IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 30, 2019, accepted June 16, 2019, date of publication June 24, 2019, date of current version July 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2924676

Cluster-Based Tensorial Semisupervised Discriminant Analysis for Feature Extraction of SAR Images

XIAOYING WU[1,2,3,4], XIANBIN WEN[1,2,3], LIMING YUAN[1,2,3], CHANGLUN GUO[1,2,3], AND HAIXIA XU[1,2,3]

[1]School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China
[2]Key Laboratory of Computer Vision and System, Ministry of Education, Tianjin 300384, China
[3]Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin 300384, China
[4]School of Information Engineering, Henan Institute of Science and Technology, Xinxiang 453003, China

Corresponding author: Xianbin Wen (xbwen317@163.com)

The work was supported in part by the National Natural Science Foundation of China under Grant 61472278, in part by the Major Project of Tianjin under Grant 18ZXZNGX00150, in part by the Key Project of Natural Science Foundation of Tianjin University under Grant 2017ZD13, in part by the Research Project of Tianjin Municipal Education Commission under Grant 2017KJ255, and in part by the Natural Science Foundation of Tianjin under Grant 18JCYBJC84800.

ABSTRACT Several features have been developed to characterize the land cover in synthetic aperture radar (SAR) data with speckle noise. Feature extraction has become an essential task for SAR image processing. However, how to preserve the original intrinsic structural information and enhance the discriminant ability to reduce the impact of noise is still a challenge in this area. In this paper, using a clustering method to maintain the nonlocal information in images and tensors with the ability to preserve spatial neighborhood structure information, a new cluster-based tensorial semisupervised discriminant analysis (CTSDA) method is proposed for feature extraction of SAR images. In the CTSDA, the block clustering algorithm is employed to generate several high-order clustering tensors of multifeature SAR images, which preserves the intrinsic nonlocal spatial information and neighborhood structure. In the multiple manifold structures of the cluster tensors, the improved discriminant analysis enhances the feature discrimination by considering the local structure and labeled and unlabeled information through the Laplace matrix, and the fusion of tensor algebraic analysis and improved discriminant analysis produces multiple new projection directions of the cluster tensors. Finally, feature extraction is achieved by rearranging the projected cluster tensors. The experimental results on the simulated SAR data and four real SAR images demonstrate the superiority of the proposed method over several state-of-the-art approaches.

INDEX TERMS Feature extraction, nonlocal, tensor, semisupervised discriminant analysis, SAR.

I. INTRODUCTION

Synthetic aperture radar (SAR) is an active microwave remote sensor that can penetrate the cloud layer and is less affected by atmospheric attenuation. SAR images contain speckle noise and strong scattering points, which makes SAR image classification very difficult. Over the past few years, several features have been developed for SAR classification, such as texture features [1], [2], geometrical features [3], [4], visual features [41], edge features [5], [6], spatial relations [7], [8], and combinations of multiple features [9]– [11]. Consequently, feature extraction has become an essential task in SAR image processing. However, how to preserve the original intrinsic structural information and enhance the discriminant ability to reduce the impact of noise is still a challenge in this area. Therefore, several feature extraction methods [12] have been proposed to address this challenge, which extract the features of effective information from high-dimensional space into low-dimensional space through a projection. These methods can be classified as nonlinear or linear. Several nonlinear methods, such as Laplacian eigenmaps (LE) [13] and t-stochastic neighbor embedding (TSNE) [14], have been developed. LE use a local perspective to build the relationship between data, and TSNE tries to keep the distribution of

The associate editor coordinating the review of this manuscript and approving it for publication was Guitao Cao.

data unchanged and maintains a strict segmentation interface. Neither of these methods has an explicit expression, so new samples cannot be expanded, and they lack discrimination. The linear manifold learning methods include principal component analysis (PCA) [15] and linear discriminant analysis (LDA) [15], [16]. PCA aims to maximize the mutual information between the original high-dimensional data sets, regardless of the label information. LDA aims to simultaneously find a projection matrix to maximize the trace of the between-class scatter matrix and minimize the trace of the within-class scatter matrix in the projected subspace, but local information cannot be embodied. In addition, there are other feature extraction methods, such as sparse and low-rank representation [17]–[20]. But traditional feature extraction methods are vector-based methods, and they ignore the global correlation and spatial relationship among neighboring pixels. Because the matrices of SAR images must first be unfolded into a one-dimensional vector structure, the reshaping process breaks the natural structure of the image.

Therefore, tensor-based methods have been proposed because they can preserve the overall spatial structure, and they have been applied in several areas, including biological [21] and medical research [22], facial recognition [23], [24], natural image processing [25], hyperspectral image analysis [26]–[28], [42] and PolSAR image processing [29], [30]. Lower rank tensor approximation (LTRA) [26] is an extension of PCA to higher-order data sets that assumes strong global correlations in the spatial and feature domains. Tensor locality preserving projection (TLPP) [28] can effectively embed the spatial structures and feature information into low-dimensional space simultaneously using a projection matrix. In these studies, a high-order tensor is used to express the research object to calculate the specific amount in the data processing. The greatest advantage of doing so is that the structural characteristics of the original data are fully maintained, and the global correlation and spatial information among neighboring pixels are fully used. However, these tensor methods simply apply the gray value information, considering that the spatial information is not comprehensive, and lack discrimination. Furthermore, all of the methods are essentially based on the hypothesis that the data are drawn from a low-dimensional subspace.

Tensors embody all of the spatial information, but in practice, a data set is often not described well by a single subspace. Therefore, it is more reasonable to consider data lying on a mixture of multiple low-dimensional subspaces, with each subspace fitting a subgroup of data. Corresponding to the following conditions, in the spatial domain, the arrangement of land covers is disordered, and a global correlation is not always guaranteed. In other words, there is a nonlocal spatial correlation, so simply projecting all of the data in the high-dimensional space to one low-rank space is not appropriate, and the classification results of the feature extraction are not desirable. In reality, there is little labeled information, and a large amount of data is unlabeled, so the information about the two should be fully utilized.

Inspired by the issues presented above, the study of clustering and semisupervised feature extraction based on a multifeature tensor framework facilitates SAR land cover classification. Using tensors as a whole frame, the semisupervised method [31], [32] is considered to learn the characteristics, which can use a small number of labeled samples and many unlabeled samples to estimate the intrinsic manifold structure of the data. To better reflect the distribution of samples, clustering is used to maintain the nonlocal information. In addition, the multimanifold local geometry structure is reflected through improved discrimination analysis.

Finally, a feature extraction method, cluster-based tensorial semisupervised discriminant analysis (CTSDA), is proposed for SAR classification, which considers the discrimination of multiple manifold structures. The contributions of the method are summarized as follows. 1) Unlike traditional vector-based methods, which treat each sample as an independent item, the samples in the CTSDA are represented in a tensor form, which can preserve the original spatial neighborhood information. 2) Traditional feature extraction methods consider that all samples are on the same manifold distribution, while the samples of the CTSDA are distributed in several unique small manifold distributions; each manifold structure generated by CTSDA reflects the local and nonlocal structure information about the image. 3) Traditional tensor feature extraction methods usually only use pixel gray values and generate a single tensor projection factor matrix in the third dimension; however, CTSDA uses multiple feature fusion expressions of the SAR image in the tensor projection and generates multiple new tensor projection factor matrices. This approach generates multiple projection directions in the third dimension, which better preserves the separation between the classes and greatly improves feature separability. 4) The improved discriminant analysis in CTSDA is semisupervised, preserves the local geometric structure to describe the within-class compactness and between-class separability on multiple manifolds, and combined with labeled and unlabeled information, results in significant feature discrimination. In addition, based on well-established search engines used for scientific research, such as Web of Knowledge and Google Scholar and to the best of our knowledge, this study is one of the few that have focused on the classification of nonpolarized SAR data using the tensor feature extraction technique.

The remainder of this paper is organized as follows. Section II describes the related work, and Section III presents the theory of the proposed method in detail. Section IV evaluates the experimental results with simulated and real data with performance analyses. Section V presents the conclusions of this paper.

II. RELATED WORK

The first part of this section presents a fundamental introduction to the tensor algebra, which will provide theoretical foundations for the following discussions. The second part of this section introduces the conventional feature extraction method LDA of manifold learning.

A. BASIC TENSOR ALGEBRA

Tensor algebra [33] is a high-dimensional extension of matrix algebra. Usually, a matrix of three or more dimensions is defined as a tensor. Let $\mathbb{X}$ be a tensor of size $I_1 \times I_2 \times \cdots \times I_m$; the order of $\mathbb{X}$ is m, and the nth dimension (or mode) of $\mathbb{X}$ is the size of I_n.

Similar to the Frobenius norm of a matrix, the norm of a tensor $\mathbb{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_m}$ is the square root of the sum of the squares of all of its elements, or the inner product with itself, and it is defined as

$$\|\mathbb{X}\|_F = \sqrt{\langle \mathbb{X}, \mathbb{X} \rangle} = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_m=1}^{I_m} \mathbb{X}_{i_1 i_2 \cdots i_m}^2}. \quad (1)$$

Matricization, or unfolding, is defined as the process of unfolding a tensor into a matrix. The mode-n unfolding of a tensor $\mathbb{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_m}$ is denoted by $X_{(n)} \in \mathbb{R}^{I_n \times (I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_m)}$. The obtained matrix $X_{(n)}$ can also be regarded as the n-mode-flattened tensor $\mathbb{X}$.

The n-mode product of a tensor $\mathbb{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots I_n \times \cdots \times I_m}$ and a matrix $U \in \mathbb{R}^{J \times I_n}$ is a tensor denoted as $\mathbb{Y}$ of size $I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_m$, which is expressed as

$$\mathbb{Y} = \mathbb{X} \times_n U \Leftrightarrow Y_{(n)} = U X_{(n)}. \quad (2)$$

where $\times_n$ denotes the mode-n tensor-matrix product operator, and $Y_{(n)}$ denotes the mode-n matricization of the resulting m-order tensor $\mathbb{Y}$.

Tucker decomposition is a widely used method of tensor decomposition, which expresses higher-order singular decomposition. This tensor decomposition is denoted as follows:

$$\mathbb{X} = C \times_1 U_1 \times_2 U_2 \times \cdots \times_n U_n = \prod_{i=1}^{n} C \times_i U_i \quad (3)$$

where C is the core tensor, and $U_i, i = 1, \ldots, n$ is the factor matrix, which is the matrix of eigenvectors associated with the covariance matrix $X_{(i)} X_{(i)}^T$, where $X_{(i)}$ is the mode-i unfolding matrix of tensor $\mathbb{X}$.

