INTERNATIONAL JOURNAL OF CIRCUITS, SYSTEMS AND SIGNAL PROCESSING
DOI: 10.46300/9106.2020.14.62 Volume 14, 2020

Semi-supervised optimization algorithm based on Laplacian Eigenmaps

Qinjuan Luo*, Jian Wen, Yutong Wu, Mingkai Wang,
The school of Technology,
Beijing Forestry University,
No.35 Tsinghua East Road, Haidian District, Beijing, 100083,
China

Received: March 27, 2020. Revised: May 8, 2020. 2nd Revised: August 6, 2020.
Accepted: August 14, 2020. Published: August 14, 2020.

Abstract—As a member of many dimensionality reduction algorithms, manifold learning is the hotspot of recent dimensionality reduction algorithm. Despite it is good at retaining the original space structure, there is no denying that its effect of classifying still has room for improvement. Based on Laplacian Eigenmap, which is one of the manifold learning algorithm, this paper committed to optimize the algorithm combined with a semi-supervised learning ideas, which can improve the recognition rate. Finally, the better method of two forms is tested in the surface electromyography system and plant leaf identification system. The experimental results show that this semi-supervised method does well in classifying.

Keywords—dimensionality reduction, Laplacian eigenmaps, semi-supervision

I. Introduction

MACHINE learning is widely used in biological information, robot control and other related application fields, showing a broad prospect for development. As one of the key links of machine learning, the dimensionality reduction of high-dimensional data has received extensive attention since the middle of the last century [1-5]. A large number of researchers have proposed different theories and methods based on various mathematical matrices and machine learning ideas, and obtained a large number of research results. Nowadays, with the vigorous development of related technologies such as computers and electronics, data collection technology and data storage technology have been greatly improved, and the era of big data is taking advantage of the trend. The generation of big data also brings high-dimensional data, which provides a reliable basis for humans to explore the essential characteristics of things. However, many high-dimensional or even super-high-dimensional data will cause "dimensionality disaster" in many cases [6]. As a result, data mining technology emerges from time to time[7].

Data dimensionality reduction has great practical application value in traditional pattern recognition and computer vision fields. High-dimensional data will not only increase storage burden and computational cost, but also easily cause dimensionality disasters. On the one hand, the number of features that need to be calculated increases exponentially. On the other hand, when the number of samples is not enough, over-fitting will also occur, which will affect the classification and recognition effect. At this time, it is necessary to apply data dimensionality reduction methods.

There are two main research ideas in data dimensionality reduction, one is feature selection and the other is feature extraction. Feature extraction changes the size and meaning of the original data, and obtains another set of low-dimensional data containing features. This is also the main research content of this article.

In different fields, we often need to retain different characteristic information. In the field of bioinformatics, gene sequence analysis requires categorization and other analysis of hundreds or thousands of dimensional data, and meaningful gene sequences are composed of fragments, which requires us to keep it in the process of dimensionality reduction mapping. Partial information. For some images with prominent geometric structure, we often need to maintain global information to avoid destroying the original characteristics after dimensionality reduction. For these different applications, researchers have done a lot of research and proposed different dimensionality reduction algorithms from different angles to adapt to different applications [8-11].

After the concept of data dimensionality reduction was put forward, more and more data dimensionality reduction algorithms followed. Common classification of dimensionality reduction algorithms is based on different data processing methods, which can reduce data dimensionality Algorithms are divided into linear and nonlinear data dimensionality reduction methods.

The linear method assumes that the data has a low-dimensional linear distribution, and the data satisfies the Gaussian distribution, calculates the global European structure of the image for dimensionality reduction analysis. The representative method is principal component analysis [12] and linear discriminant analysis [13]. Principal component analysis is a multivariate statistical analysis method in which multiple variables are linearly transformed to obtain a small number of important variables. Linear discriminant analysis is similar to the principal component analysis method. It projects

high-dimensional pattern samples into the best discriminant vector space. After projection, it is guaranteed that the pattern samples have the largest inter-class distance and the smallest intra-class distance in the new subspace to achieve The effect of extracting classification information and compressing the dimension of feature space. These linear dimensionality reduction methods have formed a complete theoretical system, and have shown good behavior in applications.

