*Journal of Network Intelligence* ©2024 ISSN 2414-8105
**Taiwan Ubiquitous Information** Volume **9**, Number **4**, November **2024**

# Image Data Classification Algorithm Based on Spatial Downscaling and Structural Information

Dong-Hua Zheng*
School of Information Technology and Engineering
Guangzhou College of Commerce, Guangzhou 511363, P. R. China
Graduate School
Management and Science University, Shah Alam 40100, Selangor, Malaysia
20190257@gcc.edu.cn

Li-Zhu Ye
School of Information Technology and Engineering
Guangzhou College of Commerce, Guangzhou 511363, P. R. China
Graduate School
Management and Science University, Shah Alam 40100, Selangor, Malaysia
20180160@gcc.edu.cn

*Corresponding author: Dong-Hua Zheng
Received January 16, 2024, revised April 22, 2024, accepted August 2, 2024.

Abstract. *Traditional image data classification algorithms often pull the original image into vectors, resulting in the destruction of spatial structure, which leads to poor classification accuracy. Aiming at the above problems, this paper offers an image data classification algorithm on the ground of Spatial Degradation & Structural Information (SDSI). Firstly, for the traditional Multilinear Principal Component Analysis (MPCA) algorithm ignoring the problem of data mean value information, the feature space representation model and projection direction of all data are obtained through model merging, so as to optimize MPCA. then based on the above optimized MPCA algorithm, spatial dimensionality reduction is performed on image data, using the matrixed Euclidean distance of the image to directly pre-embed the spatial structural information on the original image, combining with the image restructuring, and introducing the bilateral filter image smoothing method is introduced as the image smoothing strategy, and for the reorganized and downsized image matrix, the intra-class scatter is minimized while the inter-class scatter is maximized to separate the different classes as much as possible and ensure the global optimality of the solution. Finally, the performance of the algorithm is estimated on the FERT dataset, and the experimental results indicate that the accuracy, precision, and recall of the SDSI algorithm are 0.925, 0.948, and 0.927, respectively, which effectively improves the accuracy, precision, and recall of classification.*
**Keywords:** Image classification; spatial dimensionality reduction; structural information; image reorganization; bilateral filtering

1. **Introduction.** With the maturity of artificial intelligence technology, people's lives are changing day by day with unprecedented changes. The era of big data is generating a huge volume of image data information every moment, and a variety of image data information fills every corner of our life. Image data classification has a wide range of applications, whether in the field of transportation, security, or medicine, image data classification can help people live conveniently and quickly [1, 2]. Traditional image data classification techniques often use a combination of principal component analysis and

support vector machine, but the image is two-dimensional or three-dimensional data, if the traditional way to classify these image data for research, the data need to be pulled into the form of vectors and then processed [3]. This processing not only produces high-dimensional vectors, but also destroys the inherent higher-order structure and intrinsic correlation in the image data, which leads to poor classification accuracy and small sample problems. Therefore, the research work on how to make full use of spatial structure information to improve image classification performance is of great research significance.

1.1. **Related Work.** For image spatial structure information utilization strategy, Wang et al. [4] proposed an IMED classification algorithm by embedding spatial structure information into Euclidean distance. Di and Crawford [5] extended IMED by using it for multi-angle gender classification and obtained higher classification accuracy. In order to maintain the spatial structure information between pixel points as much as possible, Wang and Chen [6, 7] designed a series of classifiers by bilinear projection [8, 9] of images. Ding et al. [10] designed a bilinear support vector machine (bilinear SVM) by factorizing the regression matrix into two low-rank matrices. Hou et al. [11] utilized the multi-rank left and right projection vectors to construct decision boundaries and create interval functions to propose a multi-rank multi-linear SVM (MRMLSVM). Hossain et al. [12] improved it by proposing a similar bilinear framework for image classification, which improved the classification accuracy. Zheng et al. [13] utilized the matrix kernel paradigm as a convex approximation of the matrix rank to propose a new model SupportMatrix Machines (SMM) for matrix classification problems. Kramer et al. [14] used mapping ordered information to corresponding values to derive a distance-sensitive, predictive ordered labeled image classifier. Lei et al. [15] made a theoretical study of ordered regression and applied the principle of structural risk minimization to image classification. Sun et al. [16] introduced linear discriminant analysis in the framework of ordered regression, and proposed a discriminative method based on ordered regression. Tian et al. [17] summarized the existing strategies of utilizing the spatial structural information at that time and embedded them into the platform of ordered regression, but the accuracy rate of classification was low.

So far, spatial dimensionality reduction methods have been successfully extended to the field of image data classification. Scholars used Principal Component Analysis (PCA) algorithms to first pull tensor data into vectorial data and then perform dimensionality reduction on them. To address this problem, Yang et al. [18] proposed a 2DPCA-based image data classification algorithm to reduce the dimensionality of the second-order vector data of the image matrix. Huang et al. [19] combined the Multilinear Principal Component Analysis (MPCA) algorithm with the support of higher-order vector machines to propose an MPCA-based image data classification algorithm. Han et al. [20] combined online learning with MPCA algorithm and proposed an online multilinear principal component analysis algorithm, thus solving the problem of long running time of the algorithm. However, Zhao and Du [21] pointed out that this method only partially utilizes the spatial information of the image in the same row or column, and does not make full use of the spatial information in the whole image. Gao et al. [22] proposed a face recognition method based on the Euclidean distance of the image with the Two-Dimensional Maximum Local Variation (2DMLV), obtaining higher classification accuracy than 2DPCA. Zhu et al. [23] proposed an Implicit Spatial Regularization (ISR) strategy that is different from the explicit spatial regularization, but it is computationally expensive, and also does not guarantee the global optimality of the solution, resulting in poor classification accuracy.