B. LINEAR DISCRIMINANT ANALYSIS

LDA [34] searches for a projection matrix that makes the data points of different classes far from each other while requiring data points of the same class to be close to each other after the projection. Suppose we have a set of n samples $\{x_i, i = 1, \cdots, n\}$ belonging to c classes. The objective function of LDA is as follows:

$$P_{opt} = \arg\max_P \frac{P^T S_b P}{P^T S_w P}, \quad (4)$$

$$S_b = \sum_{k=1}^{c} n_k \left(\mu^{(k)} - \mu\right)\left(\mu^{(k)} - \mu\right)^T, \quad (5)$$

$$S_w = \sum_{k=1}^{c} \left(\sum_{i=1}^{n_k} \left(x_i^{(k)} - \mu^{(k)}\right)\left(x_i^{(k)} - \mu^{(k)}\right)^T \right), \quad (6)$$

$$S_t = S_b + S_w = \sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T. \quad (7)$$

where P is the projection matrix, n_k is the number of samples in the kth class, μ is the total sample mean vector, $\mu^{(k)}$ is the average vector of the kth class, $x_i^{(k)}$ is the ith sample in the kth class, S_w is the within-class scatter matrix, S_b is the between-class scatter matrix, and S_t is the total scatter matrix.

III. THE PROPOSED METHOD

This section presents the proposed tensor-based feature extraction method in two parts.

A. TENSOR MODELING BASED ON CLUSTERS

The starting point for our model is the multifeature selection of SAR image X, which has a size of $I_1 \times I_2$. Because SAR images are affected by speckle noise, the pixels' gray values cannot be used directly. According to [35], a gray-level co-occurrence matrix (GLCM) that can express the rich texture of SAR images is adopted, which makes full use of the imaging mechanism and the statistical characteristics of SAR images. Mean filtering [36] is used to smooth the SAR image. To protect the sharp edge of the SAR image, median filtering is used to capture the properties in the SAR image [36]. For each pixel x_{ij} in an SAR image X, many texture feature statistics can be calculated using the GLCM. Four feature statistics, namely, contrast, correlation, homogeneity and energy, are selected to express the image pixel, and they are calculated in the four directions required by GLCM: 0°, 45°, 90° and 135°. These directions are used because each measure tends to be independent of other co-occurrence statistics. In addition, each statistic is insensitive to gray level shifts [35]. The mean and median are also calculated in four directions: 0°, 45°, 90° and 135°. Because the range of gray values is different from that of the probability values, the mean value and median value must be normalized, while the original probability value of the GLCM is retained. Based on the general form of multikernel learning [37], the mean and median values are multiplied by weight coefficients for the transformation, which makes their values fall within the range of [0], [1]. Then, three types of features [35], [36] are concatenated to form a 24-dimension feature vector, at 0 degrees, the four GLCM values of contrast, correlation, homogeneity and energy are represented successively, as well as the mean value and the median value. Then the above six values are expressed at 45, 90 and 135 degrees. That is

$$\begin{aligned} F_{ij} = (&\overbrace{f_{ij1}^{0^\circ}, f_{ij2}^{0^\circ}, f_{ij3}^{0^\circ}, f_{ij4}^{0^\circ}}^{0^\circ GLCM}, \overbrace{f_{ij5}^{0^\circ}, f_{ij6}^{0^\circ}}^{0^\circ Mean,Median}, \cdots, \\ &\cdots, \overbrace{f_{ij19}^{135^\circ}, f_{ij20}^{135^\circ}, f_{ij21}^{135^\circ}, f_{ij22}^{135^\circ}}^{135^\circ GLCM}, \overbrace{f_{ij23}^{135^\circ}, f_{ij24}^{135^\circ}}^{135^\circ Mean,Median}), \\ &i = 1, 2, \cdots, I_1, \quad j = 1, 2, \cdots, I_2, \end{aligned} \quad (8)$$

where F_{ij} represents the feature vector of each pixel, and $f_{ijn}^{d^\circ}, n = 1, 2, \cdots, 24, d = 0, 45, 90, 135$ represents one feature value; since there are six different feature values in each direction, there are a total of 24 feature values.

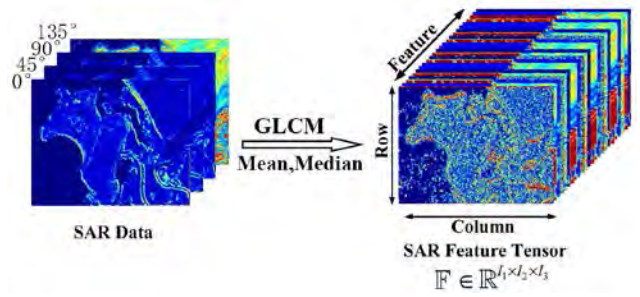


FIGURE 1. Tensor formulation of SAR data with various features.

The expression of features fully takes into account the spatial relationship of image pixel through GLCM value, while data validity and edge information are respectively reflected by the mean and median values obtained through filtering. The three types of features are independent and complementary to each other. Then, a third-order tensor $\mathbb{F} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is constructed by the feature vectors of F_{ij}, and expressed as follows:

$$\mathbb{F} = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1I_2} \\ F_{21} & F_{22} & \cdots & F_{2I_2} \\ \cdots & \cdots & \cdots & \cdots \\ F_{I_1 1} & F_{I_1 2} & \cdots & F_{I_1 I_2} \end{bmatrix}, \tag{9}$$

where I_1 and I_2 denote the spatial dimensionality of the image size, I_1 denotes the row height, I_2 denotes the column width, and I_3 denotes the feature dimensionality, which is 24 in this case. The expressed data in tensor form retains the original data spatial neighborhood information. Fig. 1 shows a tensor structure diagram.

For the SAR data set, if a tensor representation is not used, the global spatial correlation is usually not guaranteed in a particular application, but the local spatial correlation is available within a local neighborhood, except in the boundary areas between different land covers. SAR data also have another property in the spatial domain called nonlocal correlation, which is based on the fact that different areas may have the same land covers. In this paper, to utilize the nonlocal spatial correlation of SAR images, the local spatial correlation is also considered. The proposed method uses clusters to build the tensor model, segments the original SAR image into blocks, classifies the blocks into groups using the unsupervised clustering algorithm of K-means clustering, and reconstructs a new tensor representation.

Let $\mathbb{F} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ be an SAR data set with the features presented above, which is regarded as a third-order tensor corresponding to the SAR image. To utilize the properties of local spatial correlation, the tensor $\mathbb{F}$ is segmented into many blocks in the spatial domain with the fixed window $W_1 \times W_2$. Each block is a third-order subtensor $\mathbb{SF}_{kl} \in \mathbb{R}^{W_1 \times W_2 \times I_3}$, where $1 \leq k \leq m$, $1 \leq l \leq n$, $m = \lceil I_1/W_1 \rceil$, $n = \lceil I_2/W_2 \rceil$, and $\lceil . \rceil$ is a rounding function. Each pair of subtensors is mutually exclusive, and all of the subtensors cover the global tensor. Therefore, each subtensor can explore not only the local spatial similarity but also the global correlation across all of the feature items of the SAR data. Thus, the global image tensor is expressed in the matrix form of blocks, each of which is a subtensor composed of $W_1 \times W_2$ feature eigenvectors, we have

$$\mathbb{F} \simeq \begin{bmatrix} \mathbb{SF}_{11} & \mathbb{SF}_{12} & \cdots & \mathbb{SF}_{1n} \\ \mathbb{SF}_{21} & \mathbb{SF}_{22} & \cdots & \mathbb{SF}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbb{SF}_{m1} & \mathbb{SF}_{m2} & \cdots & \mathbb{SF}_{mn} \end{bmatrix}. \tag{10}$$

In the later clustering process, block form can keep local neighborhood information. To exploit the nonlocal spatial correlation, the patches with similar spatial structures must be identified. Therefore, a fourth-order tensor $\hat{\mathbb{F}} \in \mathbb{R}^{W_1 \times W_2 \times I_3 \times f}$ is constructed based on the subtensors $\mathbb{SF}_{kl}$ in $\mathbb{F}$, where $f = m \times n$ is the number of all subtensors, W_1 and W_2 are the spatial sizes of the subtensors, which are the dimensions of the window described above, and I_3 is the feature dimensionality number. Then, $\hat{\mathbb{F}}$ is converted into a matrix denoted as $\hat{F} \in \mathbb{R}^{(W_1 \times W_2 \times I_3) \times f}$, and the unsupervised clustering algorithm is employed to cluster $\hat{F}$ into q groups. The kth group is denoted as the matrix $GF_k \in \mathbb{R}^{(W_1 \times W_2 \times I_3) \times T_k}$, $k = 1, 2, \cdots, q$, $\hat{F} = \{GF_1, GF_2, \cdots, GF_q\}$, where T_k is the number of subtensors in the kth group, and $\sum_{k=1}^{q} T_k = f$, and q is the number of groups. Finally, GF_k is converted to a fourth-order tensor denoted by $\mathbb{GF}_k \in \mathbb{R}^{W_1 \times W_2 \times I_3 \times T_k}$, $k = 1, 2, \cdots, q$, then

$$\hat{\mathbb{F}} = \{\mathbb{GF}_1, \mathbb{GF}_2, \cdots, \mathbb{GF}_q\}. \tag{11}$$

Thus, all of the subtensors are clustered into different groups. The whole image is expressed by a four-order tensor $\hat{\mathbb{F}} \in \mathbb{R}^{W_1 \times W_2 \times I_3 \times f}$, which is composed of a number of four-order group tensors $\mathbb{GF}_k \in \mathbb{R}^{W_1 \times W_2 \times I_3 \times T_k}$. The fourth dimension is the number of contained three-order subtensors $\mathbb{SF}_{kl} \in \mathbb{R}^{W_1 \times W_2 \times I_3}$. This formal expression embodies the non-local relations of space.

B. KEY PART OF THE CTSDA ALGORITHM

This part presents the second half of the CTSDA method based on the abovementioned feature expression, fusing improved discriminant analysis and multilinear tensor techniques together. The second half of the CTSDA method is referred to as TSDA in the following discussion. For the clusters of subtensors presented above, TSDA is employed to project the clusters into a low-rank space and find a low-dimensionality feature representation. TSDA attempts to estimate a low-rank desired component $\mathbb{Y}_{TSDAk}$ from the noisy $\mathbb{GF}_k$ to mitigate or limit the effects of speckle noise and to reduce much of the redundancy in the feature tensor.

