a graph optimization scheme; this algorithm is termed multi-view low-rank graph optimization for dimensionality reduction. MvLRGO aims to combine projection learning and graph optimization into one framework for each single-view feature, meaning the optimal subspace and affinity graph can be constructed simultaneously. Moreover, we hypothesize that, although features from different views contain different information from a given sample, their underlying structures are similar. Therefore, we adopt a low-rank representation to construct the affinity graph for each view, as it can effectively describe their global structure and regularize the learned affinity graph for greater closeness. Based on the regularization term, MvLRGO can incorporate information from different views to obtain the optimal subspace without co-regularizing the low-dimensional representations. Numerous experiments indicate that the proposed approach can effectively explore the underlying structure of samples from given multi-view features and thereby achieve promising results. However, this work still has some limitations. First, the construction of a graph for each view means our algorithm comes with a high computational cost. In the future, the gradient projection method [47] will be considered in order to accelerate the process of graph construction. Second, our model mainly focuses on the multi-view consistency issue but often neglects inconsistency in views, which makes the graphs vulnerable to low-quality or noisy datasets.

Author Contributions: Conceptualization, H.L. and H.W.; methodology, H.L.; software, H.L.; formal analysis, H.L. and H.W.; writing—original draft preparation, H.L.; writing—review and editing, H.W.; visualization, H.L.; supervision, H.W.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science Foundation of Zhejiang Sci-Tech University: 11432932612227.

Data Availability Statement: No public involvement in any aspect of this research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhao, J.; Xie, X.; Xu, X.; Sun, S. Multi-view learning overview: Recent progress and new challenges. *Inf. Fusion* **2017**, *38*, 43–54. [CrossRef]
2. Wang, Y. Survey on deep multi-modal data analytics: Collaboration, rivalry, and fusion. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* **2021**, *17*, 1–25. [CrossRef]
3. Wang, H.; Jiang, G.; Peng, J.; Deng, R.; Fu, X. Towards Adaptive Consensus Graph: Multi-View Clustering via Graph Collaboration. *IEEE Trans. Multimed.* **2023**, *25*, 6629–6641. [CrossRef]
4. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [CrossRef]
5. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; Volume 1, pp. 886–893.
6. Torralba, A.; Murphy, K.P.; Freeman, W.T.; Rubin, M.A. Context-based vision system for place and object recognition. In Proceedings of the Ninth IEEE International Conference on Computer Vision, Nice, France, 13–16 October 2003; Volume 1, pp. 273–280.
7. Wang, H.; Wang, Y.; Zhang, Z.; Fu, X.; Zhuo, L.; Xu, M.; Wang, M. Kernelized Multiview Subspace Analysis By Self-Weighted Learning. *IEEE Trans. Multimed.* **2021**, *23*, 3828–3840. [CrossRef]
8. Zhan, K.; Nie, F.; Wang, J.; Yang, Y. Multiview Consensus Graph Clustering. *IEEE Trans. Image Process.* **2019**, *28*, 1261–1270. [CrossRef] [PubMed]
9. Wang, H.; Yao, M.; Chen, Y.; Xu, Y.; Liu, H.; Jia, W.; Fu, X.; Wang, Y. Manifold-based Incomplete Multi-view Clustering via Bi-Consistency Guidance. *IEEE Trans. Multimed.* 2024, *early access*.
10. Nie, F.; Chang, W.; Wang, R.; Li, X. Learning an Optimal Bipartite Graph for Subspace Clustering via Constrained Laplacian Rank. *IEEE Trans. Cybern.* **2023**, *53*, 1235–1247. [CrossRef] [PubMed]
11. Wold, S. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [CrossRef]
12. Balakrishnama, S.; Ganapathiraju, A. Linear discriminant analysis-a brief tutorial. *Inst. Signal Inf. Process.* **1998**, *18*, 1–8.
13. He, X.; Niyogi, P. Locality preserving projections. *Adv. Neural Inf. Process. Syst.* **2004**, *16*, 153–160.
14. He, X.; Cai, D.; Yan, S.; Zhang, H.J. Neighborhood preserving embedding. In Proceedings of the IEEE International Conference on Computer Vision, Beijing, China, 17–21 October 2005; Volume 2, pp. 1208–1213.
15. Qiao, L.; Chen, S.; Tan, X. Sparsity preserving projections with applications to face recognition. *Pattern Recognit.* **2010**, *43*, 331–341. [CrossRef]
16. Roweis, S.T.; Saul, L.K. Nonlinear dimensionality reduction by locally linear embedding. *Science* **2000**, *290*, 2323–2326. [CrossRef] [PubMed]
17. Belkin, M.; Niyogi, P. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **2003**, *15*, 1373–1396. [CrossRef]
18. Tenenbaum, J.B.; De Silva, V.; Langford, J.C. A global geometric framework for nonlinear dimensionality reduction. *Science* **2000**, *290*, 2319–2323. [CrossRef] [PubMed]
19. Zhang, L.; Qiao, L.; Chen, S. Graph-optimized locality preserving projections. *Pattern Recognit.* **2010**, *43*, 1993–2002. [CrossRef]
20. Zhang, L.; Chen, S.; Qiao, L. Graph optimization for dimensionality reduction with sparsity constraints. *Pattern Recognit.* **2012**, *45*, 1205–1210. [CrossRef]
21. Qiao, L.; Zhang, L.; Chen, S. Dimensionality reduction with adaptive graph. *Front. Comput. Sci.* **2013**, *7*, 745–753. [CrossRef]
22. Yi, Y.; Wang, J.; Zhou, W.; Fang, Y.; Kong, J.; Lu, Y. Joint graph optimization and projection learning for dimensionality reduction. *Pattern Recognit.* **2019**, *92*, 258–273. [CrossRef]
23. Wang, H.; Cui, T.; Yao, M.; Pang, H.; Du, Y. Domain adaptive person search via GAN-based scene synthesis for cross-scene videos. *Image Vis. Comput.* **2023**, *138*, 104796. [CrossRef]
24. Wang, H.; Yao, M.; Jiang, G.; Mi, Z.; Fu, X. Graph-Collaborated Auto-Encoder Hashing for Multiview Binary Clustering. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 1–13. .. [CrossRef] [PubMed]
25. Jiang, G.; Peng, J.; Wang, H.; Mi, Z.; Fu, X. Tensorial Multi-View Clustering via Low-Rank Constrained High-Order Graph Learning. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 5307–5318. [CrossRef]