1.2. **Motivation and contribution.** Although most of the existing spatial dimensionality reduction methods have been applied in the field of image data, however, almost none of them consider the utilization of spatial structural information of the image when they operate on the image directly. To address the above problems, this paper proposes an image data classification algorithm based on Spatial Dimensionality and Structural Information (SDSI). Firstly, for the existing MPCA (two-dimensional principal component analysis) algorithm which cannot handle multiple data simultaneously and ignores the mean value information of data, the MPCA algorithm is optimized by describing the data space with multiple feature space models. Then in order to overcome the inadequacy of the utilization of the existing spatial structure information, based on the optimized MPCA, implicit regularization restructures the original image divisions to reflect the smoothness of the column vectors, which implicitly utilizes the spatial structure information of the image and ensures the global optimality of the understanding. Secondly, feature extraction is performed on the reorganized image data to maintain its smoothness property while extracting features in the row direction. Finally, the results of simulation experiments show that the SDSI algorithm has higher classification performance and efficiency compared with the comparison algorithms.

## 2. Relevant theoretical analysis.

2.1. **Multilinear principal component analysis.** MPCA is a specialized method for spatial dimensionality reduction and feature extraction in image form [24]. In the process of feature extraction, MPCA can not only preserve the high-dimensional structure of the data, but also reduce the arithmetic memory requirements, especially in the processing of high-order vector data, it can get a very good effect of dimensionality reduction.

Create the set of $N$ vectors $x_1, x_2, \ldots, x_N$ used for training, assuming each vector $x_n \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_M}$, where $I_m$ is the dimension of the $m$-th mode of the vector. The purpose of MPCA is to find a multilinear transformation:

$$\left\{ \tilde{V}^{(m)} \in \mathbb{R}^{I_m \times Q_m}, I_m \geq Q_m, m = 1, 2, \ldots, M \right\} \tag{1}$$

Then map the original vector $x_j$ in space $\mathbb{R}^{I_1 \times I_2 \times \cdots \times I_M}$ to space $\mathbb{R}^{Q_1 \times Q_2 \times \cdots \times Q_M}$, where $I_m$ is the dimension of the $m$-th mode of the vector.

$$y_j = x_j \times_1 \tilde{V}^{(1)T} \times_2 \tilde{V}^{(2)T} \times \cdots \times_M \tilde{V}^{(M)T} \tag{2}$$

Finally, the reduced vector $y_j$ is obtained, where the vector $y_j$ captures the main changes in the original vector data that can be observed in $\mathbb{R}^{Q_1 \times Q_2 \times \cdots \times Q_M}$-space. In other words, the purpose of MPCA is to find $M$ projection matrices $\tilde{V}^{(m)}$ to maximize the scatter of the new vector set $\phi_y$.

$$\left\{ \tilde{V}^{(m)} \in \mathbb{R}^{I_m \times Q_m}, I_m \geq Q_m, m = 1, 2, \ldots, M \right\} = \arg\max \left\{ \tilde{V}^{(1)}, \tilde{V}^{(2)}, \ldots, \tilde{V}^{(M)} \right\} \phi_y \tag{3}$$

For Equation (3), the MPCA algorithm uses alternating iterations to solve the projection matrix, i.e., first fix $\tilde{V}^{(1)}, \ldots, \tilde{V}^{(m-1)}$, and then add the solved $\tilde{V}^{(m)}$ to the fixed sequence to solve $\tilde{V}^{(m+1)}$. In this way, the iterative updating is continuous until convergence, and then $M$ projection matrices maximizing the dispersion of the new vector set $\phi_y$ can be obtained.

2.2. **Subspace smooth learning method.** Spatial learning methods are relied on the vector pattern of the image, and thus ignore the spatial structure information inherent in the image matrix [25]. Structured Sparsity Learning (SSL) is a regularized subspace learning model, which uses regularization to penalize the relevant objective function, so as to make the optimization result of the objective function as spatially smooth as possible, thus compensating for the loss of spatial information caused by vectorization. The objective function in a general image classification algorithm is as follows.

$$\arg\max_{b} \frac{b^T XZXb}{(1-\alpha)b^T KXK^T b + \alpha L(b)} \tag{4}$$

where $b$ is the projection vector to be optimized, the regularization factor $0 \leq \alpha \leq 1$ controls the smoothness, and $L$ is a discrete Laplacian regularization function.

$$L(B) = \Delta \cdot b^2 = b^T \Delta^T \Delta b \tag{5}$$

The discrete approximation $\Delta \in \mathbb{R}^{s\times c}$ of the two-dimensional Laplacian operator in Equation (6) can be expressed as follows:

$$\Delta = D_1 \otimes I_2 + I_1 \otimes D_2 \tag{6}$$

where $I_1$ and $I_2$ are the unit matrices of $s \times s$ and $c \times c$, respectively; $\otimes$ denotes the inner product; $D_1(D_2)$ is a $s \times s(c \times c)$ second order gradient smooth operator or matrix in the image row (column) direction.

$$D_i = \frac{1}{w_i^2} \begin{pmatrix} -1 & \dots \\ \dots & 1 \end{pmatrix} \tag{7}$$

where $w$ is the width of the sample matrix in the horizontal (vertical) direction.