In this section, based on the expression formulation of LDA in tensor modality [33]. $\mathbb{F} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is an SAR data set with the features described above, the matrix $F_{(3)}$ of LDA is equivalent to the three-mode-flattened matrix of $\mathbb{F}$. According to formula (2) of tensor algebra [33], the equivalent of the

image reshaping can be written as follows:

$$Y_{SDA(3)} = BF_{(3)} \Leftrightarrow \mathbb{Y}_{SDA} = \mathbb{F} \times_3 B, \tag{12}$$

where $\mathbb{Y}_{SDA} \in \mathbb{R}^{I_1 \times I_2 \times C}$ is the third-order tensor holding the c projection features obtained by LDA, $\times_n$ is the n-mode product operator, which generalizes the product between a tensor and a matrix along the n mode, and B is the projection matrix. Similarly, according to formula (2), for each cluster tensor $\mathbb{GF}_k \in \mathbb{R}^{W_1 \times W_2 \times I_3 \times T_k}$, $k = 1, 2, \cdots, q$, the matrix $GF_{k(3)}$ is equivalent to the three-mode-flattened matrix of $\mathbb{GF}_k$, which gives

$$Y_{k(3)} = B_k GF_{k(3)} \Leftrightarrow \mathbb{Y}_k = \mathbb{GF}_k \times_3 B_k, \quad k = 1, 2, \cdots, q, \tag{13}$$

The matrix operation is transformed into tensor operation expression, which facilitates the application of matrix in tensor processing. Where $\mathbb{Y}_k \in \mathbb{R}^{W_1 \times W_2 \times p \times T_k}$ is the four-order tensor holding the p projection features obtained by improved LDA in the kth cluster tensor, and $B_k \in \mathbb{R}^{p \times I_3}$ represents the projection matrix, which derivation is illustrated below.

In the kth cluster tensor, x_{ki} represents a sample in the kth cluster tensor. To construct the within-class graph, according to the form of heat kernel with Euclidean norm in literature [40], the weight matrix is defined as

$$W_{ij}^{kw} = \begin{cases} \exp(-\left\|x_{ki} - x_{kj}\right\|^2 / t) & \text{if lable} l_{ki} = l_{kj} \\ 0 & \text{otherwise} \end{cases}, \tag{14}$$

The weight function is a strictly monotonically decreasing function with respect to the distance between samples x_{ki} and x_{kj}.

According to the purpose of the objective function (4), samples from the same class are close, and inspired by the linear representation of LE [13], [40], local structures are preserved, the within-class graph-preserving criterion is defined as

$$\arg\min B_k^T X_k L_{kw} X_k^T B_k, \tag{15}$$

where $L_{kw} = D^{kw} - W^{kw}$ is a Laplacian matrix, and D^{kw} is a diagonal degree matrix with $D_{ii}^{kw} = \sum_j W_{ij}^{kw}$. X_k represents the sample matrix in the kth cluster tensor.

For the between-class graph, M_k represents the classes' centers in the kth cluster tensor, $M_k = [\overline{m}_{k1}, \overline{m}_{k2}, \cdots, \overline{m}_{kc}]$, where $\overline{m}_{ki}$ is the mean value of the samples belonging to the ith class, and c is the number of classes in the kth cluster tensor. According to the form of heat kernel with Euclidean norm and (5), the weight matrix between the class centers reflects multiple manifolds and is defined as

$$W_{ij}^{kb} = \exp\left(-\left\|\overline{m}_{ki} - \overline{m}_{kj}\right\|^2 / t\right). \tag{16}$$

The purpose of the weight in (16) is to directly enhance contributions of classes that have smaller distances.

In order for samples from different classes to be as far away from each other as possible, and local structures among classes' centers are expressed, the between-class graph-penalizing criterion is defined as

$$\arg\max B_k^T M_k L_{kb} M_k^T B_k, \tag{17}$$

where $L_{kb} = D^{kb} - W^{kb}$ is a Laplacian matrix, and D^{kb} is a diagonal degree matrix with $D_{ii}^{kb} = \sum_j W_{ij}^{kb}$. Each cluster tensor has a unique small manifold structure, and its own unique projection matrix B_k is derived, and according to formula (4), it is expressed as follows:

$$B_k = \max_{B_k} \frac{B_k^T M_k L_{kb} M_k^T B_k}{B_k^T X_k L_{kw} X_k^T B_k}. \tag{18}$$

Because labeled and unlabeled sample information is used, the improved discriminant analysis describes the within-class compactness and between-class separability by locally preserving the multimanifolds of the cluster tensors, the features of the cluster tensors are more discriminating, and the projection direction corresponding to the projection matrix is more accurate.

Finding $\mathbb{Y}_{TSDAk}$ involves minimizing the Frobenius norm $\|\mathbb{Y}_k - \mathbb{Y}_{TSDAk}\|_F^2$, with $\mathbb{Y}_k$ defined in (13). $\mathbb{Y}_{TSDAk} \in \mathbb{R}^{W_1 \times W_2 \times p \times T_k}$ is the lower rank approximated tensor of $\mathbb{Y}_k$. Lathauwer *et al.* [38] proved that by using Tucker decomposition formula (3), minimizing the norm of cost function shown above with respect to $\mathbb{Y}_{TSDAk}$ is equivalent to maximizing the following formula with respect to $U_{kn}, n = 1, 2, 3, 4$, which is theorem 4.2 in [38]:

$$\max_{U_{k1}, U_{k2}, U_{k3}, U_{k4}} \left\|\mathbb{Y}_k \times_1 U_{k1}^T \times_2 U_{k2}^T \times_3 U_{k3}^T \times_4 U_{k4}^T\right\|_F^2. \tag{19}$$

Using the Tucker representation and (3), the reduced tensor can be decomposed by rank $(r_{k1}, r_{k2}, r_{k3}, r_{k4})$ as follows:

$$\mathbb{Y}_k = \mathbb{C}_k \times_1 U_{k1} \times_2 U_{k2} \times_3 U_{k3} \times_4 U_{k4}. \tag{20}$$

where $U_{kn}, n = 1, 2, 3, 4$ are the factor matrices, which can be thought of as the principal components in each mode, U_{kn} is the matrix of eigenvectors associated with the n-mode covariance matrix $Y_{k(n)} Y_{k(n)}^T$, where $Y_{k(n)}$ is the mode-n unfolding matrix of $\mathbb{Y}_k$, and $\mathbb{C}_k$ is a core tensor, whose entries show the level of interaction between the different components, and $r_{k4} = T_k$. Likewise, according to tucker decomposition formula (3) and formula 4.2 in [38], $\mathbb{Y}_{TSDAk}$ can also be decomposed by the same rank as follows

$$\mathbb{Y}_{TSDAk} = \mathbb{D}_k \times_1 U_{k1} \times_2 U_{k2} \times_3 U_{k3} \times_4 U_{k4}. \tag{21}$$

where $\mathbb{D}_k \in \mathbb{R}^{r_{k1} \times r_{k2} \times r_{k3} \times r_{k4}}$ is a core tensor, which can be expressed based on the minimizing criterion about formula 4.3 of theorem 4.1 in [38] as follows:

$$\mathbb{D}_k = \mathbb{Y}_k \times_1 U_{k1}^T \times_2 U_{k2}^T \times_3 U_{k3}^T \times_4 U_{k4}^T. \tag{22}$$

According to (13) and (22),

$$\mathbb{D}_k = \mathbb{GF}_k \times_1 U_{k1}^T \times_2 U_{k2}^T \times_3 U_{k3}^T B_k \times_4 U_{k4}^T. \tag{23}$$

Combining (21) and (23) so as to make labels and non-local spatial information be utilized together, and performing joint

spatial–feature processing to compress the spatial and feature modes, the feature extraction process can be expressed as

$$\mathbb{Y}_{TSDAk} = \mathbb{GF}_k \times_1 V_{k1} \times_2 V_{k2} \times_3 V_{k3}B_k \times_4 V_{k4}. \quad (24)$$

where $V_{kn} = U_{kn}U_{kn}^T$, $n = 1, 2, 3, 4$, $U_{k1} \in \mathbb{R}^{W_1 \times r_{k1}}$, $U_{k2} \in \mathbb{R}^{W_2 \times r_{k2}}$, $U_{k3} \in \mathbb{R}^{p \times r_{k3}}$, $U_{k4} \in \mathbb{R}^{T_k \times r_{k4}}$, U_{kn} is the matrix of eigenvectors associated with the n-mode covariance matrix $Y_{k(n)}Y_{k(n)}^T$, where $Y_{k(n)}$ is the mode-n unfolding matrix of $\mathbb{Y}_k$, $\mathbb{GF}_k \in \mathbb{R}^{W_1 \times W_2 \times I_3 \times T_k}$, $V_{k3}B_k \in \mathbb{R}^{p \times I_3}$, and $\mathbb{Y}_{TSDAk} \in \mathbb{R}^{W_1 \times W_2 \times p \times T_k}$ is the resulting tensor holding p intrinsic features. $V_{k3}B_k$ is the tensor projection factor matrix obtained by CTSDA, and $\Lambda^{-1/2}U_3^T$ is the projection factor matrix of the traditional tensor method with low discriminability; the two are different in derivation and quantity. The method in this paper generates multiple projection matrices, which means multiple new projection directions in the third dimension, and the traditional tensor method generates one projection matrix in the third dimension, which means only one projection direction. Λ is the $p \times p$ eigenvector diagonal matrix corresponding to the first p eigenvalues of the covariance matrix $F_{(3)}F_{(3)}^T$, U_3 consists of the first r_{k3} eigenvectors corresponding to the covariance matrix $F_{(3)}F_{(3)}^T$, and $F_{(3)}$ is the three-mode flattening matrix of $\mathbb{F} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$.

To solve (24), the joint estimation of the orthogonal matrix U_{kn}, $\forall n$ can be obtained by the alternating least square (ALS) algorithm; see step 7 of Table 1. The matrix B_k is estimated by solving the generalized eigenvalue problem of $\mathbb{GF}_k$ with p. Once the reduced feature tensors $\mathbb{Y}_{TSDAk}$ are obtained, they can also be rearranged; the pixels represented in the tensors can be arranged in the original position in the image. Then, the new feature tensor $\mathbb{Y}_{CT} \in \mathbb{R}^{I_1 \times I_2 \times p}$ is obtained after the feature extraction. The procedure of the CTSDA method is summarized in Table 1; it is a hybrid of the unsupervised block clustering algorithm and TSDA.

Fig. 2 shows a schematic of the CTSDA method for feature extraction. The SAR feature tensor $\mathbb{F} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ generates several group tensors $\mathbb{GF}_k$ by clustering, uses the improved semisupervised discriminant analysis to find B_k, then uses ALS to obtain U_{kn}, $n = 1, 2, 4$. According to equation (24), TSDA is used to obtain the feature extraction tensor $\mathbb{Y}_{CT} \in \mathbb{R}^{I_1 \times I_2 \times p}$ to complete the feature extraction. The reduced feature tensor $\mathbb{Y}_{CT} \in \mathbb{R}^{I_1 \times I_2 \times p}$ is obtained by rearranging the reduced group tensors. Its mode-3 matricization leads to a 2-D feature matrix $Y_{CT} \in \mathbb{R}^{p \times I_1 I_2}$, in which each column represents the feature vector of a pixel, and each row represents a vectorized feature of the same kind that contains spatial information about the first two dimensions of the tensor $\mathbb{Y}_{CT} \in \mathbb{R}^{I_1 \times I_2 \times p}$. In other words, $Y_{CT} \in \mathbb{R}^{p \times I_1 I_2}$ is the mode-3 unfolding matrix of the tensor $\mathbb{Y}_{CT} \in \mathbb{R}^{I_1 \times I_2 \times p}$. The matrix Y_{CT} of the extracted features is used as input to the classifier support vector machine (SVM) to obtain the final classification results.

