Journal of Advanced Digital Communications

Scilight

Article

Linear Collaborative Representation Learning Approach for Dimensionality Reduction

Ayesha Jadoon

Department of Electronics, Telecommunications and Informatics, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal; ayeshajadoon@ua.pt
Received: 19 March 2024; Revised: 30 October 2024; Accepted: 5 November 2024; Published: 18 November 2024

Abstract: In dimensionality reduction techniques an important step is to construct optimal similarity graph to achieve effective classification results. The graph construction process in many existing algorithms is manual and thus severely affects the classification performance, if the neighborhood parameter is not optimal. Moreover, existing methods that are based on Collaborative representation lack the between-class information in the embedding process. In this paper, we addressed the problem of automatic Graph construction which is datum adaptive and incorporates within-class and between-class information into the linear representation to learn optimal projection for dimensionality reduction using the Collaborative representation technique. To optimize graph construction, the proposed method used the L_2 norm graph and log-Euclidean distance. The resultant graph shows local properties by Collaborative representation and global discriminate information is represented by a Maximum Margin classifier (MMC). The MMC maximizes "between-class scatter" and minimizes "within-class scatter", without locality information. Further for effective and accurate performance for image classification real databases will be incorporated. The experimental results have demonstrated that the proposed methods achieved competitive results with compared methods.

Keywords: index terms; K-nearest neighbor; PCA; locally linear embedding (LLE); locality preserving discriminant projection (LPDP)

1. Introduction

Dimensionality reduction aim is to transform the high-dimension data into lower-dimension by discarding the un-informative variance in the data. Conversion of high-dimensional data into low-dimensional data is appreciated in the visualization and compressing of data. Dimensionality reduction plays a significant role in better classification, prediction, and visualization of data [1]. There are many applications of dimensionality reduction such as text mining, image retrieval, biological data analysis (e.g., protein classification), and face recognition [2].

Dimensionality reduction is an effective method to improve the "curse of dimensionality" problem. The large number of features in the high dimensional data set can cause an overfitting problem, which can not only increase computational complexity but also degrade the performance in classification [3]. Answering the similarity quareies in high dimensions is also a big issue due to the "curse of dimensionality". To overcome this problem "Dimensionality reduction" is a useful technique to visualize the data and compress the data.

Recently, Manifold learning approaches for dimensionality reduction have shown great interest because these techniques can find out the unknown intrinsic data structure that is hidden in high dimensional data space [3, 4], such as Locally linear embedding (LLE) [5], neighborhood preserving projection (NPE) [3] and, locality preserving discriminant projection (LPDP) [1]. We use graphs to find the objects or entities' properties and their binary relationship. In the above-mentioned techniques, the graph construction process for similarity calculation among data items plays a significant role.

Popular subspace learning methods such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) [6], LLE [7, 8], neighborhood preserving projection (NPE), and, LPDP [1] start with graph construction. These techniques depend on the manual selection of suitable parameters to get the optimal graph construction which is a difficult task and time-consuming when data is on a large scale [9, 10].

The key problem in the optimal graph construction process is how to set the neighborhood size and weight parameters. In existing methods [11], neighborhood selection and weight assignment to corresponding edges are done separately, hence in the proposed method, it is desirable to simultaneously find adaptive neighborhood and

weight assignment under a unified framework.

Most of the dimensionality reduction approaches construct the nearest neighbor graph using Euclidean distance between images but it may not find out the hidden intrinsic structure [4]. To uncover the intrinsic structures from high-dimensional data, have shown that Log-Euclidean distance is more effective than Euclidean distance for neighborhood calculation in high-dimensional space [5]. Then the proposed supervised dimensional reduction method is introduced which is intrinsic structure feature transform (ISFT) which constructs the nearest neighbor graph using log-Euclidean distance for the measurement of neighbor circles to capture the local structure. It finds a representative image for each class and it captures the global structure by using projected data samples of these representatives to maximize the between-class scatter measure. To achieve the optimal graph projection matrix it utilizes the L_2 norm regularized least square to maximize the ratio between total separability information and local compactness [11].

There are many existing Dimensionality reduction techniques like Sparse based representation classification (SRC) technique which shows outstanding performance to get optimal graph construction [12, 13] but it needs to address the L_1 norm graph optimization problem as its computational cost is expensive and its computation is slow [2, 5, 14]. To overcome the problem of SRC, our proposed method uses the L_2 norm graph regularized least square which has considerably less image computational burden and achieves outstanding performance in the classification of images [7, 15].

In our proposed method it has been shown [1, 2, 7, 16, 17] that Collaborative representation performance for image classification is much better than sparse representation-based classification techniques. To measure effective and accurate performance, some of the publically available databases such as YaleB, AR and ORL will be used in our proposed algorithm.

Our proposed Dimensionality Reduction approach emphasizes several aspects:

1. It preserves the Collaborative representation which shows better performance as compared to Sparse representation.
2. The adjacency weight matrix is constructed based on the CR framework which shows the nonlinear geometry of the data set. Linear Collaborative representation has great capability to transform the hidden intrinsic structure in the high dimensional data set.
3. It achieves the graph adjacency matrix and their corresponding weights simultaneously in a unified framework.
4. For optimal graph projection matrix, it utilizes the L_2-norm least regularized square approach which has less computational cost as compared to L_1-norm regularization.
5. To select datum adaptive neighborhood, it uses Log-Euclidean distance which shows robust performance as compared to Euclidean distance.
6. To capture global properties to discriminate information it utilizes the (MMC) maximum margin framework which maximizes the "between class scatter" and minimizes the "within class scatter" for discriminate information.
7. Then finally low embedding is obtained to best preserve such reconstruction weights.
8. We evaluated the performance of the proposed algorithm on high dimensional data which reduces it to low dimension and gives the required results.
9. The Proposed algorithm is also tested on some publically available datasets such as YaleB, ORL, and UMIST data sets.