3. **Optimized spatial dimensionality reduction algorithm on MPCA.** In this paper, the large-scale image data is decomposed into many small-scale data, some kind of representation model related to the dimensionality reduction algorithm is built on each small-scale data, and then the models are merged to get the total representation model that can represent all the data, and finally the desired projection direction is found. The improved MPCA process is shown in Figure 1. Firstly, the storage space occupied by the representation model should be as small as possible; secondly, the representation model should be easy to be merged and analyzed; and finally, the projection direction can be easily and efficiently solved by using the representation model. After obtaining the final total representation model, the final optimal projection direction is obtained using the final total representation model to realize the dimensionality reduction and subsequent classification of large-scale image data.

3.1. **Feature space representation modeling.** Assuming that there are $M$ training samples, denoted as $X_j, j = 1, 2, \dots, M$, and each training sample is represented as an image matrix of $n \times m$, the eigenspace model of the observed training samples is denoted by $\{M, \bar{n}, Q, V\}$, where $M$ denotes the number of training samples; $\bar{n}$ denotes the mean vector after averaging by columns of the image mean matrix $\bar{X} \in \mathbb{R}^{n\times m}$, where $\bar{X} = (I/M)\sum_{j=1}^{M} X_j$; $V = \operatorname{diag}(\mu_1, \mu_2, \dots, \mu_c) \in \mathbb{R}^{c\times c}$ denotes the covariance matrix's first $c$ largest eigenvalues; $Q = [q_1, q_2, \dots, q_c] \in \mathbb{R}^{n\times c}$ denotes the covariance matrix's first $c$ largest eigenvalues of the covariance matrix.

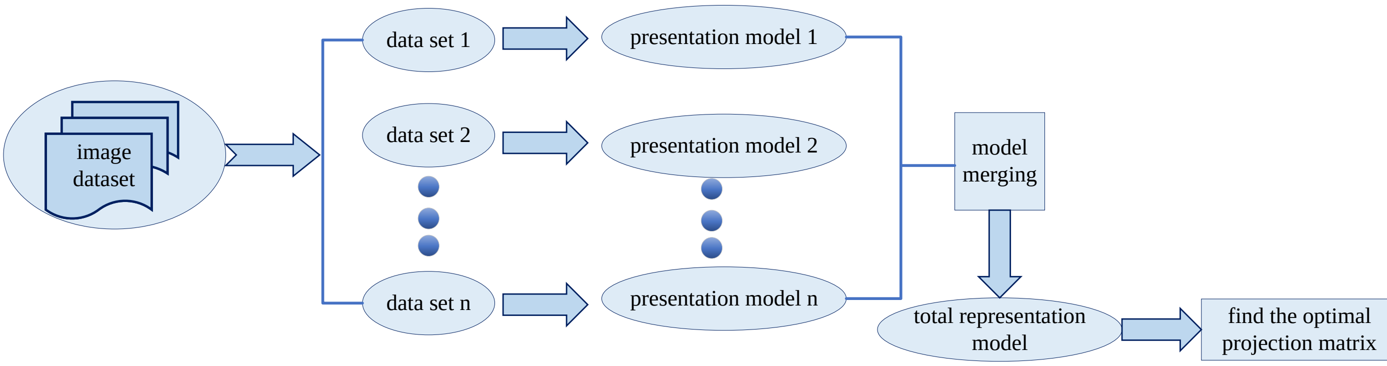


Figure 1. Improved MPCA framework

3.2. **Merging of feature space representation models.** Assuming that there are two subrepresentation models of MPCA, denoted by $x$ and $y$, respectively, their feature space models are denoted as $\{M_x, \bar{n}_x, Q_x, V_x\}$, $\{M_y, \bar{n}_y, Q_y, V_y\}$, yielding the merged sample numbers and mean vectors as follows, respectively.

$$M_z = M_x + M_y, \quad \bar{n}_z = \frac{M_x\bar{n}_x + M_y\bar{n}_y}{M_z} \tag{8}$$

Next, the combined eigenvector matrix $Q_z$ and the eigenvalue diagonal array $V_z$ are computed, assuming that $D_z$ is the total covariance matrix representing the two sets of data, which obviously satisfies $D_z \approx Q_z V_z Q_z^T$, with the significant eigenvalues taken. The equation is as follows.

$$D_z \approx \frac{M_x}{M_z} Q_x V_x Q_x^T + \frac{M_y}{M_z} Q_y V_y Q_y^T + \frac{m^2 M_x M_y}{M_z^2} (\bar{n}_x - \bar{n}_y)(\bar{n}_x - \bar{n}_y)^T \tag{9}$$

where $m$ is the number of columns in each image sample.

To obtain $Q_z$ and $V_z$, direct eigendecomposition of the covariance matrix $D_z$ may be very time-consuming or even impossible due to the dimensionality. So, we have censored the eigenvalues and retained only a small number of meaningful larger eigenvalues and their corresponding eigenvectors. We use this feature to convert the problem into a small-scale problem, i.e., an eigendecomposition problem of a small-scale matrix, which can greatly reduce the computational complexity. Assume the following equation.

$$\varphi_1 = \sqrt{\frac{M_x}{M_z}} Q_x (V_x)^{\frac{k}{2}}, \quad \varphi_2 = \sqrt{\frac{M_y}{M_z}} Q_y (V_y)^{\frac{k}{2}}, \quad \Phi = m\sqrt{\frac{M_x M_y}{M_z}} (\bar{n}_x - \bar{n}_y) \tag{10}$$

Then the covariance matrix $D_z$ can be expressed as follows.

$$D_z = [\varphi_1 \varphi_2 \Phi][\varphi_1 \varphi_2 \Phi]^T = AA^T \tag{11}$$

Let $B = A^T A$, then we have $B \in \mathbb{R}^{s \times s}$, $s_z = c_x + c_y + 1$. Obviously the size of the $B$ matrix is much smaller than the covariance matrix $D_z$, and the eigendecomposition of the matrix $B$ is easy. Let the eigenvector matrix and eigenvalue diagonal matrix of $B$ obtained after selection be $Q_B \in \mathbb{R}^{s \times c_z}$ and $V_B \in \mathbb{R}^{c_z \times c_z}$, respectively, then we have: $A^T A Q_B = Q_B V_B$.