However, since the dimensions of the actual data obtained often have a nonlinear relationship with the dimensions of the essential characteristics, and the data does not necessarily satisfy the Gaussian distribution, the linear dimensionality reduction method can no longer obtain satisfactory results, and it is necessary to use nonlinear Method for dimensionality reduction. There are many methods for nonlinear dimensionality reduction, which can be divided into dimensionality reduction methods based on kernel functions and eigenvalues. Representative kernel-based methods include kernel-based principal component analysis (KPCA) [14], and kernel-based independent component analysis (KICA) [15]. The linearly inseparable data on the dimension is mapped to the high dimension, and then the linear dimension reduction method is used to reduce the dimensionality of the data. The dimensionality reduction method based on eigenvalues is also known as the Manifold Learning algorithm. The dimensionality reduction algorithm of manifold learning was first proposed by ST Roweis and JB Tenenbaum, and it has gradually become a research hotspot of data dimensionality reduction [16]. This type of method assumes that the high-dimensional data is distributed on an essentially low-dimensional nonlinear manifold. By learning the local neighborhood structure of the sample in the high-dimensional space, the low-dimensional manifold is found from the high-dimensional space to achieve dimensionality reduction. simple. Studies have shown that high-dimensional biometric data such as faces, fingerprints, iris, etc. are distributed in a high-dimensional space in a manifold manner [17], which further confirms the advantages of manifold learning in intelligent recognition. The representative algorithms are Isomatric Mapping (Isomap) [18], Locally Linear Embedding (LLE) [19], Laplacian Eigenmaps (LE) [20] Although the implementation methods of these algorithms are different, their implementation principles can be unified into the graphics embedding framework. The difference between them lies in the structure of the graph structure and the calculation of edge weights. That is to say, the focus of the dimensionality reduction method of manifold learning lies in the different construction methods of graph structure.

The above-mentioned traditional manifold learning methods are all unsupervised, and in actual classification problems, the samples projected by these algorithms show the defect of low separability of the data. For this reason, researchers first proposed a batch of extended LLE algorithms to improve the classification performance of LLE algorithms. The traditional manifold learning algorithm uses Euclidean distance or geodesic distance to construct the nearest neighbor graph, and when the label information is introduced, the local inter-class divergence matrix and the local intra-class divergence matrix can be constructed. Supervised LLE (SLLE) algorithm is one of the representatives [21]. This type of algorithm introduces the sample category information of the original data in the LLE learning process to guide the classification of the data. On this basis, researchers have successively proposed Locality Preserving Discriminant Projections (LPDP) [22], Linear Discriminant Projection (LDP) [23], and other supervised manifold learning algorithms. LPDP effectively introduces class label information, but lacks attention to outlier data; LDP can pay attention to class label information while maintaining the locality of submanifolds, and the robustness of outliers is also stronger.

Since most of the data in reality have nonlinear characteristics, the selection process mainly considers the manifold learning algorithm or first uses the linear dimensionality reduction method to eliminate redundancy, and then uses the nonlinear method for projection recognition. It can be known from the experimental experience of previous researchers that the use of supervised dimensionality reduction methods, that is, adding label information in training, can achieve satisfactory results in classification. Manifold learning cannot directly obtain the spatial mapping from high-dimensional to low-dimensional, so how to construct the relationship matrix between new data (unlabeled data) and known data is a difficult point of thinking.

The main research content of this paper is to summarize and analyze the classic dimensionality reduction algorithm, and select appropriate evaluation indicators to evaluate the performance of the classic dimensionality reduction algorithm. In addition, in accordance with the research hotspots of dimensionality reduction algorithms, this article is also committed to implementing a semi-supervised dimensionality reduction method with improved classification performance and applying it to gesture recognition systems based on surface EMG signals and plant leaf recognition systems.

II. Methods

A. Laplacian eigenmaps

The basic idea of LE is that in high dimensional space, if the distance between two sample points is very small, then when the two points are projected into a low dimensional space, their corresponding image distance is also small. The weighted distance between each point and the image is defined as a loss function, and then the loss function is solved by the spectral properties of the Tula Plath operator. The optimal low dimensional representation can be obtained under the condition of preserving some local information of the original dataset.