In the end, we also analyzed the performance of the algorithm through dimensionality reduction by projecting it from high dimensional data to lower dimensionality reduction. Our algorithm has incomparable results compared to other state-of-the-art algorithms.

2. Related Work on Dimensionality Reduction Technique

Security sectors and law enforcement agencies need image classification applications that are robust and can perform accurate classification with less computational time [18, 19]. Considering these real-world scenarios, the given test image may be partially noised (contiguous occlusion) and the training image samples may be insufficient to model all possible variations of a given test sample so under these conditions the face recognition problem becomes a very challenging task, and remain an open research problem [8, 20]. Real-world face images include a large number of variations; face images are high dimensional having several variations. The variation includes

factors like noise, illumination, and pose variations [21]. In these cases, identifying the accurate class of the test sample from the training images dictionary is a complicated task. Partially noised (contiguous occlusion) images make the face image classification task more complicated [10, 22, 23].

2.1. Graph-Based Approach

The Graph construction process for similarity calculation among data items plays a significant role. To find the properties of the object and their binary relationship we use graphs. The graph represents the pairwise relationship between data samples in which nodes represent the data points and edges represent the distance between the points. Data points are represented in the form of a graph, called a similarity graph. Where nodes of the graph represent the data points and edges represent their corresponding weights according to the distance between the points. Similar points carry higher weights and are grouped into the same class while the dissimilar points that carry low edge weights are grouped into different classes.

In existing methods [7, 8, 15], neighborhood selection and weight assignment to corresponding edges are done separately, hence in the proposed method, it is desirable to simultaneously find adaptive neighborhood and weight assignment under a unified framework. Some graph construction methods have been used by different researchers which include manual techniques as well as some automatic graph construction techniques [10, 24, 25]. In manual techniques, some researchers construct the graph using b-matching, and some construct using K-nearest neighbors and ϵ-ball neighborhoods [26, 27]. Automatic graph construction techniques include sparse representation-based framework (SRC) and collaborative representation-based framework (CRC) [28, 29] .

2.2. K-Nearest Neighbors (KNN) and ϵ-Ball Neighborhood

Graph construction techniques involve adjacency graph construction based on KNN, ϵ-ball neighborhoods, or fully connected graphs. The ϵ-ball neighborhood is considered an unweighted graph as all those points are connected whose pairwise distance is smaller than ϵ value [27]. In K–nearest- neighbor we connect vertex i_{th} with vertex j_{th} if they are the nearest neighbor depending on the value 'k' [26, 30].

In a fully connected graph, the graph shows the local neighborhood relationship in which we simply connect all data points with the same similarity with each other and is the parameter that controls the width of the neighbors in the fully connected graph as parameter controls in ϵ-ball neighborhood (function of both the parameter is the same). The distance between all the edges is measured as edges do not contain much information. The drawbacks of these methods are that these methods are manual and require the selection of appropriate parameters, the selection of subspace dimensions, etc. [31].

3. Proposed Method for Collaborative Neighborhood Representation Discriminant Embedding

We propose a collaborative representation framework for dimensionality reduction as shown in Figure 1. The proposed algorithm is formulated for data points that lie in a union of linear subspaces. The proposed algorithm focuses on adaptive graph representation which deals with the problems related to the automatic construction of graphs. It uses the self-expression property of data using CR to obtain automatic construction of graphs as well as assigning weights to edges automatically and embeds the high dimensional data into low dimensional space.

Our proposed algorithm consists of two major steps as shown in Figure 1. First, we find a few other data points that correspond to the same subspace for each sample using the proposed optimization program i-e collaborative neighborhood representation (CNR) that gives information about members of the data points. The collaborative representation learning method for dimensionality reduction has much interest because this technique has great capability to find out the hidden intrinsic data structure in the high dimensional data space [32, 33]. For optimal graph projection matrix, it utilizes the L_2-norm least regularized square approach which is less computational cost as compared to L_1-norm regularization [11, 34]. To capture global properties to discriminate information it utilizes the (MMC) maximum margin framework which maximizes the "between class scatter" and minimizes the "within class scatter" for discriminate information [3]. Then finally low embedding is obtained to best preserve such reconstruction weights. Secondly, the projection matrix is a "Collaborative representation" framework that shows the nonlinear geometry of the data set, and conveys discriminate information "within class scatter" and "between class scatter".

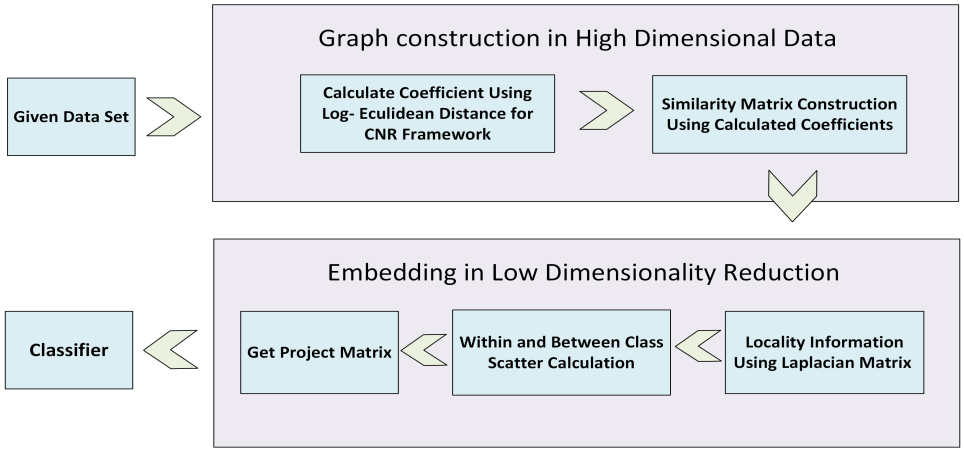

Figure 1. Block Diagram of Proposed System Model.