TABLE 1. Algorithm of the cluster-based tensorial semisupervised discriminant analysis.

Input: $\mathbb{F} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is an SAR feature tensor, equivalent to the sample set $F = \left[\mathbf{f}_1, \cdots, \mathbf{f}_l, \mathbf{f}_{l+1}, \cdots, \mathbf{f}_{I_1 I_2}\right] \in \mathbb{R}^{I_3 \times I_1 I_2}$ with l label information.

$Mlab \in \mathbb{R}^{I_1 \times I_2}$ is the label matrix corresponding to $\mathbb{F} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ with l label information.

r_{k1}, r_{k2} are rank parameters.

p is the dimension number of the extracted features, $p < I_3$.

W_1, W_2 is the block size.

nc is the cluster number.

Output: $\mathbb{Y}_{CT} \in \mathbb{R}^{I_1 \times I_2 \times p}$ is the SAR feature tensor after feature extraction.

Step 1: Segment $\mathbb{F}$ into blocks in the spatial domain with the fixed window $W_1 \times W_2$ and obtain f subtensors, $\mathbb{SF}_j \in \mathbb{R}^{W_1 \times W_2 \times I_3}$, $j = 1, 2, \cdots, f$.

Step 2: Segment $Mlab \in \mathbb{R}^{I_1 \times I_2}$ into blocks in the spatial domain with the fixed window $W_1 \times W_2$ and obtain f submatrices of the label matrix, $Slab_j \in \mathbb{R}^{W_1 \times W_2}$, $j = 1, 2, \cdots, f$.

Step 3: Use the unsupervised clustering algorithm to cluster all of the blocks into groups $\mathbb{GF}_k \in \mathbb{R}^{W_1 \times W_2 \times I_3 \times T_k}$, $k = 1, 2, \cdots, q$, according to nc from $\mathbb{SF}_j \in \mathbb{R}^{W_1 \times W_2 \times I_3}$, $j = 1, 2, \cdots, f$. T_k is the number of subtensors in the kth group.

Step 4: Find groups $\mathbb{GL}_k \in \mathbb{R}^{W_1 \times W_2 \times T_k}$ from $Slab_j \in \mathbb{R}^{W_1 \times W_2}$, $j = 1, 2, \cdots, f$ according to the block's index of clusters; the equivalent of that submatrix of labels generates clusters.

Step 5: TSDA initialization. Set the rank parameters in each mode of cluster $\mathbb{GF}_k \in \mathbb{R}^{W_1 \times W_2 \times I_3 \times T_k}$ as $\left(r_{k1}, r_{k2}, p, r_{k4}\right)$, $r_{k4} = T_k$.

Step 6: Estimate the matrix $B_k \in \mathbb{R}^{p \times I_3}$ by solving the improved SDA generalized eigenvalue problem of $\mathbb{GF}_k$ with p and $\mathbb{GL}_k$.

Step 7: Compute $\mathbb{Y}_k = \mathbb{GF}_k \times_3 B_k$. The ALS loop is

$i = 0, \forall n = 1, 2, 3, 4, U_{kn,0} \leftarrow$ first r_{kn} or p $(n = 3)$ eigenvectors of matrix $E\left[Y_{k(n)}Y_{k(n)}^T\right]$.

Repeat

ALS loop:

For all n = 1, 2, 3, 4, do

$\hat{\mathbb{Y}}_{k,i} \leftarrow \mathbb{Y}_k \times_q U_{kq,i}^T \times_r U_{kr,i}^T \times_s U_{ks,i}^T$ with $q, r, s \neq n$

$E\left[\hat{Y}_{k(n),i}\hat{Y}_{k(n),i}^T\right]$ eigenvalue decomposition

$U_{kn,i+1} \leftarrow$ first r_{kn} or p $(n = 3)$ eigenvectors

$i \leftarrow i + 1$

End

Until equation (19) converges;

$\mathbb{Y}_k$ is decomposed as follows:

$\mathbb{Y}_k = \mathbb{C}_k \times_1 U_{k1} \times_2 U_{k2} \times_3 U_{k3} \times_4 U_{k4}$.

Step 8: Compute $V_{kn} = U_{kn}U_{kn}^T, n = 1, 2, 3, 4$.

Step 9: Calculate $\mathbb{Y}_{TSDAk} \quad \mathbb{GF}_k \times_1 V_{k1} \times_2 V_{k2} \times_3 V_{k3}B_k \times_4 V_{k4}$.

Step 10: Rearrange the reduced tensors $\mathbb{Y}_{TSDAk} \in \mathbb{R}^{W_1 \times W_2 \times p \times T_k}$ back and return the dimensionality-reduced SAR feature tensor $\mathbb{Y}_{CT} \in \mathbb{R}^{I_1 \times I_2 \times p}$.

IV. RESULTS AND DISCUSSION

In this section, the proposed method is applied to the classification of simulated and real SAR images. To validate the performance of the proposed method, we use both

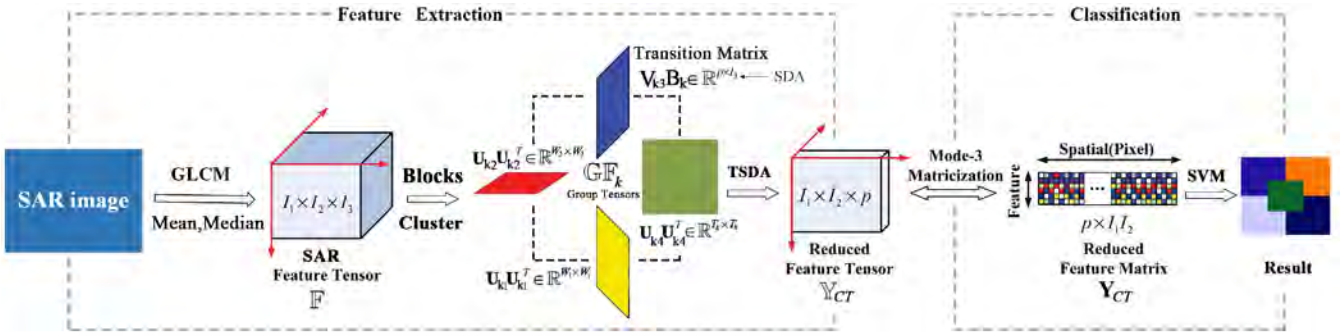


FIGURE 2. Schematic of the proposed CTSDA method for feature extraction.

types of images for quantitative evaluation and visualization. We mainly compare our results with those of previous studies [14], [15], [26], [28], [31], in which their parameters are tuned to obtain the best results. SVM [39] is chosen as the classifier to perform the supervised classification. No more than 10% of the samples are selected as the training set, and all of the samples are used as the testing set for the classifier. The experimental results are obtained by running the code on a DELL computer with an Intel (R) Core (TM) i7 CPU, 3.4 GHz, 16 GB RAM with MATLAB 2016(a) on Windows 10 (64-bit operating system).

The simulated SAR images in the experiments have inhomogeneous intensities and are generated by adding multiplicative noise with a gamma distribution to pure images. The n-view simulated SAR image is generated by averaging n independent realizations of speckle noise. The simulated SAR images have different textural regions of the same number of views. As shown in the first column of Fig. 7, the number of views is 3, from top to bottom, and the numbers of texture types are as follows: 2, 3, 4, and 5. The size of each image is 256 × 256. The test images are named Syn1, Syn2, Syn3 and Syn4.

In addition, four real SAR images (SAR1, SAR2, SAR3 and SAR4) are tested, shown in Fig. 10a. SAR1 has a size of 475 × 446 and covers the China Lake Airport, California, with a Ku-band radar with a 3-m resolution. SAR2 has a size of 321 × 258 and covers a pipeline over the Rio Grande River near Albuquerque, New Mexico, with a Ku-band radar with a 1-m resolution. SAR3 has a size of 256 × 256 and covers a suburb of Beijing, China. SAR4 has a size of 505 × 476 and covers an area of the south Russian steppes northeast of the Black Sea with a 15-m resolution X-band radar. Based on the GLCM, mean values and median values, 24 features are extracted as described in Section III (A). Each image is then formulated as a third-order tensor for our method.

A. MULTIFEATURE AVAILABILITY AND EXTRACTION VALIDITY

In this part, the multifeature extraction is analyzed to verify its effectiveness in obtaining satisfactory results. We perform

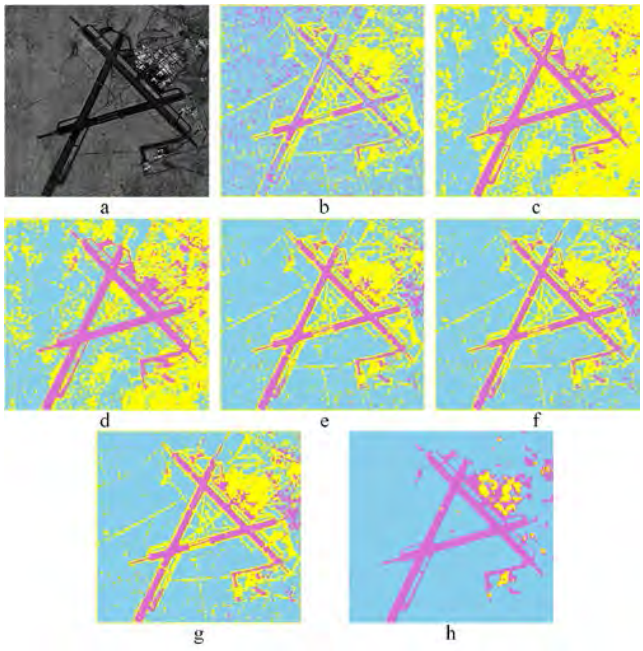


FIGURE 3. Comparison of (a) the original SAR1 image with (b) gray-level co-occurrence matrix (GLCM), (c) mean, (d) median, (e) GLCM and mean (Gmean), (f) GLCM and median (Gmedian), (g) multifeature, and (h) multifeature extraction of CTSDA.

an experiment on the original image SAR1 (Fig. 3a) to show the influence of multifeature extraction and the impact of the classification results. In Table 2, the accuracy of extracted feature of CTSDA is the best, followed by the one with multifeature. Fig. 3h has fewer miscellaneous points than Figs. 3b–g because the essential features of the multiple features are extracted, and information about the distribution and statistics is included to obtain better results. Therefore, the multifeature extraction is important. In Table 3, the classification running time of the extracted features is lower than that of the multifeature, and the feature extraction time is not included. These results show that feature extraction improves the classification results and reduces the time of classification.