First, construct the adjacency graph and define the adjacency matrix *W*. In order to minimize the distance between the connected points, we need to satisfy the following objective functions:

$$\begin{aligned} E &= \min \sum_{i,j} \left\| y_i - y_j \right\|^2 W_{i,j} \\ &= 2tr\left[Y^T \left(D - W \right) Y \right] \end{aligned} \tag{1}$$

Where W is the weight matrix, which can be calculated using a thermonuclear function:

$$W_{i,j} = e^{\frac{\|x_i - x_j\|^2}{t}} \quad (2)$$

It can also be simplified to:

$$W_{i,j} = \begin{cases} 1 & i, j \text{ is relevant} \\ 0 & i, j \text{ is not relavant} \end{cases} \quad (3)$$

Here we use the properties of the Laplacian matrix to solve. Make $L = D - W$, the rewrite (1) is:

$$\min tr\left(Y^T LY\right), s.t. Y^T DT = 1 \quad (4)$$

Next, solve the objective function:

$$LY = -DY\Lambda \quad (5)$$

Where Λ is a diagonal matrix, L and D are symmetric matrices, which can be considered as solving generalized eigenvalues. The eigenvector corresponding to the first d non-zero minimum eigenvalues is the dimensionally reduced low-dimensional data.

B. Semi-supervised Laplacian eigenmaps based on probability theory

When dealing with data with category information and data without category information, the first simple idea is to use supervised dimensionality reduction on data with category information, and still use the original unsupervised dimensionality reduction on data without category information. mode. However, such an approach cannot stand scrutiny. To map the two separately, it is difficult to ensure that we map to the same low-dimensional space. It is meaningless to do so even if you get a good recognition result in the end. Therefore, in order to introduce the idea of probability to achieve semi-supervised dimensionality reduction, this paper proposes two solutions.

First, the data with category information is dimensioned in a supervised manner, and then the data of the unclassified markers is analogically mapped using probability coefficients. The low-dimensional map obtained by dimensionality reduction is multiplied by the probability coefficient to represent the space without the category information, where the probability coefficient is represented by the distance, that is:

$$P(A,H) = \frac{distance(A,H)}{distance(A,H) + distance(A,J)} \quad (6)$$

Where A is a data point that does not contain category information, and H and J are respectively two nearest neighbor data points with category information at point A.

When this probability idea is added to the LE algorithm, there are two ways to deal with it. One is to construct a new neighbor graph when calculating the k-nearest neighbor distance:

$$G(A) = P(A,H) \times G_b(H) + P(A,J) \times G_b(J) \quad (7)$$

Where G(H) and G(J) represent their distance matrices with k-nearest neighbors, respectively.

The second way to deal with it is to use probability directly as a coefficient to calculate a low-dimensional projection of unlabeled data:

$$dr\_data = P(A,H) \times DR\_data(H) + P(A,H) \times DR\_data(J) \quad (8)$$

Where DR_data refers to the low-dimensional mapping after the dimensional reduction of the data points with category tags.

This way of processing is to directly express the low-dimensional space of the sample data without class-marking information using similarity relations. Similarly, this is also an assumption in principle, using the similarity relationship to construct a virtual data point projection.

C. Semi-supervised Laplacian eigenmaps based on sample class center point idea

In 2005, Costa et al. [24] proposed a theory when discussing the dimension reduction based on manifold learning. Assuming each class has a center point, we can determine whether the data point belongs to this class by judging whether the data point is close to the center of the class. Because this method has some defects, we can't completely rely on the class center point, but we can save some of the original LE calculations. Thus, we define a new distance expression.

Define the center point as:

$$c_k = \frac{\sum P_k}{n}$$

Where c_k represents the coordinates of the center point of type k, P_k is the coordinates of the samples of type k, and n is the total number of the samples of type k.

So the original cost function is optimized as:

$$E = \alpha \sum_{i,j} W_{ij} \|y_i - y_j\|^2 + \sum_{k,l} V_{kl} \|c_k - y_i\|^2 \quad (9)$$

Where W_{ij} is the weight matrix, α represents the distribution coefficient, and when α is 0, the equation (9) is completely consistent with the unsupervised LE. c_k represents the coordinates of the center point and $c_k = \frac{\sum p_k}{n}$. V_{kl} is the relationship matrix:

$$V_{kl} = \begin{cases} 1 & if\ x_i \in k; \\ 0 & others. \end{cases} \quad (10)$$