3.1. Steps for Dimensional Reduction

Following are the steps are:

3.1.1. Graph Construction

Data points are represented in the form of a graph, called a similarity graph. Where nodes of the graph represent the data points and edges represent their corresponding weights according to the distance between the points. The similar points carry higher weights and are grouped into the same class while the dissimilar point carries lower weights and are grouped in a different class [35].

An important step in existing [36] state-of-the-art Dimensionality reduction techniques is constructing an optimal similarity graph to achieve effective classification results. The graph construction process in many existing algorithms is manual and thus severely affects the classification performance, if the neighborhood parameter is not optimal [37]. Moreover, existing methods are based on Collaborative representation that lacks the between-class information in the embedding process [38]. In this work, we will address the problem of automatic graph construction which is datum adaptive and incorporates within-class and between-class information into the linear representation to learn optimal projection for dimensionality reduction using Collaborative representation. The adjacency weight matrix is constructed based on the CR framework which shows the nonlinear geometry of the data set.

3.1.2. Collaborative Nearest Neighborhood (CNR) Optimization

Our proposed algorithm has a special feature that uses the self-expressiveness property of the data which represents each data point in the union of the subspace, by the combination of the other data points in the same subspace. The first step of our algorithm is the selection of datasets that are available publicly. Datasets give information on various samples/data points or images and their corresponding number of classes, image sizes, etc [39]. A collaborative neighborhood representation (CNR) framework is used to calculate the effective coefficients to construct a similarity graph. Secondly, the projection matrix is the "Collaborative representation" framework which shows the nonlinear geometry of the data set, and conveys discriminate information "within class scatter" and "between class scatter". Then finally low embedding is obtained to preserve such reconstruction weights.

3.1.3. CNR Coefficients Calculation

The proposed algorithm makes use of the self-expressive property of data i-e it enforces automatic graph construction, inspired by Collaborative neighborhood representation-based classification (CNR) . CNR represents each data point as a linear combination of nearest neighbors. It chooses the nearest neighbor ideal to represent data points from the complete dictionary. Optimum selected neighbors lie in the same subspace as sample x_i. Linear representation of query sample among dictionary is represented by

Let $\{x_i\}^{n_{i=0}}$ be the data points that lie in the union of n subspaces having unknown dimensions.

$$x_i \approx D\beta \quad \texttt{Where} \quad D \in \mathbb{R}^{m \times n} \left(m > n\right) \tag{1}$$

D is an over-complete dictionary that is it represents the matrix that concatenates all the data points, and β represents the coefficients. Concatenating all the samples in D, and D_k reduced dimension in lower dimension as shown in below the equations. It can be presented as:

$$D = \begin{bmatrix} D_1 & D_2 & \dots & D_n \end{bmatrix} \in \mathbb{R}^{m \times n}$$

$$D_k = \begin{bmatrix} d_{k1} & d_{k2} & \dots & d_{kn_k} \end{bmatrix} \in \mathbb{R}^{m \times n_k}$$

$$\beta = \underset{\beta}{\operatorname{argmin}} \left(\frac{1}{2} \|x - D\beta\|_2^2 + \sigma \sum_{i=1}^{n} \beta_i^2 \|x - x_i\|_2^2 + \lambda \|\beta\|_2^2 \right) \tag{2}$$

The first term of the Equation (2) in brackets shows the linear reconstruction error calculated against each class, x represents the data point, D represents the dictionary, and β represents the coefficient computed against each data point. x_i represents each data point in dictionary D. where σ and λ are regularization terms having positive terms to introduce some sparsity and β_i^2 is a coefficient vector having positive values. The full proof of Equation (2) the equation is shown in the Appendix A.

The Proposed solution of the objective function is as follows

$$\beta = \left(D^T D + \lambda I + \sigma \begin{bmatrix} \|x - x_1\|_2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \|x - x_n\|_2 \end{bmatrix} \right)^{-1} D^T x \tag{3}$$

where $\beta_k = [0, \dots, 0, \beta_{k1}, \beta_{k2}, \dots, \beta_{kn_k}, 0, \dots, 0]^T$ is a vector whose only nonzero entries corresponding to the k_{th} class are the coefficients of β.

The diagonal matrix in Equation (3) representing distance between data point and all other data points of the dictionary. Obtained coefficients are used to build similarity graph. Obtained CNR coefficients from the diagonal matrix is used to generate weight or similarity matrix representing relationship between data points. If we receive the highest value of weight than it means data points are more similar points. Based on the corresponding coefficient weights similarity graph is constructed. Data points that have higher weights lie in the same space than others.

3.1.4. Embedding (Discriminate Projection Matrix)

The second step of dimensionality reduction is graph embedding in a low-dimensional subspace [33]. In our case, we built this similarity graph using our proposed method CNR. Obtained CNR coefficients from the diagonal matrix (as explained above) are used to generate a weight or similarity matrix representing the relationship between data points [40]. Similar data points receive more weight than others. Based on the corresponding weights similarity graph is constructed. Data points that have higher weights lie in the same space as others.

To capture global properties to discriminate information it utilizes the MMC framework which maximizes the "between class scatter" and minimizes the "within class scatter" for discriminate information [41]. Then finally low embedding is obtained to best preserve such reconstruction weights.