26. Hu, J.; Lu, J.; Tan, Y.P. Sharable and Individual Multi-View Metric Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 2281–2288. [CrossRef] [PubMed]
27. Wang, H.; Feng, L.; Meng, X.; Chen, Z.; Yu, L.; Zhang, H. Multi-view metric learning based on KL-divergence for similarity measurement. *Neurocomputing* **2017**, *238*, 269–276. [CrossRef]
28. Zhang, C.; Liu, Y.; Liu, Y.; Hu, Q.; Liu, X.; Zhu, P. FISH-MML: Fisher-HSIC multi-view metric learning. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18, Stockholm, Sweden, 13–19 July 2018; pp. 3054–3060.
29. Wang, H.; Peng, J.; Fu, X. Co-regularized multi-view sparse reconstruction embedding for dimension reduction. *Neurocomputing* **2019**, *347*, 191–199. [CrossRef]
30. Li, H.; Su, Z.; Wang, H.; Liu, X. Autoweighted multi-view smooth representation preserve projection for dimensionality reduction. *J. Electron. Imaging* **2021**, *30*, 023022. [CrossRef]
31. Kumar, A.; Rai, P.; Daume, H. Co-regularized Multi-view Spectral Clustering. In *Proceedings of the Advances in Neural Information Processing Systems, Granada, Spain, 12 December 2011*; Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K., Eds.; Curran Associates, Inc.: Glasgow, UK, 2011; Volume 24.
32. Rupnik, J.; Shawe-Taylor, J. Multi-view canonical correlation analysis. In Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2010), Ljubljana, Slovenia, 12 October 2010; Volume 473, pp. 1–4.
33. Hardoon, D.R.; Szedmak, S.; Shawe-Taylor, J. Canonical correlation analysis: An overview with application to learning methods. *Neural Comput.* **2004**, *16*, 2639–2664. [CrossRef]
34. Xia, T.; Tao, D.; Mei, T.; Zhang, Y. Multiview spectral embedding. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **2010**, *40*, 1438–1446.
35. Zhang, C.; Fu, H.; Hu, Q.; Zhu, P.; Cao, X. Flexible multi-view dimensionality co-reduction. *IEEE Trans. Image Process.* **2016**, *26*, 648–659. [CrossRef] [PubMed]
36. Liu, G.; Lin, Z.; Yan, S.; Sun, J.; Yu, Y.; Ma, Y. Robust recovery of subspace structures by low-rank representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 171–184. [CrossRef] [PubMed]
37. Xie, D.; Gao, Q.; Zhao, Y.; Yang, F.; Song, W. Consistent graph learning for multi-view spectral clustering. *Pattern Recognit.* **2024**, *154*, 110598. [CrossRef]
38. Wang, H.; Yang, Y.; Liu, B. GMC: Graph-Based Multi-View Clustering. *IEEE Trans. Knowl. Data Eng.* **2020**, *32*, 1116–1129. [CrossRef]
39. Yan, S.; Xu, D.; Zhang, B.; Zhang, H.J.; Yang, Q.; Lin, S. Graph embedding and extensions: A general framework for dimensionality reduction. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *29*, 40–51. [CrossRef]
40. Lee, D.D.; Seung, H.S. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, *401*, 788–791. [CrossRef] [PubMed]
41. Hoyer, P.O. Modeling receptive fields with non-negative sparse coding. *Neurocomputing* **2003**, *52*, 547–552. [CrossRef]
42. Zhuang, L.; Gao, H.; Lin, Z.; Ma, Y.; Zhang, X.; Yu, N. Non-negative low rank and sparse graph for semi-supervised learning. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA 16–21 June 2012; pp. 2328–2335.
43. Wang, H.; Feng, L.; Yu, L.; Zhang, J. Multi-view Sparsity Preserving Projection for dimension reduction. *Neurocomputing* **2016**, *216*, 286–295. [CrossRef]
44. Qiang, Q.; Zhang, B.; Wang, F.; Nie, F. Multi-view unsupervised dimensionality reduction with probabilistic neighbors. *Neurocomputing* **2022**, *500*, 203–216. [CrossRef]
45. Schütze, H.; Manning, C.D.; Raghavan, P. *Introduction to Information Retrieval*; Cambridge University Press: Cambridge, UK, 2008; Volume 39.
46. Cai, D.; He, X.; Han, J. Document clustering using locality preserving indexing. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 1624–1637. [CrossRef]
47. Duchi, J.; Shalev-Shwartz, S.; Singer, Y.; Chandra, T. Efficient projections ontro the l 1-ball for learning in high dimensions. In Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, 5–9 July 2008; pp. 272–279.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.