In order to find the low-dimensional map using the properties of the Laplacian matrix [10], we make $Q = \begin{bmatrix} I & D \\ D^T & \alpha W \end{bmatrix}$, where $D_{kl} = V_{kl} \|c_k - y_l\|^2$, $m_n = [c_1, \cdots, c_k, y_1, \cdots, y_l]$, so we will write (9) as:

$$E = \min \sum_{i,j} \|m_i - m_j\|^2 Q_{i,j} \quad (11)$$

In general, the semi-supervised Laplacian feature map dimensionality reduction method based on the idea of class

center point is to adjust and optimize the objective function to a certain extent. Changed the weight matrix of two adjacent sides when constructing the neighbor graph. This reduces the degree of interference from different types of sample points during mapping.

III. Expetiment

In order to verify the effectiveness of the two optimization algorithms, we experimentally verified them on the five datasets. A detailed description of the datasets is given in Table Ⅰ. The final experimental results were expressed by the recognition rate of the SVM classifier.

Table Ⅰ: The description of the experimental data set

Dataset	Total number of	Sample dimension	Number of categories
USPS	1000	256	10
Wine	144	12	3
Iris	160	4	3
Libras_Movement	360	89	15
DNA	840	180	3

A. The experiment of semi-supervised Laplacian eigenmaps based on probability theory

We first experimented with the first processing method, repeating 10 times, and comparing with the recognition rate of unsupervised Laplacian feature mapping under the same parameters. In order to be more realistic, we reduced the training data of the supervised algorithm to 70%. The specific experimental parameters and experimental results are shown in Table Ⅱ:

Table Ⅱ: The results of solution 1

	dim	*sigma*	*k*	Semi-super-visory (%)	Supervisi-on (%)	Unsuper-vised (%)
USPS	25	2	12	77.5	77.0	89.0
Wine	5	2	5	66.7	80.6	72.1
Iris	2	2	5	45.4	96.7	70.0
Libras_ Movem-ent	25	2	3	53.4	36.7	40.0
DNA	40	5	30	40.1	53.8	77.5

It can be seen from the data in Table 2 that the implementation of this semi-supervised method is higher than the unsupervised method in the Libras_Movement data set, and the recognition rate is slightly lower than the unsupervised in the other four data sets. And compared with the supervised Laplace, this method has no advantage. Experiments show that there is a large error between the feature data obtained by this method and the actual sample features, which means that the similarity of sample categories does not play a key role in the reconstruction of neighbor graphs.

Next, we try to obtain the projection matrix in the low-dimensional space directly from the probability value as the proportional coefficient. The experiment was repeated 10 times, and the average recognition rate of 10 experiments was taken as the final result. The results are shown in Table Ⅲ:

Table Ⅲ: The results of solution 2

	dim	*sigma*	*k*	Semi-super-visory (%)	Supervi-sion (%)	Unsuper-vised (%)
USPS	25	2	12	83.5	77.0	89.0
Wine	5	2	5	83.3	80.6	72.1
Iris	2	2	5	96.7	96.7	70.0
Libras_Movem-ent	25	2	3	56.3	36.7	40.0
DNA	40	5	30	67.5	53.8	77.5

It can be seen from Table 3 that the implementation method is outstanding in the Wine dataset and the Iris dataset, and is much better than the unsupervised dimensionality reduction method, and has a slight advantage compared with the supervised method. Although this method does not achieve a high recognition rate on the Libras_Movement data set, it has also been improved to some extent. The recognition results on the USPS data set and the DNA dataset substantially maintain the original recognition level. From the experimental data, the method can obtain good dimensionality reduction effect on some data sets and proves the feasibility of the scheme under certain conditions.

B. The experiment of semi-supervised Laplacian eigenmaps based on sample class center point idea

Then we experimentally verify the second processing method, and obtain the experimental results of the method on the five datasets, and compare it with the unsupervised Laplacian feature mapping method. The experimental parameter setting and recognition results are all shown in Table Ⅳ.