Multiply both sides by $A$ to left, and you get $AA^T A Q_B = A Q_B V_B$, which is $D_z A Q_B = A Q_B V_B$.

Thus the eigenvector matrix and eigenvalue diagonal array of the covariance matrix $D_z$ can be derived as follows, respectively.

$$Q_z = AQ_B (V_B)^{\frac{1}{2}}, \quad V_z = V_B \tag{12}$$

In this way, we obtain a new feature space representation model by merging the two subrepresentation models, in which the column vector of $Q_z$ in the representation model represents the feature direction of the merged feature space, i.e., the optimal projection direction of the two sets of data after merging.

## 4. Image data classification algorithm based on SDSI.

**4.1. Smooth structural information of the original image.** The main problems in the current image data classification algorithms are: (1) although one of the bilateral MPCA reaches the best current results of similar methods, its objective function is non-convex, and it can only be solved by using an alternating iteration optimization algorithm, which is computationally expensive, and at the same time does not guarantee the global optimum of the solution; and (2) the implicit method of divisional reorganization, although it intuitively better maintains the local spatial structure of the image smooth, its space is still undersmoothed due to the lack of explicit smoothing enforced by it.

To overcome the shortcomings in the above classification methods, this paper designs an efficient image data classification algorithm based on SDSI, using the matrixed image Euclidean distance to directly pre-embed spatial structural information on the original image, combined with image restructuring, spatial dimensionality reduction of the restructured data based on the optimized MPCA algorithm mentioned above, and the introduction of the bilateral filter image smoothing method as the image smoothing strategy, which not only greatly reduces the computational complexity, and also ensures global optimality of the solution. The flow of SDSI is shown in Figure 2.

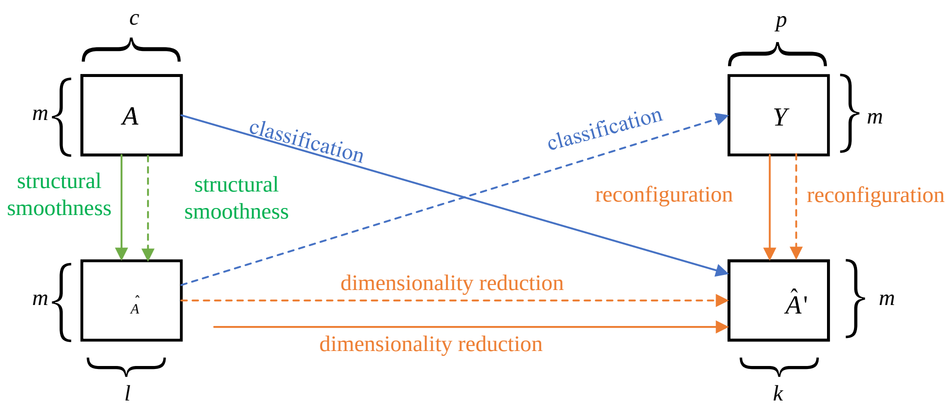


Figure 2. The flow of SDSI algorithm

In this paper, Bilateral Filter (BLF) [26] is introduced to process the gray values of neighboring pixel points of image data, which takes into account the geometric spatial proximity and similarity in gray values, and the blurred edge information can be maintained while the image is smooth.

The original image $A_k$ is smoothed by bilateral filtering to obtain the image $\hat{A}_k$. The gray value $\hat{A}_{ij}$ of the pixel at spatial coordinate $(i, j)$ is transformed as follows:

$$\hat{A}_{ij} = \frac{\sum_{p,q \in T_{i,j}} h(p,q) f(p,q)}{\sum_{p,q \in T_{i,j}} h(p,q)} \tag{13}$$

where $T_{i,j}$ denotes the neighborhood of size $(2N+1)^2$ with $(i,j)$ as the center point, $N$ is the half-width of the filter, and the larger the value of $N$, the stronger the smoothing effect. $h(p,q)$ is the weighting coefficient at $(i,j)$, which consists of the product of two factors: $h(p,q) = h_r(p,q)h_t(p,q)$. Among them, the spatial proximity factor $h_r(p,q)$ and the gray value similarity factor $h_t(p,q)$ are expressed as follows:

$$h_r(p,q) = \exp\left(\frac{|p-i|^2 + |q-j|^2}{2\vartheta_r^2}\right) \tag{14}$$

$$h_t(p,q) = \exp\left(-\frac{|A(p,q) - A(i,j)|}{2\vartheta_t^2}\right) \tag{15}$$

where $r$ and $t$ control the degree of attenuation of $h_r(p,q)$ and $h_t(p,q)$, respectively.

4.2. **Image data segmentation and reorganization and spatial dimensionality reduction.** A given smooth image of size $s \times d$ is partitioned into $LL$ spatial windows of the same size $n \times m$, $LL = sd/nm$, and each spatial window is sequentially pulled into vectors thereby forming a new matrix of dimension $nm \times LL$. The matrix is then reorganized by dividing and reorganizing the spatial windows into columns. By dividing and reorganizing, the columns of the new matrix (the corresponding spatial windows) are usually smooth, and the implicit local spatial relations are highly and fully preserved.