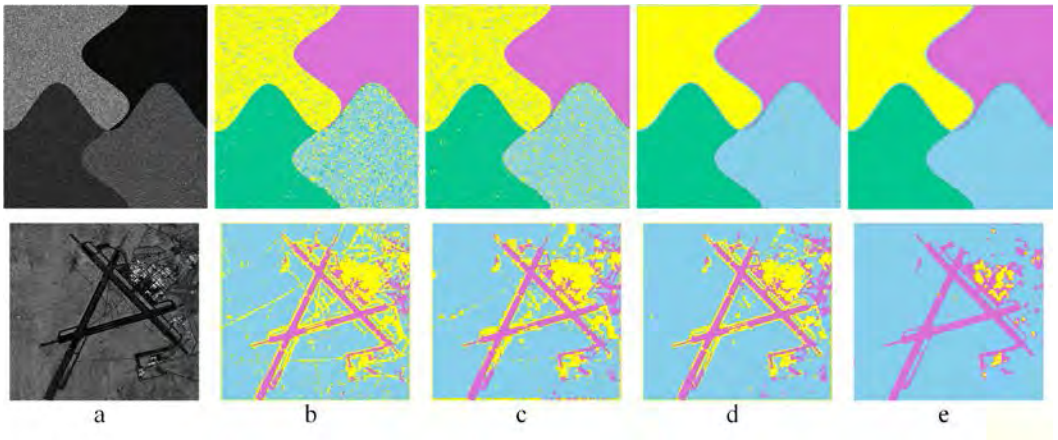


FIGURE 4. Comparison of (a) the original image with classification of (b) no added information, (c) non-local spatial information, (d) labels information, and (e) combined information.

TABLE 2. Classification accuracies of different features.

GLCM	Mean	Median	Gmean	Gmedian	Multi-feature	Extracted feature
65.89	52.87	60.01	75.0635	75.0602	75.69	95.07

TABLE 3. Classification running times of different features.

GLCM	Mean	Median	Gmean	Gmedian	Multi-feature	Extracted feature
51.42s	10.28s	8.87s	53.88s	54.25s	58.83s	30.52s

B. VALIDITY OF LABELS AND NON-LOCAL SPATIAL INFORMATION ASSOCIATION

To evaluate the effectiveness of labels and non-local spatial information association, we compared the results of no added information, non-local spatial information, labels information and combined information, respectively corresponding to the classification of tensor, cluster-based tensorial principal component analysis (CTPCA), tensor semisupervised discriminant analysis, and cluster-based tensorial semisupervised discriminant analysis (CTSDA). For convenience, we conduct the experiments on the Syn3 and SAR1. We set the block size is 10, the rank parameters are set to the block size for each cluster, the extracted feature dimensionality is set to 10, and the numbers of clusters are set as 4 and 3 on Syn3 and SAR1 respectively.

In Table 4, the accuracy of joint information is the best, followed by the one with labels information, followed by the one of non-local spatial information, and the worst is the one without any information. Fig.4e shows the minimum number of misclassification points, because the labeling information improves the discrimination of features, while the spatial non-local information promotes the tightness within classes and the distance between classes, and the combination of the two makes the classification result optimal. From the results, we can see that information association indeed enhances the classification performance.

TABLE 4. Accuracies of different information on images.

Image	No added	Non-local spatial	labels	combined
Syn3	89.43	91.63	95.98	97.72
SAR1	75.69	80.11	82.68	95.07

C. INVESTIGATION OF BLOCK SIZE AND CLUSTER NUMBER

When the tensor-based SAR data set is segmented into blocks in the spatial domain with the fixed window $W_1 \times W_2$, W_1 and W_2 determine the number of pixels for each subtensor. We set $W_1 = W_2$, and when the subtensors are grouped into clusters, the cluster number nc determines the number of clusters. Fig. 5 shows the overall accuracy (OA) of the proposed method with the variations in W_1, W_2, and nc when the extracted feature dimensionality is set to 10 and the rank parameters r_{k1} and r_{k2} of the spatial mode are set to the block size for each cluster on Syn3 and SAR1. The results show that, when the block size is too large (e.g., larger than 11), because different land covers may be included in a subtensor, the local correlation will be destroyed, which will degrade its performance. Based on Fig. 5, we set the block size to 10 because it gives the best precision in many experiments. In addition, there are some variations when nc takes different values. The nonlocal correlation cannot be reflected well when the cluster number nc is too large (e.g., larger than 8) because the same original classes of different regions are divided into different categories. The accuracy is lower than other after clustering when the cluster number is 1, which shows that the nonlocal correlation is effective. When the block size is fixed at 10, similar performance is obtained when nc is set to 2, 4, and 6 on Syn3 and when nc is set to 3, 7, and 9 on SAR1. Considering the performance and computational complexity, we set the block size to 10 and set the cluster number to the number of categories contained in the images, which is 4 in Syn3 and 3 in SAR1.

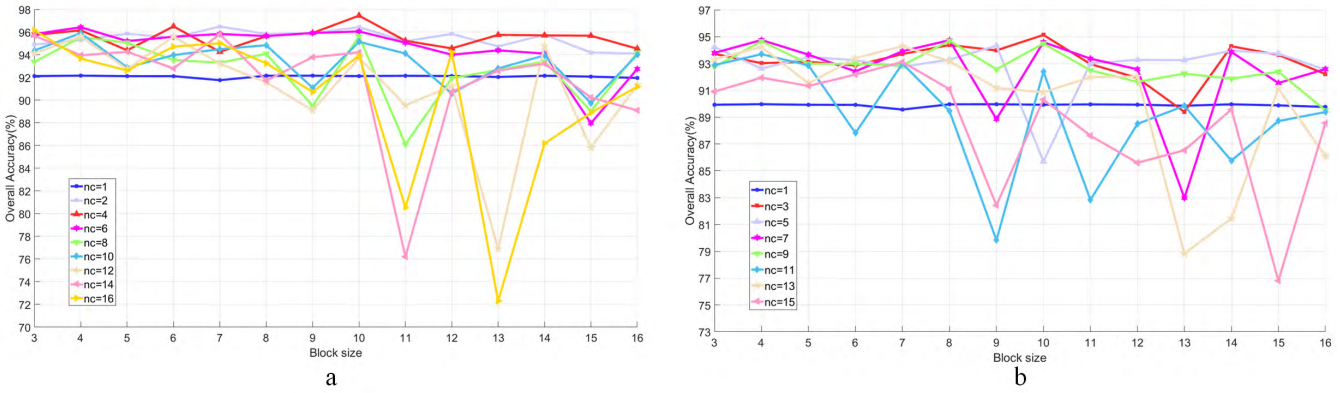


FIGURE 5. OA with the variations in block size and number of clusters: (a) Syn3; (b) SAR1.

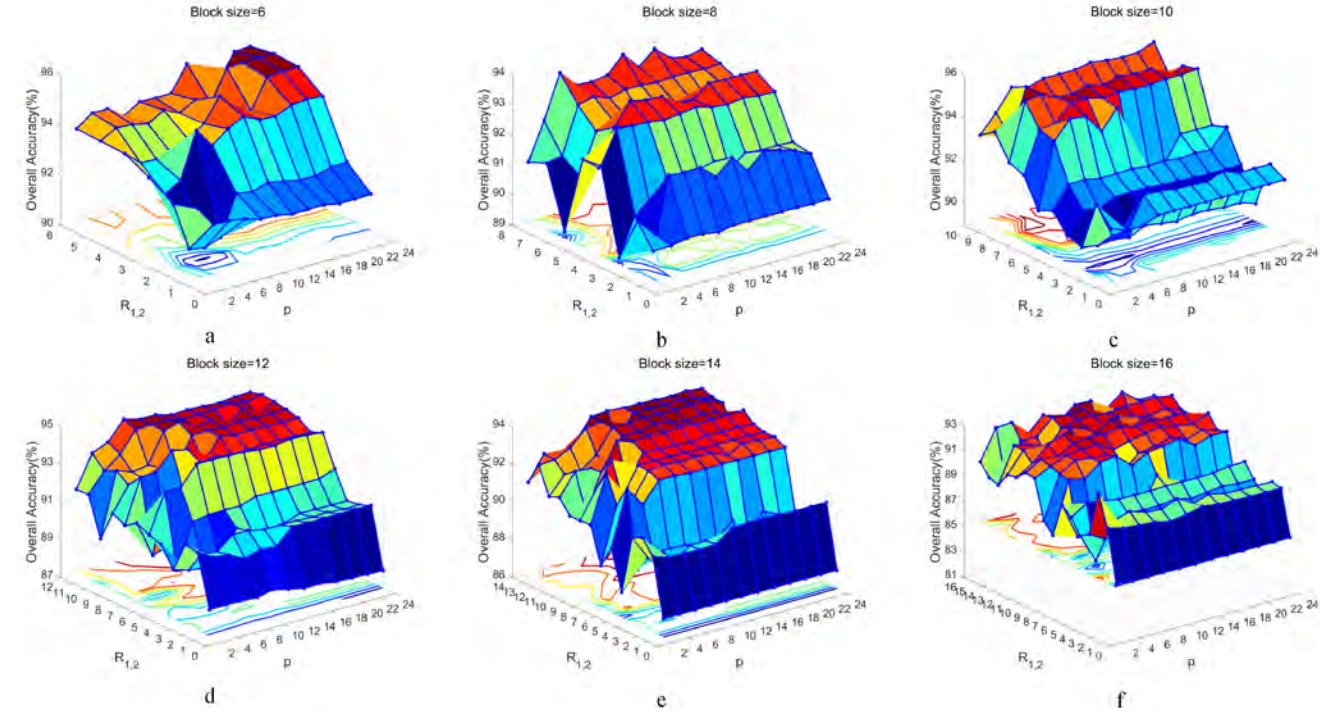


FIGURE 6. OA with the variation in the spatial rank $R_{1,2}$ on different block sizes.