Optimization of the objection function is defined as follows:

$$\begin{cases} \min\limits_{i,j} \sum\limits_{i,j} \|y_i - y_j\|^2 W_{ij} \\ \text{s.t.} \quad P^T P = I \end{cases} \tag{4}$$

3.1.5. Graph Laplacians

The next step of the method is a calculation of graph Laplacian which is the main tool of locality information. Laplacian can be obtained from degree matrix D of W. Degree matrix is a diagonal matrix $d_{ii} = \sum_{j=1}^{n} W_{ij}$ and holds the information about the number of edges that are attached to each vertex.

Laplacian Matrix Is Defined as

$$L = D - W$$

For locality information: Construct Laplacian matrix as $L = D - W$. where D is a diagonal matrix whose entries are row sums of *W*, where $d_{ii} = \sum_{j=1}^{n} W_{ij}$ When we have connected components the adjacency matrix *W* forms the block diagonal and the same is also true for *L*.

$$L = \begin{bmatrix} L_1 & 0 & \dots & 0 \\ 0 & L_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L_n \end{bmatrix}$$

As *L* is a block diagonal that represents the Eigenvalues and corresponding Eigenvectors in corresponding blocks has *n* non-negative values. The smallest eigenvalue is zero. We find the largest eigenvectors in our graph.

Solving this equation:

$$R = P(D_{\text{row}} - W)P^T + P(D_{\text{col}} - W^T)P \tag{5}$$

where $(D_{\text{row}} - W) + (D_{\text{col}} - W^T)$ Laplacian matrix. D_{row} and D_{col} are diagonal matrices whose elements are the row and column sums of the *W* matrix, respectively.

For Between-Class Information and Within-Class Information

For each class x_i find representative data points according to SMRS and calculate between class scatter S_B using

$$S_B = \sum_{i=1}^{c} c_i(\mu_i - \mu)(\mu_i - \mu)^T \tag{6}$$

$$S_W = \sum_{i=1}^{c} \left(\sum_{j=1}^{c_i} (x_j^{(i)} - \mu_i)(x_j^{(i)} - \mu_{(i)})^T \right) \tag{7}$$

where μ is the total mean of class, μ_i is the mean of i_{th} class is the number of samples in the i_{th} class, c is the number of the classes c^i is the average vector of the i_{th} class, x_j^i is the j_{th} sample of i_{th} class.

Eigen Solver

The next step of the method is to compute the eigenvalues and eigenvectors for the problem.

$$S_B P = \lambda S_W P \tag{8}$$

where *P* is formed from the optimal eigenvectors corresponding to the largest eigenvalues.

In our proposed method, we use the top eigenvector of the Laplacian matrix. These eigenvectors are used to define the optimal cut.

3.1.6. Projection

Compute Projection matrix composed of optimal 'r' smallest eigenvectors

$$DLD^T P = \lambda \frac{\rho}{1-\rho} S_B P \tag{9}$$

where ρ is the parameter that balances local and global contribution for data classification.

Classification

The projection is composed of the optimal projected vectors corresponding to the smallest eigenvalues Assign the points to one or more classes based on sorted vectors. Once the eigenvectors are chosen next we sort the components of the reduced vector and identify the classes by splitting the sorted vector.

4. Simulation Setup and Results

The details of our experimental settings and simulation environment to generate low-dimensional space. It also includes comparison methods that are used to compare our proposed method and their description. Different benchmark datasets are used for dimensional reduction; their description and experimental setting are also explained in the following section.

4.1. Simulation Setup

We have implemented the whole work on the Fujitsu Core i3 machine which has a 2.30 GHz processor with 4 GB RAM. MATLAB software is used for simulation results. The main performance parameter of our proposed method includes accuracy.

4.2. Comparison with Previous Methods

We evaluated the performance of our proposed graph construction algorithm Collaborative neighborhood representation (CNR) [32] with different graph construction algorithms, Collaborative representation classifier [11], Locally linear embedding (LLE) [3, 5], neighborhood preserving projection (NPE) and, locality preserving projection (LPP) [1]. Also, we compare the performance of the proposed algorithm on different data sets.

4.3. Datasets Used

We evaluated the performance of our proposed graph construction technique on different datasets which include YaleB [42], UMIST database, and ORL [43] database. A detailed description of each dataset is presented below:

4.4. Extended YaleB Database

The extended YaleB database contains 2414 images of 38 individuals with 64 images in each class and the dimension of each image is 1024 (32×32). Each image varies under different lighting conditions for other images of the same and different classes. Figure 2 shows some example images of the extended YaleB database.

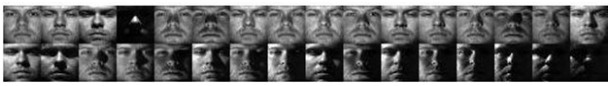

Figure 2. Example images of YaleB database.

4.4.1. Results on YaleB Database

We evaluated the performance of our algorithm on different dimensions of samples extracted from the YaleB database. We evaluated the results through a Collaborative representation framework (CR). Results clearly show that the performance of the proposed algorithm is better in terms of accuracy than comparison algorithms as shown in Table 1 and in Figure 3:

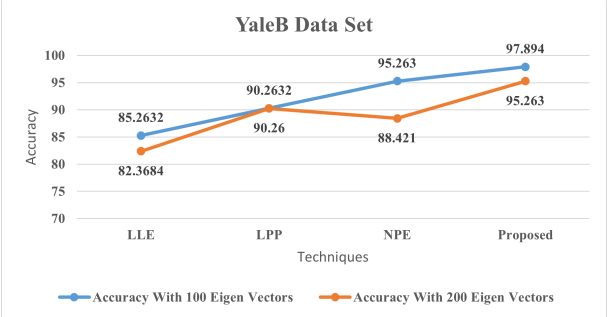

Figure 3. YaleB Data Set Accuracy graph.