The optimized MPCA algorithm is then adopted to spatially downscale the restructured image matrix while merging multiple feature space models. Assume that there are $s$ feature space representation models: $K_j = \{M_j, \bar{n}_j, Q_j, V_j | j = 1, 2, \ldots, s\}$, while merging these representation models to obtain the total feature space model $f_n = (M_{f_n}, \bar{n}_{f_n}, Q_{f_n}, V_{f_n})$. The total number of samples $M_{f_n}$ with the total mean vector $\bar{n}_{f_n}$ is as follows.

$$M_{f_n} = \sum_{j=J}^{s} M_j, \bar{n}_{f_n} = \frac{J}{M}\sum_{j=J}^{s} M_j\bar{n}_j \tag{16}$$

The total covariance matrix $D_{f_n}$ is shown below.

$$D_{f_n} = \sum_{j=J}^{s} \frac{M_j}{M} Q_j V_j Q_j^T + \sum_{j=J}^{s-k}\sum_{j=J+i}^{j+J} \frac{n^2 M_i M_j}{M^2}(\bar{n}_i - \bar{n}_j)(\bar{n}_i - \bar{n}_j)^T \tag{17}$$

In the same way, we order:

$$\varphi_j = \sqrt{\frac{M_j}{M_{f_n}}} Q_j (V_j)^{k/2}, j = 1, 2, \ldots, s \tag{18}$$

$$\Phi_{ij} = \frac{m\sqrt{M_i M_j}}{M_{f_n}}(\bar{n}_j - \bar{n}_i), j = 1, 2, \ldots, s-1, j = i+1, \ldots, s \tag{19}$$

$$W = [\varphi_1, \varphi_2, \ldots, \varphi_k, \ldots, \Phi_{(s-1)s}] \tag{20}$$

Then we have $D_{f_n} = WW^T$, the eigen decomposition of $WW^T$ can get the main eigenvalue $V_w$ and the main eigenvector $Q_w$, and finally get the main eigenvalue diagonal array of $D_{f_n}$ and the corresponding eigenvector matrix as follows.

$$V_{f_n} = V_w, Q_{f_n} = WQ_w(V_w)^{k/2} \tag{21}$$

The total representation model is obtained after merging all the feature space representation models, and the column vectors of the feature vector matrix in the total representation model represent the projection direction.

4.3. **Image data feature extraction and classification.** For the reorganized and dimensionality reduced image matrix, feature extraction method is used to extract the features. Let $\hat{A}' = [\hat{A}'_1, \hat{A}'_2, ..., \hat{A}'_M]$ be a set of dimensionalities reduced image matrices divided into $L$ classes. Class $i$ contains $m^i$ image samples. Construct an undirected weighted graph $G' = [\hat{A}', H']$ with $\hat{A}'$ as the vertex set, where the element $H'_{i,j}$ of the weight matrix $H'$ is the similarity between the fixed points $\hat{A}'_i$ and $\hat{A}'_j$, and $V$ is the diagonal matrix with diagonal elements $V_{i,i} = \sum_{i,j} H'_{i,j}$. Let $Y = [Y_1, Y_2, ..., Y_M]$ be the extracted feature set, and its objective function can be expressed as follows:

$$\min \sum_{i,j} \|Y_j - Y_l\|^2 H'_{i,j} \tag{22}$$

In this paper, features are extracted based on Linear Discriminant Analysis (LDA) [26]. LDA maximizes the inter-class scatter by minimizing the intra-class scatter while maximizing the inter-class scatter in order to separate the different classes as much as possible. The weight matrix $H'$ is as follows:

$$H'_{i,j} = \begin{cases} \frac{1}{m^s}, & \text{if } \lambda_i = \lambda_j = s \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

The matrix $\hat{A}'$ obtained after the above steps is smooth in the column direction, while maintaining its smoothness, the features are extracted in the row direction using $Y_j = \hat{A}'_j V$, and brought into Equation (22) to obtain the objective function.

$$\min \sum_{i,j} \|Y_i - Y_j\|^2 H'_{i,j} = \min \sum_{i,j} \|\hat{A}'_j V - \hat{A}'_i V\|^2 \tag{24}$$

By a simple transformation of the formula, the above equation is rewritten as follows.

$$\max \operatorname{tr} \left( V^T \sum_{i,j} H'_{i,j} \hat{A}'_i \hat{A}'_j \right) V \tag{23}$$

Compute the eigenvectors $Q_1, Q_2, ..., Q_c$ corresponding to the first $C$ largest eigenvalues. The optimization of the projection matrix $U$ can be found by solving the eigendecomposition problem of Equation (26).

$$\sum_{i,j} H'_{i,j} \hat{A}'_i \hat{A}'_j U = \lambda \sum_i \left( v_i \hat{A}'_i \hat{A}'_i \right) U \tag{24}$$

Given a test image matrix $Q$, spatial smoothing and dimensionality reduction are performed to obtain a new matrix $\hat{Q}'$, after which LDA is used to extract features from it: $F = \hat{Q}'V$. Assuming that $Y_1, Y_2, ..., Y_M$ is the feature matrix of the corresponding training sample, using the nearest neighbor classifier algorithm [27], the image $Q$ is discriminated to the class to which the training sample image $Y_s$ belongs if $dis(Y_s, F) = \min(dis(Y_j, F))$, for all $j = 1, 2, ..., M$.