Table Ⅳ: The applications of semi-LE on 5 datasets

	dim	*sigma*	*k*	Semi-su-pervisor y (%)	Supervi-sion (%)	Unsuper-vised (%)
USPS	25	2	0.5	12	89.5	89.0
Wine	11	5	0.1	5	86.7	75.0
Iris	3	10	10	2	80.0	75.6
Libras_M-ovement	40	2	1	3	75.0	65.1
DNA	40	5	5	30	85.4	77.5

From Table 4, we can see that the optimization algorithm based on the class center point has obvious advantages compared with the unsupervised method. It still maintains a high recognition rate in the USPS data with high recognition

rate and increases nearly 10% in the Libras_Movement data with low original recognition rate, which proves the reliability of the method to some extent.

The dim parameter of this experiment selects the value when the highest recognition rate is obtained in the previous experiment. Since there are two other variable parameters α and sigma in this method, this section uses the Wine dataset as an example to carry out two sets of control variable experiments.

First, the sigma parameter is controlled to 2, and the α parameter is changed. The obtained result is shown in Figure 1:

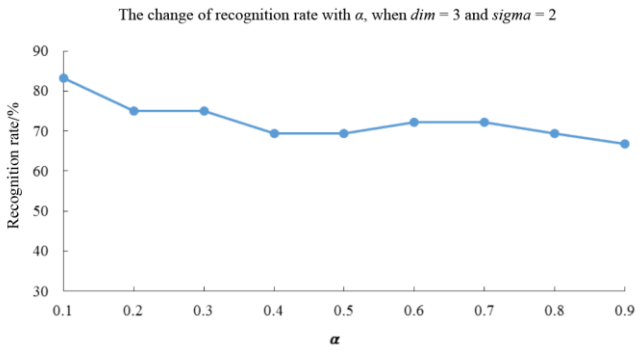

Fig.1 The recognition rate varies with α when dim=3, sigma=2

It can be concluded from Fig. 1 that when α changes from 0.1 to 0.9, the average recognition rate is 72.51%, which is significantly higher than the original unsupervised Laplacian feature mapping algorithm. The maximum value of the recognition rate occurs when α = 0.1, which is as high as 83.3%.

The value of α is kept at 0.1, the value of *dim* is 3, and the value of sigma is changed to obtain a change in the recognition rate.

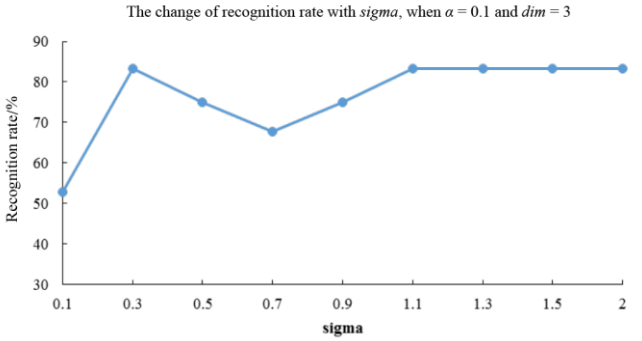

Fig.2 The recognition rate varies with sigma when dim=3, α =2

It can be seen from Fig. 2 that when both dim and α remain unchanged, the recognition rate does not change much with the increase of sigma, and the approximate recognition rate can reach 83.3%, which is also larger than the unsupervised LE. So we speculate that the value of sigma has little effect on the dimension reduction effect. In the subsequent experiments, the *sigma* takes a fixed value of 2. At the same time, the values of α and dim were changed and experiments were carried out. Finally, we found that when α=1 and *dim*=2, the recognition rate is the highest, up to 87%, which is a certain improvement compared with the highest 76% in the unsupervised algorithm.

C. Results and analysis

For the first semi-supervised dimensionality reduction method based on probability theory, from the experimental results, the first processing method does not get satisfactory classification recognition effect. The reason for this phenomenon is that the projection of the structure is not very consistent with the actual situation. Secondly, the label data and the unlabeled data with different dimensionality reduction are not projected in the same low-dimensional space. The second treatment method has obtained satisfactory results, and the recognition rate on the two data sets has been significantly improved. For the second semi-supervised algorithm based on class-center points, the results of the experimental verification on the Wine dataset show that the effect of classification recognition is significantly improved compared with the unsupervised algorithm, and the effectiveness of the algorithm is proved initially. In addition, the control variables were also tested on the two parameters α and *sigma* that affect the dimensionality reduction effect. We find that when α changes from 0.1 to 0.9, the average recognition rate is significantly higher than that of the unsupervised LE algorithm, but the value of *sigma* has little effect on the dimensionality reduction effect. Joint experiments show that when α=1 and *dim*=2, the dimension reduction effect is the best.