Table 1. The highest accuracy achieved among comparison methods on the YaleB dataset.

Dimensionality Reduction	**LLE**	**LPP**	**NPE**	**Proposed**	**LLE**	**LPP**	**NPE**	**Proposed**
Reduced-dimension	100	100	100	100	200	200	200	200
Classes	38	38	38	38	38	38	38	38
Accuracy	85.2632	90.2632	95.263	97.894	82.3684	90.26	88.421	95.263
Correct Prediction	324	343	362	372	313	343	336	362
Training-data	380	380	380	380	380	380	380	380
Testing -data	380	380	380	380	380	380	380	380
Classifier	KNN	KNN	KNN	KNN	KNN	KNN	KNN	KNN

4.5. Umist Database

UMIST database contains 564 images of 20 individuals with 20 images in each class and the dimension of each image is 1024 (32 × 32). Each image varies under different pose variations for other images of the same and different classes. Figure 4 shows some example images of the UMIST database.

Figure 4. Example images of UMIST database.

4.5.1. Results on UMIST Database

We evaluated the performance of our algorithm on different samples extracted from the UMIST database. We evaluate the result on different training datasets and on dimension dimensions. We evaluated the results through a Collaborative representation framework (CR). Results clearly show that the performance of the proposed algorithm is better in terms of accuracy than the comparison algorithms in Figure 5 and in Table 2. On different dimensions 50, 80, and 100 when training and test data 320/80, 200/180 are used, the best results are achieved compared to existing methods, as shown in Figures 6 and 7.

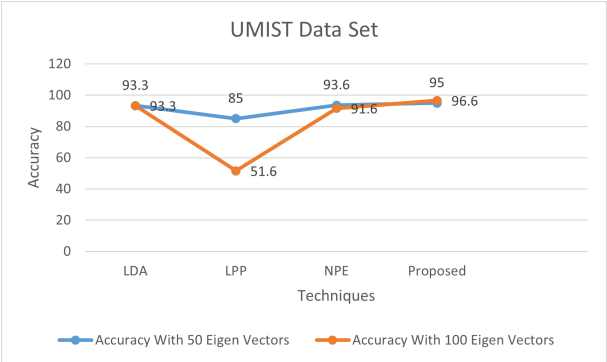

Figure 5. UMIST Data Set Accuracy graph.

Table 2. The highest accuracy achieved among comparisons methods on UMIST database.

Dimensionality Reduction	**LDA**	**LPP**	**NPE**	**Proposed**	**LDA**	**LPP**	**NPE**	**Proposed**
Red-dim	50	50	50	50	100	100	100	100
classes	20	20	20	20	20	20	20	20
Accuracy	93.3	85	93.6	95	93.3	51.6	91.6	96.6
Correct Prediction	55	45	55	57	55	40	53	58
Training-data	320	320	320	320	320	320	320	320
Testing -data	80	80	80	80	80	80	80	80
Classifier	KNN	KNN	KNN	KNN	KNN	KNN	KNN	KNN

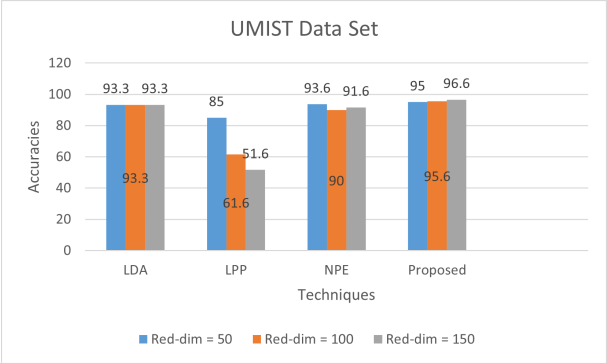

Figure 6. Results on different Training and Test data 320/80 achieves best results as compared to existing methods.

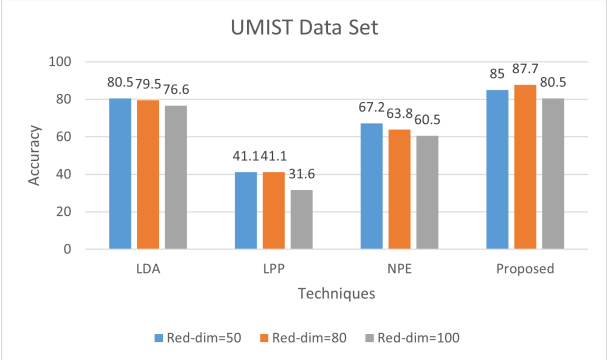

Figure 7. Results on different Training and Test data 200/180 achieves best results as compared to existing methods.

4.6. ORL Database

ORL database used for the experimental setting was proposed by Olivetti Research Laboratory (ORL) whose Director is Prof Andy Hopper FR Eng. It contains 400 images of size 1024 (32 × 32) for each image. ORL database is divided into 40 classes having 10 images respectively. Each image is different from others in terms of varying the lighting, facial expressions (open/closed eyes, smiling/not smiling), and facial details (glasses/no glasses). Some images of a single class are shown in Figure 8.

Figure 8. Example images of ORL database.

4.6.1. Results on ORL Database

We evaluated the performance of our algorithm on different images extracted from the ORL database. We evaluate the result using different training datasets and dimensions. We evaluated the results through a Collaborative representation framework (CR). Results clearly show that the performance of the proposed algorithm is better in terms of accuracy than the comparison algorithm as shown in Table 3 and 9. On different dimensions 50, 80, 100, and 150 when training and test data 240/160, 280/120 are used, and the best results are achieved compared to existing methods, as shown in Figures 10 and 11.