Define the distance between the identity matrix $Y_1 = [y_1^1, y_1^2, ..., y_1^c]$ and the identity matrix $Y_2 = [y_2^1, y_2^2, ..., y_2^c]$ as follows:

$$dis(Y_1, Y_2) = \sum_{j=1}^{c} \|y_1^j - y_2^j\|^2 \tag{25}$$

## 5. Algorithm performance testing and analysis.

5.1. **Analysis of downscaling results.** To estimate the performance of the SDSI algorithm designed in this article, this article compares it with other existing algorithms for simulation and experimentation, and all the experiments are done on Python platform in personal computer. The computer configuration is as follows: Windows 10 operating system, 8GB RAM and Intel I7 processor. The experiments were performed on the FERET face database and the algorithms compared were FGIA [12], FETD [21], and RESM [28]. We selected a portion of data from the FERET face repository [29] for our experiments. This part of the data contains 1400 images of 200 individuals, 7 images per person. The images of each person were taken at different times, lighting and facial expressions. All images were cropped to 32 × 32 size. Figure 3 shows some of the images of two people. For ease of description, the accuracy is denoted as Acc and the recognition rate is denoted as Rec.

Figure 3. The flow of SDSI algorithm

In the FERET database, we randomly select 4 images for each class as training samples and the rest as test samples, and repeat the process 100 times as well. Table 1 lists the average recognition rate and standard deviation of the SDSI algorithm and the comparison algorithm designed in this paper. We can see that the average recognition rate of the SDSI algorithm is 92.78%, and the FGIA, FETD, and RESM algorithms are 78.16%, 87.41%, and 82.95%, respectively, and the SDSI algorithm is higher than the other three 2D algorithms, and in the actual updating of the existing model, since our algorithm can handle multiple new samples at a time, while the FGIA, FETD, and RESM algorithms can only handle one new sample at a time, which makes our algorithm significantly reduce the number of updates to the model.

Table 1. Average Recognition Rate and Standard Deviation for SDSI and Comparison Algorithms

| Algorithm | SDSI | FGIA | FETD | RESM |
|---|---|---|---|---|
| Recognition Rate (%) | 92.78 | 78.16 | 87.41 | 82.95 |
| Standard Deviation (%) | 0.35 | 0.42 | 0.47 | 0.40 |

Then the dimension sizes derived above are applied to the following experiments to compare the effects of the original data and the data after dimensionality reduction by the SDSI algorithm on the classifier training time and accuracy, respectively, where the training set is 50 samples and 150 samples, and the test set is 200 samples. The results are shown in Table 2 below.

From the perspective of training time, the original data contains more information, so the training time is longer; for the SDSI algorithm, the data is downsized, the redundant information is removed, and the data information is compressed to 98% of the original,

Table 2. Comparison of experimental results with different number of training samples

| Training samples' number | Method | Training time (s) | Acc (%) |
|---|---|---|---|
| 50 | original data | 517 | 78.93 |
| 50 | SDSI+original data | 41 | 88.16 |
| 150 | original data | 3194 | 81.42 |
| 150 | SDSI+original data | 492 | 93.28 |

so the data dimensions become smaller, and therefore the training time of the classifier in training the model is inevitably shortened. This also shows the advantage that the SDSI algorithm can significantly reduce the training time after dimensionality reduction of the data. From the accuracy point of view, the accuracy rate obtained by using the SDSI algorithm exceeds that obtained by the original data classification, and even exceeds about 5% when the training sample is 150.

As can be seen from Figure 4, we can find that both SDSI algorithm and FETD algorithm converge to the MPCA projection direction in the end, but compared with algorithm FETD, our incremental algorithm converges faster and has better convergence effect. This may be because our algorithm is based on the combination of group data. Compared with FGIA algorithm and RESM algorithm, one sample is updated once. Our method reduces the number of updates, thus reducing the cumulative error in the updating process, and making the projection direction correction faster in each model update, thus speeding up the convergence.

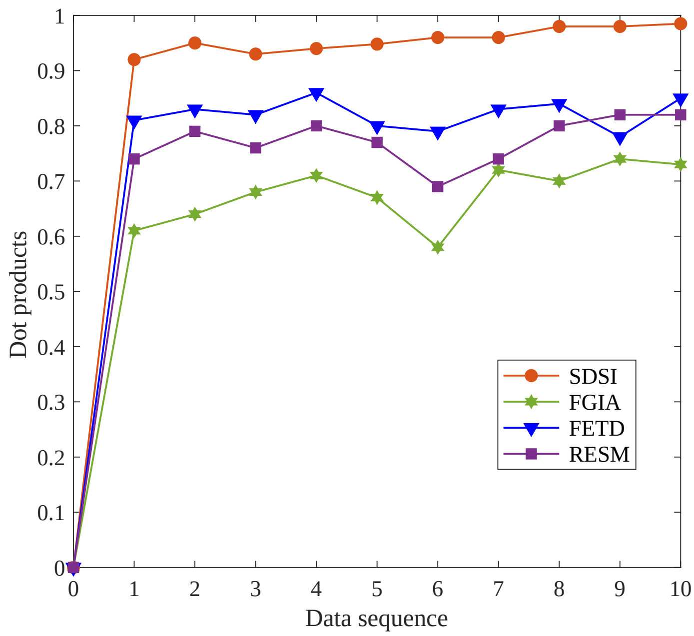


Figure 4. Comparison of the convergence on the FERT library in the projection direction

5.2. **Classification Performance Comparison.** This experiment records the results of training time, classification accuracy, and classification recognition rate of 200 test samples under 400 training sample numbers under four algorithms. The data after dimensionality reduction using the multilinear principal component analysis algorithm has a core vector dimension size of 13*14*12 when using Tucker decomposition. The outcome is indicated in Table 3 below.