D. Engineering example

To further verify the practicability of the semi-supervised method based on class-center points, we performed experiments on sEMG-based gesture recognition systems and plant leaf recognition systems, respectively. The experimental results were compared with the results of the unsupervised LE algorithm. The evaluation indicator still selects the recognition rate of the SVM classifier.

E. Application on gesture recognition system based on surface EMG signals

The sEMG signal control system determines the action by extracting the characteristics of the surface EMG signal, thereby realizing the control of the device. The experimental data is derived from the laboratory subject of surface electromyography signal classification for forestry machine gesture control [25]. Through the work of the previous research team on the EMG signal, we finally obtained 900×252 matrix data as the dimension reduction object. In order to verify the reliability of the Laplacian optimization algorithm, we randomly selected 80% of these data as training data and 20% as testing data for experimental analysis. A total of five comparison experiments were carried out. The parameters of the SVM were set to 1 and 0.007, respectively. The other experimental parameters were selected by two sets of control variables: *dim*=100, α=10, *sigma*=2, *k*=12. The experimental results are shown in Figure 3.

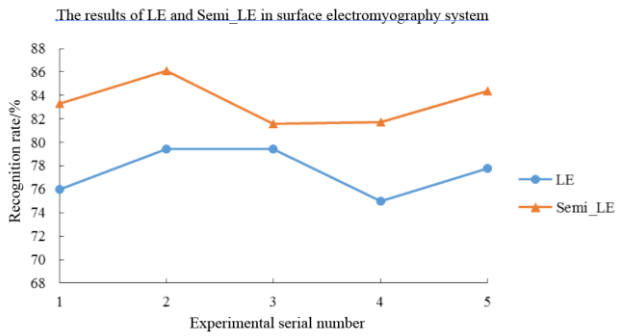

Fig. 3 The results of LE and Semi_LE in surface electromyography system

From Figure 3 we can see that the semi-supervised LE method has 5.9% advantages compared with the unsupervised dimensionality reduction method. It can be said that the semi-supervised LE algorithm based on the idea of the class center point has been successfully applied in the gesture recognition system based on surface EMG signals, and achieved good results.

In order to further verify the reliability of the improved algorithm in gesture recognition based on surface EMG signals, this section uses the isometric mapping (ISOMAP) [7] of the manifold learning algorithm to perform dimensionality reduction experiments. The results are shown in table V.

Table V the results of three-dimensionality reduction methods in surface electromyography system

Recognition rate（%）	1	2	3	4	5	Average
LE	76.0	79.4	79.4	75.0	77.8	77.5
Semi_LE	83.3	86.1	81.6	81.7	84.4	83.4
ISOMAP	81.1	83.3	82.6	79.8	83.0	82.0

It can be seen from the data in Table 5 that the optimized supervised LE algorithm has certain advantages compared with the unsupervised LE or the ISOMAP

F. Application on plant leaf recognition system

The plant leaf recognition system mainly extracts the features by processing the image of the plant leaves and then obtains the recognition result through the classification learning algorithm. The experimental data comes from the plant leaf data collected in the early stage of the laboratory plant leaf recognition system project. We selected 32 kinds of blades, 50 images of each blade, and selected the Hu invariant moment, wavelet features and Fourier features extracted in the early stage of the laboratory to form a feature matrix of 1600×613. We randomly selected 90% as the training data and 10% as the test data for the experimental analysis. The experiment was repeated five times. The experimental parameters were also selected through two sets of control variable experiments: dim=200, α=0.5, $sigma$=2, k=8, and the experimental results are shown in Figure 4.

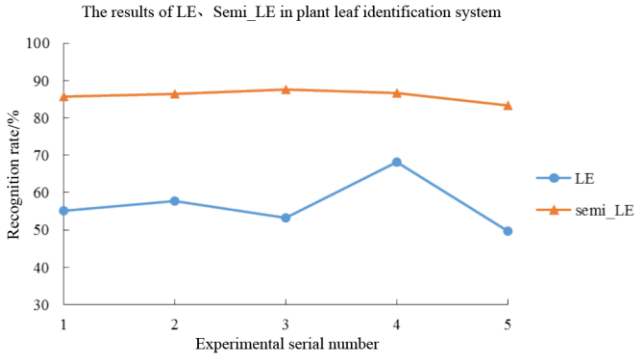

Fig. 4 The results of LE、Semi_LE in plant leaf identification system

From Figure 4, we can see that the semi-supervised LE algorithm is significantly improved compared with the unsupervised LE, and its average recognition rate has a stable advantage as a whole. Compared with the unsupervised method, the semi-supervised method is much better on the plant leaf system, but from the application point of view, the recognition rate is not high when the feature dimension is not much reduced, partly because of the problem of feature calculation, you can replace or add new features to improve recognition rate.