D. INVESTIGATION OF SPATIAL RANK WITH DIFFERENT BLOCK SIZES

We also conduct experiments to investigate the effects of the spatial rank on the proposed method. The $R_{1,2}$ value represents the spatial rank, which is the number of first eigenvectors of the covariance matrix in the algorithm. For n = 1 and 2, we set $r_{k1} = r_{k2} = R_{1,2}$ of each cluster tensor in the experiments. The influence of the quantity number p of the retained features is conceded. Fig. 6 shows the evolution of the OA value when $R_{1,2}$ varies from 1 to the block size and p varies from 2 to 24 with a step of 2 on SAR1.The rank parameters reflect the degree of spatial correlation. If the spatial correlation is weak, the rank will be large or even be a full rank; otherwise, the rank will be small. In the proposed method, in the spatial domain, each subtensor consists of neighbor blocks. Moreover, as shown in the previous analysis, to reduce the variation and constrain the classification accuracy to be higher than 92%, the spatial block size should not be too large. Thus, for an appropriate block size, it is reasonable that if the spatial rank parameters are set high to reduce the spatial correlation, more main components are extracted. This is consistent with the results shown in Fig. 6. When the block size is smaller than 14, if the value of $R_{1,2}$ is too low, the OA value decreases; as $R_{1,2}$ increases, the value of OA improves, when the value of $R_{1,2}$ exceeds the median block size, the value of OA stabilizes, and the

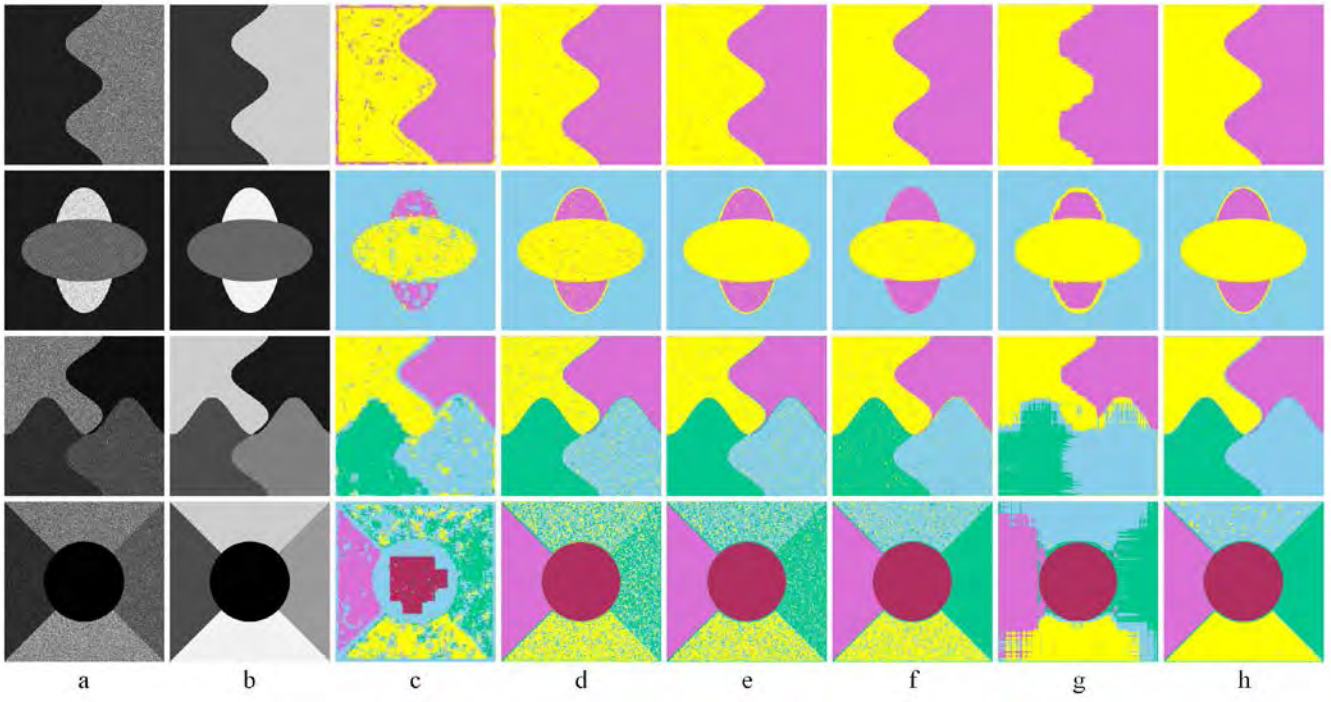


FIGURE 7. Results from different methods with simulated SAR images: (a) simulated SAR images; (b) ground truths; (c) TSNE; (d) PCA; (e) SDA; (f) TLPP; (g) LRTA$_{dr} - (K_1, K_2, p)$; and (h) CTSDA.

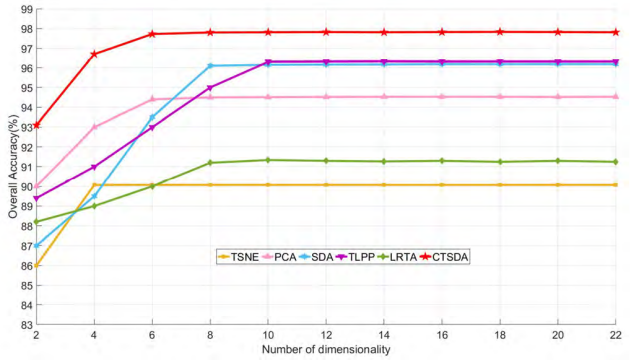


FIGURE 8. Variations in the OA value with the extracted features for different methods: TSNE, PCA, SDA, TLPP, LRTA (LRTA$_{dr} - (K_1, K_2, p)$), and CTSDA.

proposed method will have better performance. Therefore, we set the spatial rank parameters to be greater than the median of the block size, $R_{1,2} = 8$, in all of the experiments because the block size is 10.

E. EXPERIMENT ON SIMULATED DATA

In this section, experiments are conducted on the four simulated SAR images to demonstrate the performance of the proposed feature extraction method for SAR classification. As mentioned previously, five extracted feature methods are used for comparison. The ground truths are used to calculate the accuracy of the classification results to evaluate the algorithms. In our method, we set the block size $W_1 = 10$, $W_2 = 10$, the cluster number to the number of classes in the image, and the spatial rank parameters $r_{k1} = r_{k2} = 8$. To evaluate the classification performance, training samples are selected randomly in all of the comparison methods. We record the average results of 10 runs when the extracted feature dimensionality is set to 10. The classification results for the overall accuracy and the Kappa coefficient are shown in Tables 5 and 6, respectively, and the classification maps are shown in Fig. 7. The results show that the proposed method achieves much better classification performance in terms of OA and Kappa than the other state-of-the-art methods (Tables 5 and 6). For example, the proposed method

TABLE 5. Accuracies of different algorithms on simulated SAR images.

image	TSNE	PCA	SDA	TLPP	LRTA	CTSDA
Syn1	91.54	97.73	98.5	99.27	97.85	99.35
Syn2	90.76	96.66	97.75	98.11	96.77	98.06
Syn3	90.07	94.51	96.15	96.19	91.34	97.72
Syn4	59.28	73.96	78.97	89.84	87.06	96.06

TABLE 6. Kappa values of different algorithms on simulated SAR images.

image	TSNE	PCA	SDA	TLPP	LRTA	CTSDA
Syn1	0.83	0.95	0.97	0.99	0.96	0.99
Syn2	0.80	0.93	0.95	0.96	0.93	0.96
Syn3	0.87	0.93	0.95	0.95	0.88	0.97
Syn4	0.49	0.67	0.74	0.87	0.84	0.95

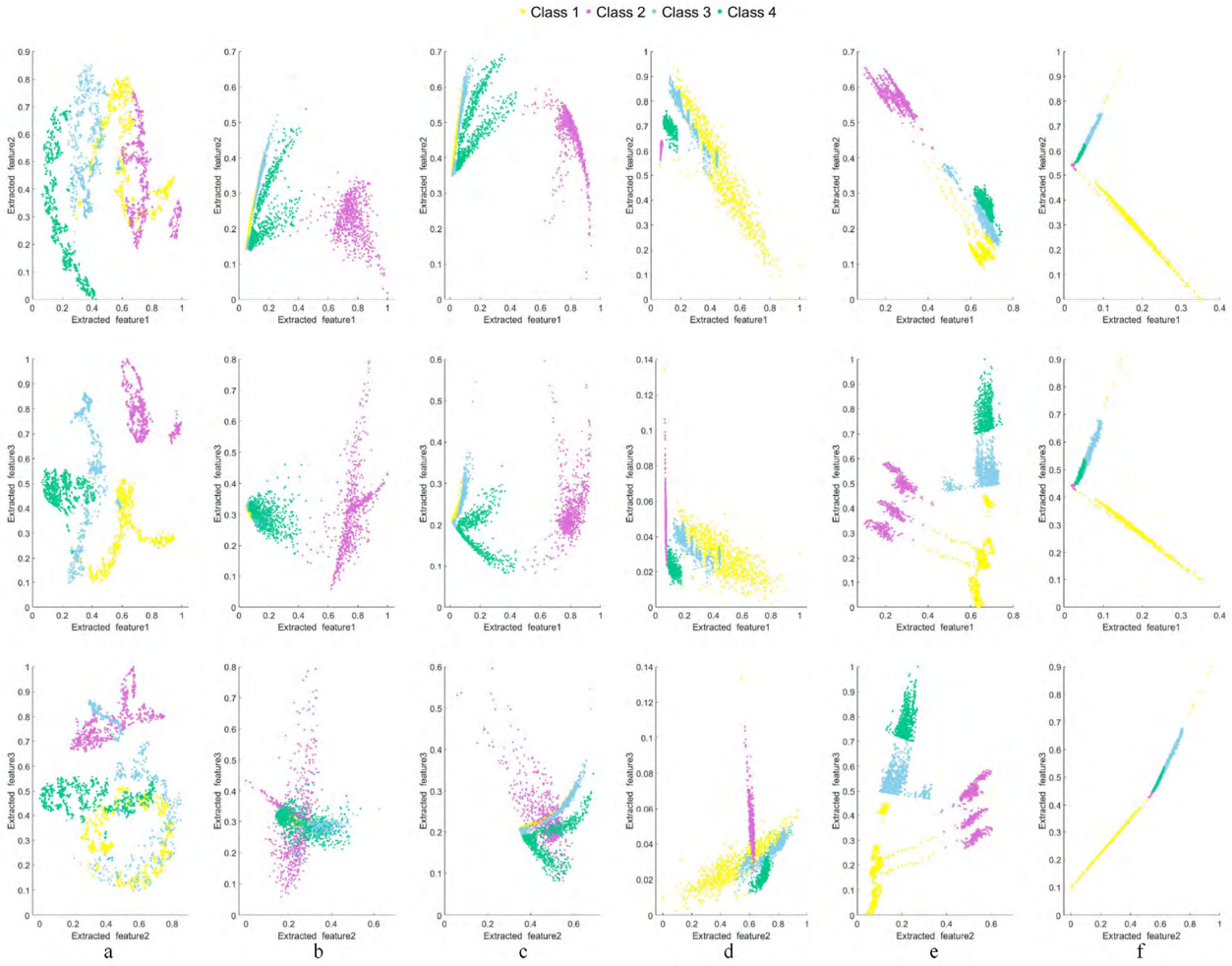


FIGURE 9. Scatterplots of the extracted features using: (a) TSNE, (b) PCA, (c) SDA, (d) TLPP, (e) LRTA$_{dr}$ − (K_1, K_2, p), and (f) CTSDA.

is approximately 6% and 0.08 better than the second-best method for OA and Kappa, respectively, on Syn4. This result demonstrates that the proposed method is an effective discriminative feature extraction method. The results also show that the tensor-based methods achieve better performance than the vector-based methods.