Table 3. Comparison of experimental results of four algorithms

| | **SDSI** | **FGIA** | **FETD** | **RESM** |
|---|---|---|---|---|
| Training time (s) | 34 | 153 | 109 | 76 |
| Acc (%) | 92.51 | 75.42 | 85.16 | 78.34 |
| Rec (%) | 95.82 | 79.14 | 87.65 | 83.91 |

As shown in Table 3, under different numbers of training samples, compared with the comparison algorithms, the SDSI algorithm used in this paper has lower training time and higher classification accuracy for image classification, and the FGIA algorithm has the longest training time and the lowest classification accuracy and recognition rate, which is due to the fact that a large amount of time is spent on the selection of feature vectors in the downscaling of dimensionality decomposition to obtain the inner product of vectors. the classification accuracy of the FETD algorithm performs better than the other two algorithms. performs better than the other two algorithms, probably because the data in tensor form is pulled into vector form, which leads to too high data dimensionality, and at the same time, the number of training samples is relatively small, and the model captures too little feature information, which leads to longer training time.The recognition rate of the RESM algorithm performs poorly, which is due to the fact that it doesn't take into account the structural information inside the image, which leads to bias in the recognition. Therefore, SDSI algorithm is more advantageous for image data classification problem.

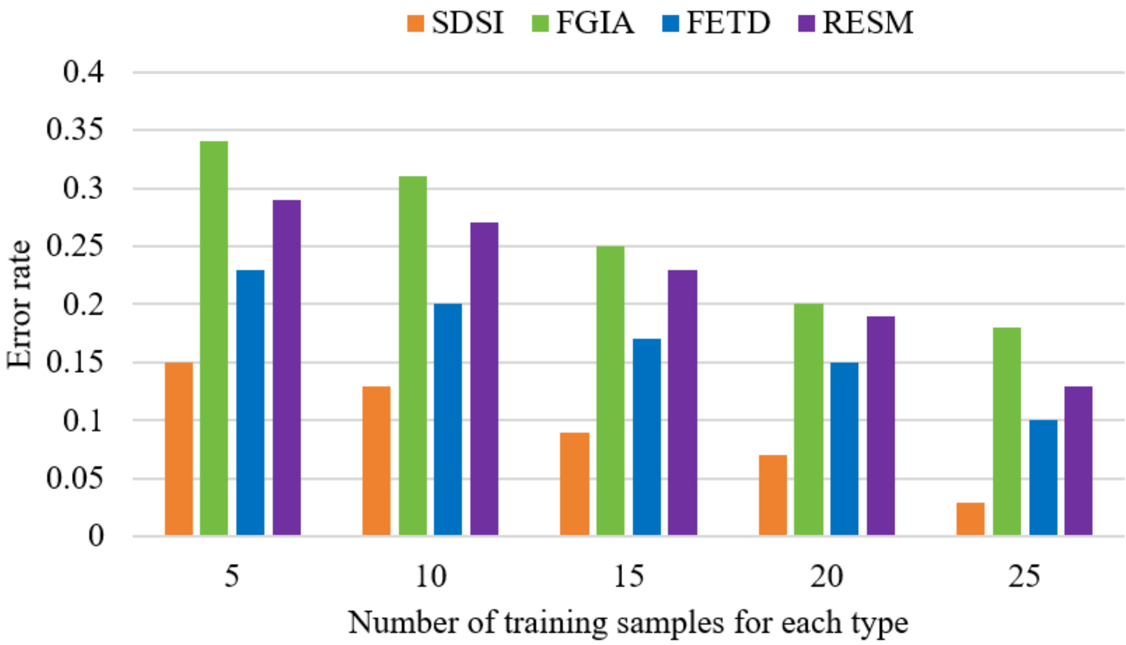


Figure 5. Comparison of error rate of different algorithms

Figure 5 shows the misclassification rate of SDSI algorithm compared with the comparison algorithms.The results after 100 random sampling experiments show that,when the number of samples of each type is 10,the misclassification rate of SDSI is 13%, FGIA 31%, FETD 20% and RESM 27%.S therefore locally structured smooth SDSI algorithm achieves optimal classification performance over FGIA, FETD and RESM algorithms for all the training divisions.

6. **Conclusion.** Aiming at the issue of low accuracy of existing image data classification methods, this paper proposes an image data classification algorithm based on spatial downscaling and structural information. The method first optimizes the MPCA algorithm by obtaining the feature space representation model and projection direction of all data through model merging. Then, based on the above optimized MPCA algorithm, spatial dimensionality reduction is performed on the image data, multiple feature space models are merged at the same time, and combined with the spatial smooth structure information, the original image is divided and reorganized by implicit regularization to reflect the smoothness of the column vectors, and the spatial structure information of the image is implicitly utilized to ensure the global optimality of the understanding. Finally, the experimental results show that the method proposed in this paper effectively improves the accuracy, precision, and recall of classification, and can be better applied to the field of image data classification.

**Acknowledgment.** This work is supported by the 2023 Guangdong Provincial Education Research Higher Education Project (No. 2023GXJK407), the 2022 Guangdong Province Education Department Project (No. 2022SJJXGG992), the 2021 Guangzhou Philosophy and Social Sciences Development Project (No. 2021GZGJ145), the 2021 Higher Vocational Education Project (No. 2021GZJG05), the 2022 School-level Quality Engineering Construction Project(No. 2022JXGG29), the 2022 School-level Quality Engineering Construction Project(No. 2022JXGG68), the 2021 University-level Higher Education Teaching Reform Project (No. 2021JXGG15), and the 2023 University-level Research Project (No. 2023XJYB33).

## REFERENCES

[1] D. Lu, and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*, vol. 28, no. 5, pp. 823-870, 2007.

[2] C. Affonso, A. L. D. Rossi, F. H. A. Vieira, and A. C. P. de Leon Ferreira, "Deep learning for biological image classification," *Expert Systems with Applications*, vol. 85, pp. 114-122, 2017.