In view of the fact that the isometric mapping (ISOMAP) feature extraction algorithm and the Laplacian feature extraction algorithm are both manifold learning algorithms. Therefore, in order to further verify the reliability of the improved algorithm, isometric mapping is used to perform dimensionality reduction experiments on samples. The experimental results are shown in Table VI

Table VIthe results of three-dimensionality reduction methods in plant leaf identification system

Recognition rate（%）	1	2	3	4	5	Average
LE	55.1	57.7	53.3	68.1	49.7	56.8
Semi_LE	85.6	86.3	87.5	86.6	83.3	85.7
ISOMAP	3.2	13.7	5.4	11.3	9.8	8.7

It can be seen from the data in Table 6 that the supervised Laplacian feature mapping method has a relatively stable and obvious advantage compared with the unsupervised Laplacian feature mapping and the isometric mapping method, which further illustrates the rationality of the improved algorithm.

IV. Conclusion

In this paper, based on LE algorithm and semi-supervised ideas, the optimization algorithms based on probability theory and based on the sample class center point are presented. The experimental results show that the optimization algorithm based on probability theory has little difference in dimension reduction compared with the unsupervised method, and the optimization algorithm based on class center point has better overall improvement than the unsupervised method. The optimization algorithm based on the center point is applied to

the gesture recognition system based on sEMG signals and the plant leaf recognition system. The results show that the recognition rate of sEMG signals is increased by 5.9%, and the recognition rate of plant leaves is also significantly improved. The reliability of the method in actual engineering is further confirmed.

But the research of this article still has some parts that need to be improved. This article does not explore the influence of the number of neighbors in the LE algorithm on the experimental results, but only selects the parameter based on accumulated experience. Experiments in this article show that some parameters in the algorithm have linear effects on the classification effect, but some are irregular, and the settings of these parameters, especially the prediction of eigendimensions, are very different for different data sets. According to It is challenging for experimenters to choose parameters through experience or through multiple experiments. Therefore, in future research, you can consider reducing the parameter settings, or obtaining adaptive parameter values through a certain learning algorithm. In addition, many methods currently perform dimensionality reduction on the same amount of different types of data, which wastes a large part of the data we collect. Therefore, it is urgent to propose a dimensionality reduction method for uneven samples. For the consideration of computational cost, real-time online dimensionality reduction is also the next further work goal.

REFERENCES

[1] Symeonidis P . Content-based Dimensionality Reduction for Recommender Systems[C]// Data Analysis, Machine Learning & Applications- Conference of the Gesellschaft Für Klassifikation Ev, Albert-ludwigs-universitt Freiburg, March. DBLP, 2008.

[2] Cheng W , Ng C A . Using Machine Learning to Classify Bioactivity for 3486 Per- and Polyfluoroalkyl Substances (PFASs) from the OECD List[J]. Environmental Science & Technology, 2019, 53(23):13970-13980.

[3] Shamir, Lior, Yerby, et al. Classification of large acoustic datasets using machine learning and crowdsourcing: Application to whale calls.[J]. Journal of the Acoustical Society of America, 2014.

[4] Junghyun Namkung, Robert C. Elston, Jun-Mo Yang,等. Identification of gene-gene interactions in the presence of missing data using the multifactor dimensionality reduction method[J]. Genetic Epidemiology, 2009.

[5] Prathan Chommaungpuck, Siwanu Lawbootsa, Jiraphon Srisertpol, Fault Detection of Linear Bearing in Auto Core Adhesion Mounting Machine Using Artificial Neural Network, WSEAS Transactions on Systems and Control, pp. 31-42, Volume 14, 2019.