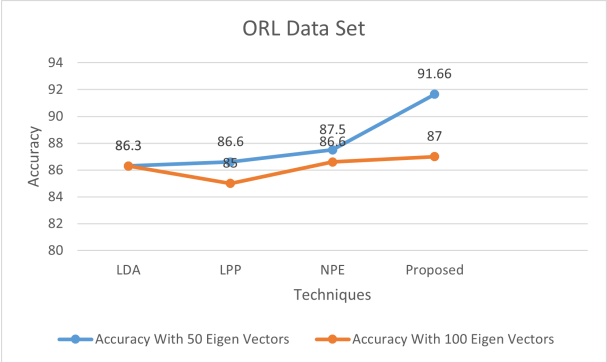

Figure 9. ORL Data Set Accuracy graph.

Table 3. The highest accuracy achieved among comparisons methods on ORL database.

Dimensionality Reduction	**LDA**	**LPP**	**NPE**	**Proposed**	**LDA**	**LPP**	**NPE**	**Proposed**
Red-dim	50	50	50	50	100	100	100	100
classes	40	40	40	40	40	40	40	40
Accuracy	86.3	86.6	87.5	91.66	86.3	85	86.6	87
Correct Prediction	99	102	105	110	99	100	103	104
Training-data	280	280	280	280	280	280	280	280
Testing -data	120	120	120	120	120	120	120	120
Classifier	KNN	KNN	KNN	KNN	KNN	KNN	KNN	KNN

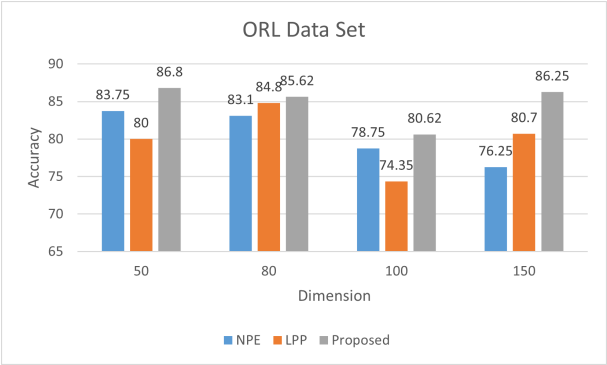

Figure 10. Results on different Training and Test data 240/160 achieves best results as compared to existing methods.

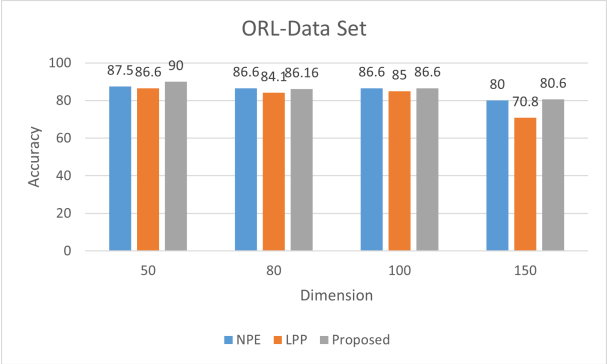

Figure 11. Results on different Training and Test data 280/120 achieves best results as compared to existing methods.

5. Conclusion and Future Work

In this paper, we proposed graph construction using an affinity matrix through collaborative representation (CR framework) and transform data in low dimensional space which is useful for many discriminant tasks. The proposed graph is used to categorize detected faces in a supervised approach. In Dimensionality reduction 'The graph construction process" plays a significant role in transforming the high-dimensional data set into the low dimensional data set. The main problem in the optimal graph construction process is how to set the neighborhood size and weight parameters. In existing methods [11, 32, 44], neighborhood selection, and weight assignment to corresponding edges are done separately, hence it is desirable to simultaneously find adaptive neighborhood and weight assignment under a unified framework.

In this work, we will address the problem of automatic Graph construction which is datum adaptive and incorporates "within class" and "between class" information into the linear representation to learn the optimal projection matrix for dimensionality reduction using the Collaborative representation technique. We also proved that projecting data from a higher dimension to a lower dimension gives better results as compared to higher dimensional data. Calculated results show that the proposed graph construction algorithm gives better accuracy than other graphs construction algorithms including NPE, Local LLE, and LPP. Embedding in reduced space has better results with less computational burden than embedding in high-dimensional feature space.

In the future, our work will address reducing the computational time of CNR, because when the number of samples in each class increases, the construction of a graph becomes computationally expensive because of finding relationships with all other samples in the database. Our research will cover more difficult databases.

Funding: This research received no external funding.
Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.
Data Availability Statement: The data supporting this paper will be made available upon request and are also freely accessible on the internet.
Acknowledgments: The author gratefully acknowledges the invaluable support, patience, and encouragement of their family

and teachers, whose guidance was essential to the completion of this research and manuscript.
Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Closed-form Solution of Equation (2)

The closed-form solution of Equation (2) minimizes the linear reconstruction error calculated against each class. The first term in Equation (2), within brackets, represents the linear reconstruction error calculated against each class. Here, x represents the data point, D represents the dictionary, and β denotes the coefficient computed for each data point. Additionally, x_i represents each data point in the dictionary D. The parameters σ and λ are regularization terms with positive values, introducing sparsity. Finally, β_i^2 is a coefficient vector comprising positive values.

The objective function is defined as follows:

$$\beta = \underset{\beta}{\operatorname{argmin}} \left(\tfrac{1}{2}\|x - D\beta\|_2^2 + \sigma \sum_{i=1}^{n} \beta_i^2 \|x - x_i\|_2^2 + \lambda \|\beta\|_2^2 \right)$$

The above equation is derived by taking the partial derivatives of the objective function and setting them equal to zero, resulting in the following:

$$-D^T(x - Dx) + \sigma \begin{bmatrix} \|x - x_1\|_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|x - x_n\|_2 \end{bmatrix} \beta + \lambda\beta = 0$$

Here,

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{bmatrix}$$

represents an n-dimensional coefficient vector with positive values associated with vector x.