[3] J. Song, S. Gao, Y. Zhu, and C. Ma, "A survey of remote sensing image classification based on CNNs," *Big Earth Data*, vol. 3, no. 3, pp. 232-254, 2019.

[4] L. Wang, Y. Zhang, and J. Feng, "On the Euclidean distance of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1334-1339, 2005.

[5] W. Di, and M. M. Crawford, "Active learning via multi-view and local proximity co-regularization for hyperspectral image classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 618-628, 2011.

[6] Z. Wang, and S. Chen, "New least squares support vector machines based on matrix patterns," *Neural Processing Letters*, vol. 26, pp. 41-56, 2007.

[7] Z. Wang, and S. Chen, "Matrix-pattern-oriented least squares support vector classifier with AdaBoost," *Pattern Recognition Letters*, vol. 29, no. 6, pp. 745-753, 2008.

[8] Z. Wang, C. Zhu, D. Gao, and S. Chen, "Three-fold structured classifier design based on matrix pattern," *Pattern Recognition*, vol. 46, no. 6, pp. 1532-1555, 2013.

[9] Z. Wang, S. Chen, J. Liu, and D. Zhang, "Pattern representation in feature extraction and classifier design: matrix versus vector," *IEEE Transactions on Neural Networks*, vol. 19, no. 5, pp. 758-769, 2008.

[10] Y. Ding, W. K. Wong, Z. Lai, and Z. Zhang, "Bilinear supervised hashing based on 2D image features," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 2, pp. 590-602, 2019.

[11] C. Hou, F. Nie, C. Zhang, D. Yi, and Y. Wu, "Multiple rank multi-linear SVM for matrix data classification," *Pattern Recognition*, vol. 47, no. 1, pp. 454-469, 2014.

[12] S. Hossain, S. Umer, R. K. Rout, and M. Tanveer, "Fine-grained image analysis for facial expression recognition using deep convolutional neural networks with bilinear pooling," *Applied Soft Computing*, vol. 134, pp. 109997, 2023.

[13] Q. Zheng, F. Zhu, J. Qin, B. Chen, and P.-A. Heng, "Sparse support matrix machine," *Pattern Recognition*, vol. 76, pp. 715-726, 2018.

[14] S. Kramer, G. Widmer, B. Pfahringer, and M. De Groeve, "Prediction of ordinal classes using regression trees," *Fundamenta Informaticae*, vol. 47, no. 1-2, pp. 1-13, 2001.

[15] Y. Lei, H. Zhu, J. Zhang, and H. Shan, "Meta ordinal regression forest for medical image classification with ordinal labels," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 7, pp. 1233-1247, 2022.

[16] B.-Y. Sun, J. Li, D. D. Wu, X.-M. Zhang, and W.-B. Li, "Kernel discriminant learning for ordinal regression," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 6, pp. 906-910, 2009.

[17] Q. Tian, S. Chen, and X. Tan, "Comparative study among three strategies of incorporating spatial structures to ordinal image regression," *Neurocomputing*, vol. 136, pp. 152-161, 2014.

[18] J. Yang, D. Zhang, A. F. Frangi, and J.-y. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131-137, 2004.

[19] X. Huang, X. Guan, J. A. Benediktsson, L. Zhang, J. Li, A. Plaza, and M. Dalla Mura, "Multiple morphological profiles from multicomponent-base images for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 12, pp. 4653-4669, 2014.

[20] L. Han, Z. Wu, K. Zeng, and X. Yang, "Online multilinear principal component analysis," *Neurocomputing*, vol. 275, pp. 888-896, 2018.

[21] W. Zhao, and S. Du, "Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4544-4554, 2016.

[22] Q. Gao, X. Hao, Q. Zhao, W. Shen, and J. Ma, "Feature extraction using two-dimensional neighborhood margin and variation embedding," *Computer Vision and Image Understanding*, vol. 117, no. 5, pp. 525-531, 2013.

[23] Y. Zhu, S. Chen, and Q. Tian, "Spatial regularization in subspace learning for face recognition: implicit vs. explicit," *Neurocomputing*, vol. 173, pp. 1554-1564, 2016.

[24] T.-Y. Wu, A. Shao, and J.-S. Pan, "CTOA: Toward a Chaotic-Based Tumbleweed Optimization Algorithm," *Mathematics*, vol. 11, no. 10, 2339, 2023.

[25] T.-Y. Wu, H. Li, and S.-C. Chu, "CPPE: An Improved Phasmatodea Population Evolution Algorithm with Chaotic Maps," *Mathematics*, vol. 11, no. 9, 1977, 2023.

[26] E. K. Wang, C.-M. Chen, M. M. Hassan, and A. Almogren, "A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain," *Future Generation Computer Systems*, vol. 108, pp. 135-144, 2020.

[27] C. H. Park, and H. Park, "A comparison of generalized linear discriminant analysis algorithms," *Pattern Recognition*, vol. 41, no. 3, pp. 1083-1097, 2008.

[28] A. Kataria, and M. Singh, "A review of data classification using k-nearest neighbour algorithm," *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, no. 6, pp. 354-360, 2013.

[29] H. Tan, B. Yang, and Z. Ma, "Face recognition based on the fusion of global and local HOG features of face images," *IET Computer Vision*, vol. 8, no. 3, pp. 224-234, 2014.

[30] H. Touvron, P. Bojanowski, M. Caron, M. Cord, A. El-Nouby, E. Grave, G. Izacard, A. Joulin, G. Synnaeve, and J. Verbeek, "Resmlp: Feedforward networks for image classification with data-efficient training," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 5314-5321, 2022.