[6] Saquer J, Saquer J, Saquer J. AN APPROACH TO MAKING DYNAMIC PROGRAMMING EASIER FOR STUDENTS IN THE COMPUTER SCIENCE CURRICULUM[J]. 2016, 2(4):157-165.

[7] Xiao J , Zhou Z , Hu D , et al. Self-organized Locally Linear Embedding for Nonlinear Dimensionality Reduction[C]// International Conference on Natural Computation. Springer, Berlin, Heidelberg, 2005.

[8] Renard, Nadine, Bourennane, et al. Dimensionality Reduction Based on Tensor Modeling for Classification Methods.[J]. IEEE Transactions on Geoscience & Remote Sensing, 2009.

[9] Wu G , Feng T T , Zhang L J , et al. Inexact Krylov Subspace Algorithms for Large Matrix Exponential Eigenproblem from Dimensionality Reduction[J]. Mathematics, 2015.

[10] Raymer M L , Punch W F , Goodman E D , et al. Dimensionality reduction using genetic algorithms[J]. IEEE Transactions on Evolutionary Computation, 2000, 4(2):164-171.

[11] Zhang T , Chen W , Li M . Classification of inter-ictal and ictal EEGs using multi-basis MODWPT, dimensionality reduction algorithms and LS-SVM: A comparative study[J]. Biomedical Signal Processing and Control, 2019, 47(JAN.):240-251.

[12] JOLLIFFE I T. Principal component analysis[M]. Berlin:Springer Series in Statistics. Springer-Verlag, 1986.

[13] DUDA RO, HART P E, STORK D H. Pattern classification[M]. 2nd ed. [S.1.]: Wiley Interscience, 2000.

[14] SCHOLKOPF B, SMOLA A, MULLER K. Nonlinear component analysis as a kernel eigenvalue problem[J]. Neural Computation, 1998, 10(5): 1299-1319.

[15] Yu X, Hu D, Xu J. Kernel independent component analysis[J]. Journal of Machine Learning Research, 2003, 4(1):1-48.

[16] Lin T , Zha H . Riemannian Manifold Learning[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2008, 30(5):796.

[17] Weinberger K Q, Saul L K. Unsupervised Learning of Image Manifolds by Semidefinite Programming[J]. International Journal of Computer Vision, 2006, 70(1):77-90.

[18] TENENBAUM J B，SILVA V，UNGFORD J C. A global geometrie framework for nonlinear dimensionality reduction[J]. Science, 2000, 290(12): 2319-2323.

[19] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.

[20] BELKIN M, NIYOGI P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. Neural Computation, 2003, 15(6): 1373-1396.

[21] Turchetti C, Falaschetti L. A manifold learning approach to dimensionality reduction for modeling data[J]. Information Sciences, 2019.

[22] Gui Jie, Jia Wei, Zhu Ling, et al. Locality Preserving Discriminant Projections for Face and Palmprint Recognition[J]. Neurocomputing, 2010, 73(13/15): 2696-2707.

[23] Li T , Zhu S , Ogihara M . Topic hierarchy generation via linear discriminant projection[C]// ACM, 2003.

[24] Raich R, Costa JA, Damelin SB, et al. Classification Constrained Dimensionality Reduction[C]// IEEE International Conference on Acoustics. IEEE, 2005.

[25] Hu XH, Li WB, Kan JM. Gesture control technology based on surface electromyography for automatic pruning machine. Journal of Beijing Forestry University 2017; 117-124 (article in Chinese with an abstract in English).

Qinjuan Luo, with the school of Technology, Beijing Forestry University, Beijing,100083, China (corresponding author, e-mail: luoqinjuan@126.com).

Jian Wen, with the school of Technology, Beijing Forestry University, Beijing, 100083, China (e-mail: wenjian@bjfu.edu.cn).

Yutong Wu, with the school of Technology, Beijing Forestry University, Beijing, 100083, China (e-mail: wuyutong_1996@bjfu.edu.cn).

Mingkai Wang, with the school of Technology, Beijing Forestry University, Beijing, 100083, China (e-mail: misslittlesheep@163.com).

This research is financially supported by the National Natural Science Foundations of China (Grant No.31600588 and Grant No.31600589) and the Fundamental Research Funds for the Central Universities (NO.2015ZCQ-GX-01, NO. 2017ZY27). The authors have no conflict of interest to declare.