To obtain the optimal solution, replace the x_i coefficients with the corresponding elements of the solution vector x. The final solution is expressed as:

$$\beta^* = \left(D^T D + \sigma \begin{bmatrix} \|x - x_1\|_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|x - x_n\|_2 \end{bmatrix} + \lambda I \right)^{-1} D^T x$$

The diagonal matrix represents the distance between the data point and all other data points in the dictionary. Obtained coefficients are used to build a similarity graph.

References

1. Hua, J.; Wang, H.; Ren, M.; Huang, H. Dimension reduction using collaborative representation reconstruction based projections. *Neuro-Comput.* **2016**, *193*, 1–6.
2. Jadoon, W.; Zhang, H. Locality features encoding in regularized linear representation learning for face recognition. In Proceedings of the 11th International Conference on Frontiers of Information Technology, Islamabad, Pakistan, 16–18 December 2013; pp. 189–194.
3. Yuan, M.-D.; Feng, D.-Z.; Liu, W.-J.; Xiao, C.-B. Collaborative representation discriminant embedding for image classification. *J. Vis. Commun. Image Represent.* **2016**, *41*, 212–224.
4. Zhou, Z.; Waqas, J. Intrinsic structure based feature transform for image classification. *J. Vis. Commun. Image Represent.* **2016**, *38*, 735–744.
5. Waqas, J.; Yi, Z.; Zhang, L. Collaborative neighbor representation based classification using l2-minimization approach. *Pattern Recognit. Lett.* **2013**, *34*, 201–208.
6. Liang, L.; Xia, Y.; Xun, L.; Yan, Q.; Zhang, D. Class-Probability Based Semi-Supervised Dimensionality Reduction for Hyperspectral Images. In Proceedings of the 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 23–25 November 2018; pp. 460–463. https://doi.org/10.1109/ICSESS.2018.8663892.

7. Zhang, L.; Yang, M.; Feng, X.; Collaborative Representation based Classification for Face Recognition. *arXiv* **2012**, arXiv:1204.2358.
8. Arunasakthi, K.; KamatchiPriya, L. A review on linear and non-linear dimensionality reduction techniques. *Mach. Learn. Appl. Int. J.* **2014**, *1*, 65–76.
9. Huang, W.; Yin, H. On nonlinear dimensionality reduction for face recognition. *Image Vis. Comput.* **2012**, *30*, 355–366.
10. Zhang, H.; Gabbouj, M. Feature Dimensionality Reduction with Graph Embedding and Generalized Hamming Distance. In Proceedings of the 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 1083–1087. https://doi.org/10.1109/ICIP.2018.8451089.
11. Yang, W.; Wang, Z.; Sun, C. A collaborative representation-based projections method for feature extraction. *Pattern Recognit.* **2015**, *48*, 20–27.
12. Qiao, L.; Chen, S.; Tan, X. Sparsity preserving discriminant analysis for single training image face recognition. *Pattern Recognit. Lett.* **2010**, *31*, 422–429.
13. Zhou, Y.; Ding, Y.; Luo, Y.; Ren, H. Sparse Neighborhood Preserving Embedding via L2,1-Norm Minimization. In Proceedings of the 2016 9th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 10–11 December 2016; pp. 378–382. https://doi.org/10.1109/ISCID.2016.2096.
14. Chen, S.; Li, S.; Ji, R.; Yan, Y.; Zhu, S. Discriminative local collaborative representation for online object tracking. *Knowl.- Based Syst.* **2015**, *100*, 13–24.
15. Waqas, J.; Zhang, Y.I.; Zhang, L.E.I. Graph-Based Features Extraction Via Datum Adaptive Weighted Collaborative Representation for Face Recognition. *Int. J. Pattern Recognit. Artif. Intell.* **2014**, *28*, 2.
16. Jadoon, W.; Zhang, L.; Zhang, Y. Extended collaborative neighbor representation for robust single-sample face recognition. *Neural Comput. Appl.* **2015**, *26*, 1991–2000.
17. Liu, B.D.; Shen, B.; Gui, L.; Wang, Y.X.; Li, X.; Yan, F.; Wang, Y.J. Face Recognition using Class Specific Dictionary Learning for Sparse Representation and Collaborative Representation. *Neurocomputing* **2016**, *204*, 198–210.
18. Yang, J.; Yu, K.; Gong, Y.; Huang, T. Linear spatial pyramid matching using sparse coding for image classification. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, 20–25 June 2009; pp. 1794–1801. https://doi.org/10.1109/CVPR.2009.5206757.
19. Han, P.Y.; Yin, O.S.; Ling, G.F. Semi-supervised generic descriptor in face recognition. In Proceedings of the 2015 IEEE 11th International Colloquium on Signal Processing and Its Applications (CSPA), Kuala Lumpur, Malaysia, 6–8 March 2015; pp. 21–25.
20. Weinberger, K.Q.; Packer, B.D.; Saul, L.K. Nonlinear dimensionality reduction by semi-definite programming and kernel matrix factorization. Tenth International Workshop on Artificial Intelligence and Statistics (AISTAT 2005). *PMLR* **2005**, *2005*, 381–388.
21. Bhele, S.G.; Mankar, V.H. A Review Paper on Face Recognition Techniques. *Int. J. Adv. Res. Comput. Eng. Technol.* **2012**, *1*, 339–346.
22. Shermina, J. Application of locality preserving projections in face recognition. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, **2010**, *1*(3), 82. http://ijacsa.thesai.org/.
23. Meena, M.K.; Meena, H.K. A Literature Survey of Face Recognition Under Different Occlusion Conditions. In Proceedings of the 2022 IEEE Region 10 Symposium (TENSYMP), Mumbai, India, 1–3 July 2022; pp. 1–6. https://doi.org/10.1109/TENSYMP54529.2022.9864502.
24. Li, W.; Feng, F.; Li, H.; Du, Q. Discriminant Analysis-Based Dimension Reduction for Hyperspectral Image Classification: A Survey of the Most Recent Advances and an Experimental Comparison of Different Techniques. *IEEE Geosci. Remote Sens. Mag.* **2018**, *6*, 15–34. https://doi.org/10.1109/MGRS.2018.2793873.
25. Shinwari, A.R.; Balooch, A.J.; Alariki, A.A.; Abdulhak, S.A. A Comparative Study of Face Recognition Algorithms under Facial Expression and Illumination. In Proceedings of the 2019 21st International Conference on Advanced Communication Technology (ICACT), PyeongChang, Korea, 17–20 February 2019; pp. 390–394. https://doi.org/10.23919/ICACT.2019.8702002.
26. Guo, Z.; Zhang, Z.; Xing, E.; Faloutsos, C. Semi-Supervised Learning Based on Semiparametric Regularization. *SDM* **2008**, 132–142. https://doi.org/10.1137/1.9781611972788.
27. Jebara, T.; Wang, J.; Chang, S.F. Graph construction and b-matching for semi-supervised learning. In Pro-