As shown in Fig. 7, the results from the first to third rows show that the categories are basically distinguished; in addition, the classifications based on CTSDA and TLPP are clearly better than the other methods for the same image. The classifications from the different methods are similar for image Syn1, as shown in the first row, due to the high contrast between the two classes of pixels. For images Syn2 and Syn3, which are shown in the second and third rows, respectively, TSNE is not very effective because the different classes of pixels are confused in the classifications, which means that the nonlinear methods are not suitable for SAR image classification. As shown in the fourth row (Syn4), the classification based on CTSDA is best and the most similar to the ground truth image, and the noise is minimal. The results of TLPP and LRTA$_{dr}$ − (K_1, K_2, p) are good, and the results of the other three methods are not as good. These results demonstrate the superiority of the CTSDA method in complex scenarios.

1) SENSIBILITY ANALYSIS FOR THE DIMENSION OF THE EXTRACTED FEATURES

We also investigate the effect of the dimensionality of the extracted features on the performance of our method. Fig. 8 shows the OA results with the SVM classifier on Syn3 when the dimensionality ranges from 2 to 22 with a step of 2. The OA improves with increasing dimensionality and stabilizes when the dimensionality exceeds 8 of our method. As a result, we set the dimensionality to 10. Fig. 8 also illustrates the advantage of the proposed method over the other comparison methods when the dimensionality is low, which further demonstrates the ability of the proposed method in feature extraction.

2) VISUALIZATION OF FEATURE EXTRACTION

To intuitively visualize the effects of the various methods on the feature extraction, scatterplots for Syn3 are plotted

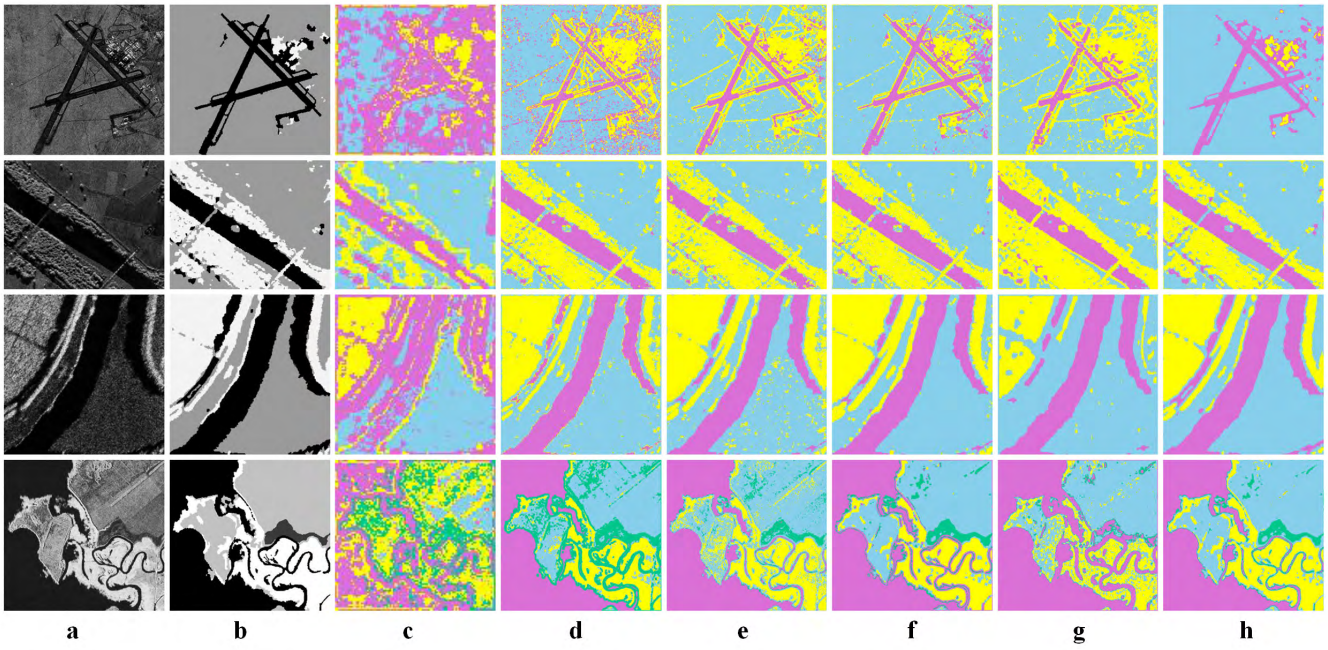


FIGURE 10. Results of different methods with real SAR images: (a) real SAR images; (b) ground truths; (c) TSNE; (d) PCA; (e) SDA; (f) TLPP; (g) LRTA$_{dr}$ $-$ (K_1, K_2, p); and (h) CTSDA.

and analyzed. For illustration, we set the dimensionality of the intrinsic extracted feature subspace to p $=$ 3. In Fig. 9, the points represent all of the samples, and the four colors (yellow, purple, blue and green) represent four different types of textures.

Fig. 9a shows the scatterplots for TSNE, which show much correlations among the different classes, making the following classification difficult. Fig. 9b shows scatterplots of the principal components extracted by PCA. The principle components are very close in the feature space. The scatterplots of the three distinctive components obtained by SDA are shown in Fig. 9c. The distinctive components are partially superimposed, indicating unsatisfactory classification performance. For comparison, two tensor manifold learning methods, the TLPP and the LRTA$_{dr}$ $-$ (K_1, K_2, p), are also applied. The scatterplots of the features extracted by TLPP are shown in Fig. 9d. All of the samples achieved separation except for small intersections, which indicates that the TLPP method captures the discriminative information between each class well. Fig. 9e shows that the features extracted by LRTA$_{dr}$ $-$ (K_1, K_2, p) are more distinct than those of the previous feature extraction methods. However, samples of a class are not grouped together, then the classification results may be unstable. Finally, the proposed CTSDA method is applied to the feature tensor to obtain an intrinsic feature set. The scatterplots are shown in Fig. 9f. The extracted features are more concentrated and discriminable. The samples of each class are clustered and distinct from the other classes, which indicates better classification performance. The proposed CTSDA method produces more separable features, which facilitates the classification process.

F. EXPERIMENT ON REAL DATA

In this section, the four real SAR images described previously are used for further analysis. The methods are the same as those used on the simulated SAR images. The results are shown in Fig. 10. These original real images (SAR1, SAR2, SAR3 and SAR4) have three, three, three and four types of different regions, respectively, as shown in Fig. 10a. The ground truths of SAR1, SAR2, SAR3 and SAR4 are shown in Fig. 10b. The evaluation of classification method is based on the visual inspection of the classification, the run time, the accuracy, and the Kappa coefficient. The extracted feature dimensionality is set to 10, which gives the best results. The results in Fig. 10 show that the CTSDA method gives the best classification results of all of the comparison methods. The CTSDA method gives clearer contours, more precise locations, and better regional consistency in every section, and the results are most similar to the ground truth images with minimal noise. The results of TLPP (Fig. 10f) are better than all of the other methods except CTSDA; the results have favorable regional consistency, but there are small misclassifications, and the terrain boundary classification is fuzzy in SAR4. The classification results of LRTA$_{dr}$ $-$ (K_1, K_2, p) (Fig. 10g) are similar to those of TLPP, but there are more misclassified points, and the classification results are unstable. The results of SDA and PCA (Fig. 10e and Fig. 10d, respectively) are similar; their advantages and weaknesses in different SAR images are different. The classification results of TSNE (Fig. 10c) are the poorest of all of the methods. The results shown in Fig. 10 demonstrate that the proposed method is suitable for SAR image classification and obtains the optimal results. Tables 7-9 provide quantitative

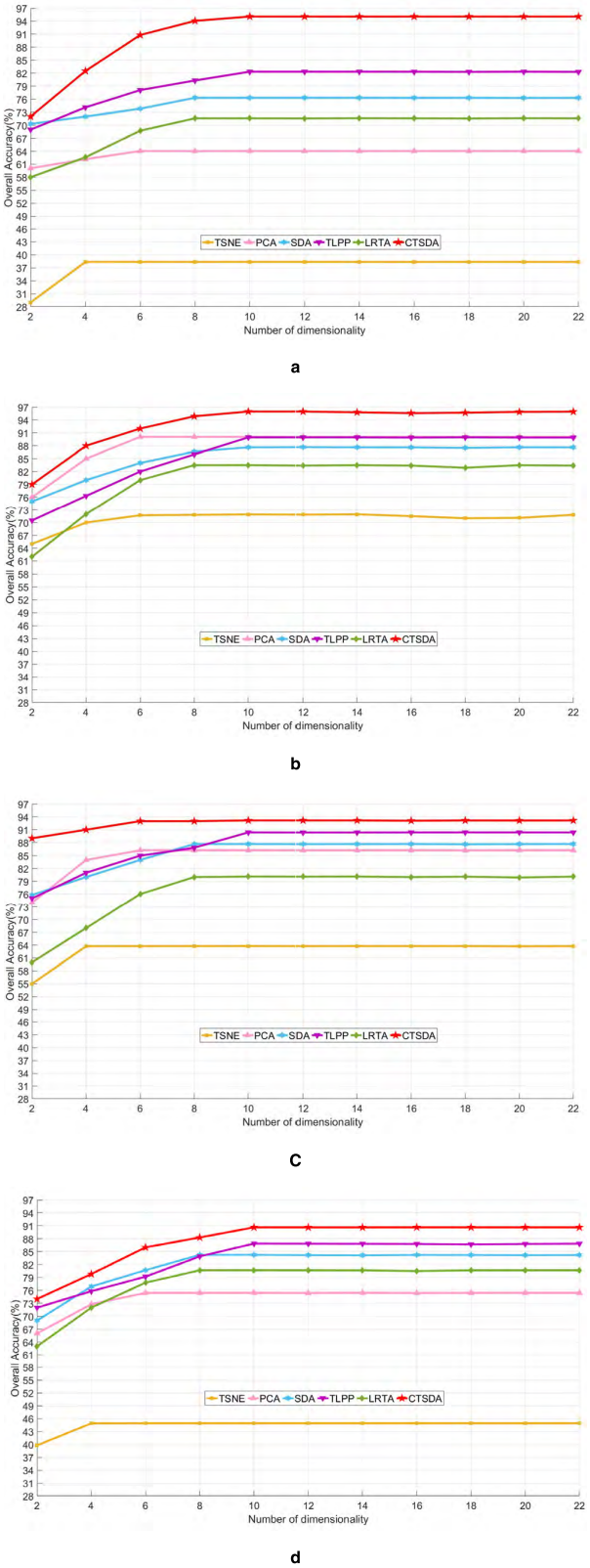


FIGURE 11. OA versus extracted dimensionality of different images: (a) SAR1; (b) SAR2; (c) SAR3; (d) SAR4.

evaluations of the different methods. Although CTSDA has a greater running time than the other methods, it has a shorter running time on complex multiclass images such as SAR4.