ceedings of the 26th Annual International Conference on Machine Learning, Montreal, QC, Canada, 14–18 June 2009; pp. 441–448.

28. Wright, J.; Ganesh, A.; Yang, A.; Ma, Y. Robust Face Recognition via Sparse Representation. *TPAMI* **2008**, *in press.*
29. Wu, Y. Vansteenberge Jarich, Masayuki Mukunoki, and Michihiko Minoh. Collaborative Representation for Classification, Sparse or Non-Sparse? *arXiv* **2014**, arXiv:1403.1353.
30. Wang, G.; Forsyth, D.; Hoiem, D. Improved Object Cate- gorization and Detection Using Comparative Object Similarity. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2442–2453.
31. Cheng, B.; Yang, J.; Yan, S.; Fu, Y.; Huang, T. Learning with l1-graph for image analysis. *IEEE Trans. Image Process.* **2009**, *19*, 858–866.
32. Qiao, L.S.; Chen, S.C.; Tan, X.Y. Sparsity preserving projections with applications to face recognition. *Pattern Recognit.* **2010**, *43*, 331–341.
33. Cai, D.; He, X.; Han, J. Semi-supervised Discriminant Analysis. In Proceedings of the 2007 IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 14–17 October 2007; pp. 1–7. https://doi.org/10.1109/ICCV.2007.4408856.
34. Huang, W.; Wang, X.; Zhu, Y.; Li, J. Iterative Collaborative Representation based Classification for Face Recognition. *Signal Process. Res.* **2015**, *4*, 12–19.
35. Gong, G. Variational Quantum Isometric Feature Mapping. In Proceedings of the 2024 4th International Symposium on Computer Technology and Information Science (ISCTIS), Xi'an, China, 12–14 July 2024; pp. 558–563. https://doi.org/10.1109/ISCTIS63324.2024.10699116.
36. Hou, X.; Yao, G.; Wang, J. Semi-Supervised Classification Based on Low Rank Representation. *Algorithms* **2016**, *9*, 48.
37. Maitra, S.; Hossain, T.; Hasib, K.M.; Shishir, F.S. Graph Theory for Dimensionality Reduction: A Case Study to Prognosticate Parkinson's. In Proceedings of the 2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 4–7 November 2020; pp. 134–140. https://doi.org/10.1109/IEMCON51383.2020.9284926.
38. Hu, H.; Feng, D.; Yang, F. A Promising Nonlinear Dimensionality Reduction Method: Kernel-Based Within Class Collaborative Preserving Discriminant Projection. *IEEE Signal Process. Lett.* **2020**, *27*, 2034–2038. https://doi.org/10.1109/LSP.2020.3037460.
39. Yuan, X.T.; Yan, S.C. (2010), "Visual classification with multitask joint sparse representation. *IEEE Trans. Image Process.* **2012**, *21*, 4349-4360.
40. Nie, F.; Wang, Z.; Wang, R.; Li, X. Submanifold-Preserving Discriminant Analysis With an Auto-Optimized Graph. *IEEE Trans. Cybern.* **2020**, *50*, 3682–3695. https://doi.org/10.1109/TCYB.2019.2910751.
41. Qiang, T.; Liu, Z.; Huang, Q.; Zhang, Z.; Chen, Z.; Chen, H. Dimensionality reduction by regularized least squares weighted discriminant projection. In Proceedings of the 2021 CIE International Conference on Radar (Radar), Haikou, Hainan, China, 15–19 December 2021; pp. 2220–2223. https://doi.org/10.1109/Radar53847.2021.10027971.
42. Elhamifar, E.; Vidal, R. Robust Classification using Structured Sparse Representation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, USA, 20–25 June 2011.
43. Huang, W.; Yin, H. Linear and nonlinear dimensionality reduction for face recognition. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; pp. 3337–3340. https://doi.org/10.1109/ICIP.2009.5413898.
44. Van der Maaten, L.J.P.; Postma, E.O.; van den Herik, H.J. *Dimensionality Reduction: A Comparative Review*; MICC, Maastricht University: Maastricht, The Netherlands, 2009.