TABLE 7. Running times of different algorithms on real SAR images.

image	TSNE	PCA	SDA	TLPP	LRTA	CTSDA
SAR1	2033.11s	0.56s	971.44s	608.96s	0.76s	1309.09s
SAR2	1205.27s	0.64s	133.31s	111.39s	2.57s	141.71s
SAR3	2232.27s	0.23s	82.41s	74.61s	0.66s	74.86s
SAR4	1951.37s	0.38s	1295.24s	1184.61s	1.78s	952.35s

TABLE 8. Accuracies of different algorithms on real SAR images.

image	TSNE	PCA	SDA	TLPP	LRTA	CTSDA
SAR1	38.37	64.06	76.32	82.34	71.60	95.07
SAR2	71.90	90.07	87.63	89.95	83.51	95.98
SAR3	63.77	86.25	87.66	90.38	80.12	93.17
SAR4	44.95	75.42	84.27	86.87	80.69	90.65

TABLE 9. Kappa values of different algorithms on real SAR images.

image	TSNE	PCA	SDA	TLPP	LRTA	CTSDA
SAR1	0.12	0.29	0.46	0.55	0.39	0.84
SAR2	0.53	0.84	0.79	0.83	0.74	0.93
SAR3	0.45	0.79	0.81	0.85	0.68	0.89
SAR4	0.27	0.66	0.77	0.81	0.71	0.86

Because of its good classification performance, the running time of CTSDA is acceptable. The proposed method has the highest accuracy, which is approximately 12% higher than that of the second-best method (TLPP) on SAR1. Moreover, the proposed method has the highest robustness, which is reflected by the Kappa coefficient. In general, the method outperforms the others in terms of accuracy and robustness.

1) SENSIBILITY ANALYSIS OF THE DIMENSIONS OF THE EXTRACTED FEATURES

The four real SAR images are used to further analyze the classification performance with respect to the dimensions of the extracted features. Fig. 11 shows the OAs of the six algorithms with increasing dimensionality from 2 to 22. The OA of the proposed algorithm remains stable in the range of [10]–[22]. This observation means that the proposed algorithm is not sensitive to the feature dimensionality. In addition, the proposed algorithm provides better performance, especially when the feature dimensionality is low.

V. CONCLUSION

In this paper, we presented a new method for feature extraction of SAR images that is based on clusters. The results validate that adding multiple features, labels and nonlocal spatial manifolds can benefit the feature extraction results and increase the information about the images. Furthermore, using the tensor form can decrease the loss of spatial neighborhood information due to the use of vector samples. The cluster strategy was introduced to fuse the labeling tensor

projection and the semisupervised improved discriminant analysis to form multiple new tensor projection factor matrices (i.e., multiple new projection directions). Due to the added benefits, the robustness was enhanced, and the classification accuracy was improved significantly by using the extracted features. Comparison experiments based on simulated and real SAR images clearly demonstrate the efficiency and advantages of the proposed feature extraction method. Moreover, the proposed method has an acceptable computational cost for complex multiclass SAR images. These added benefits are general for SAR image feature extraction and may be suitable for use in other applications in the field of SAR image feature extraction as well as in other areas where tensor feature extraction methods, such as LRTA and TLPP, can be applied.

This method provides an improvement in feature discrimination for SAR feature extraction and applications. Future research will focus on developing more efficient algorithms for use with large-scale SAR images and combining deep learning technology to realize automatic feature extraction.

ACKNOWLEDGMENT

The authors would like to thank Jinliang An for providing real SAR image data, technical support, and proofreading. They would also like to thank the editors and anonymous reviewers for their insightful comments and suggestions, which have significantly improved this paper.

REFERENCES

[1] T. L. M. Barreto, R. A. S. Rosa, C. Wimmer, J. R. Moreira, L. S. Bins, F. A. M. Cappabianco, and J. Almeida, "Classification of detected changes from multitemporal high-res Xband SAR images: Intensity and texture descriptors from superpixels," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 12, pp. 5436–5448, Dec. 2016.

[2] K. Kayabol and J. Zerubia, "Unsupervised amplitude and texture classification of SAR images with multinomial latent model," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 561–572, Feb. 2013.

[3] H. Lang and S. Wu, "Ship classification in moderate-resolution SAR image by naive geometric features-combined multiple kernel learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1765–1769, Oct. 2017.

[4] V. Liesenberg, C. R. de Souza Filho, and R. Gloaguen, "Evaluating moisture and geometry effects on L-band SAR classification performance over a tropical rain forest environment," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 12, pp. 5357–5368, Dec. 2016.

[5] F. Gao, F. Ma, J. Wang, J. Sun, E. Yang, and H. Zhou, "Visual saliency modeling for river detection in high-resolution SAR imagery," *IEEE Access*, vol. 6, pp. 1000–1014, Feb. 2018.

[6] F. Chen, H. Zhou, C. Grecos, and P. Ren, "Segmenting oil spills from blurry images based on alternating direction method of multipliers," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 6, pp. 1858–1873, Jun. 2018.

[7] S. C. Steele-Dunne, H. McNairn, A. Monsivais-Huertero, J. Judge, P.-W. Liu, and K. Papathanassiou, "Radar remote sensing of agricultural canopies: A review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 2249–2273, May 2017.

[8] G. Dong-Dong, T. Tang, L. Zhao, and J. Lu, "A feature combining spatial and structural information for SAR image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Milan, Italy, Jul. 2015, pp. 4396–4399.

[9] X. Zhang, Y. Wang, D. Li, Z. Tan, and S. Liu, "Fusion of multifeature low-rank representation for synthetic aperture radar target configuration recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1402–1406, Sep. 2018.

[10] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 153–158, Feb. 1997.

[11] B. Bigdeli and P. Pahlavani, "High resolution multisensor fusion of SAR, optical and LiDAR data based on crisp vs. fuzzy and feature vs. decision ensemble systems," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 52, pp. 126–136, Jun. 2016.

[12] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Boston, MA, USA: Elsevier, 2009, pp. 13–249.

[13] L. Shi, L. Zhang, L. Zhao, L. Zhang, P. Li, and D. Wu, "Adaptive Laplacian eigenmap-based dimension reduction for ocean target discrimination," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 7, pp. 902–906, Jul. 2016.

[14] M. Yu, L. Zhao, S. Zhang, B. Xiong, and G. Kuang, "SAR target recognition using parametric supervised t-stochastic neighbor embedding," *Remote Sens. Lett.*, vol. 8, no. 9, pp. 849–858, May 2017.

[15] A. K. Mishra, "Validation of PCA and LDA for SAR ATR," in *Proc. Tencon IEEE Region 10 Conf.*, Hyderabad, India, Nov. 2008, pp. 1–6.

[16] R. Huan, R. Liang, and Y. Pan, "SAR target recognition with the fusion of LDA and ICA," in *Proc. IEEE Int. ICIECS*, Wuhan, China, Dec. 2009, pp. 1–5.

[17] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[18] B. Ren, B. Hou, Z. Wen, W. Xie, and L. Jiao, "PolSAR image classification via multimodal sparse representation-based feature fusion," *Int. J. Remote Sens.*, vol. 39, no. 22, pp. 7861–7880, Jun. 2018.

[19] N. Huyan, X. Zhang, H. Zhou, and L. Jiao, "Hyperspectral anomaly detection via background and potential anomaly dictionaries construction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2263–2276, Apr. 2019.

[20] X. Zhang, C. Li, J. Zhang, Q. Chen, J. Feng, L. Jiao, and H. Zhou, "Hyperspectral unmixing via low-rank representation with space consistency constraint and spectral library pruning," *Remote Sens.*, vol. 10, no. 2, p. 339, Feb. 2018.

[21] S. Ji, "Computational network analysis of the anatomical and genetic organizations in the mouse brain," *Bioinformatics*, vol. 27, no. 23, pp. 3293–3299, Oct. 2011.

[22] C. Sauvage, S. Poirriez, M. Manto, P. Jissendi, and C. Habas, "Reevaluating brain networks activated during mental imagery of finger movements using probabilistic tensorial independent component analysis (TICA)," *Brain Imag. Behav.*, vol. 5, no. 2, pp. 137–148, Jun. 2011.

[23] S. Yuan, X. Mao, and L. Chen, "Multilinear spatial discriminant analysis for dimensionality reduction," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2669–2681, Jun. 2017.

[24] X. He, D. Cai, and P. Niyogi, "Tensor subspace analysis," in *Proc. Adv. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, 2005, pp. 499–506.

[25] X. Han and L. Clemmensen, "Regularized generalized eigen-decomposition with applications to sparse supervised feature extraction and sparse discriminant analysis," *Pattern Recognit.*, vol. 49, pp. 43–54, Jan. 2016.

[26] N. Renard and S. Bourennane, "Dimensionality reduction based on tensor modeling for classification methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1123–1131, Apr. 2009.

[27] J. An, X. Zhang, and L. C. Jiao, "Dimensionality reduction based on group-based tensor model for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 10, pp. 1497–1501, Oct. 2016.

[28] Y.-J. Deng, H.-C. Li, L. Pan, L.-Y. Shao, Q. Du, and W. J. Emery, "Modified tensor locality preserving projection for dimensionality reduction of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 277–281, Feb. 2018.

[29] M. Tao, F. Zhou, Y. Liu, and Z. Zhang, "Tensorial independent component analysis-based feature extraction for polarimetric SAR data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2481–2495, May 2015.

[30] X. Huang, H. Qiao, B. Zhang, and X. L. Nie, "Supervised polarimetric SAR image classification using tensor local discriminant embedding," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2966–2979, Jun. 2018.

[31] S. Wang, J. Lu, X. Gu, H. Du, and J. Yang, "Semi-supervised linear discriminant analysis for dimension reduction and classification," *Pattern Recognit.*, vol. 57, pp. 179–189, Sep. 2016.

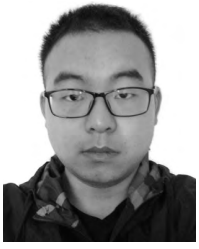
[32] D. Cai, X. He, and J. Han, "Semi-supervised discriminant analysis," in *Proc. IEEE Int. Conf. Comput. Vis.*, Rio de Janeiro, Brazil, Oct. 2007, pp. 14–21.

[33] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
[34] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
[35] D. A. Clausi, "Comparison and fusion of co-occurrence, Gabor and MRF texture features for classification of SAR sea-ice imagery," *Atmos.-Ocean*, vol. 39, no. 3, pp. 183–194, Sep. 2001.
[36] C. G. Rafael and E. W. Richard, *Digital Image Processing*. Upper Saddle River, NJ, USA: Pearson, 2017, pp. 188–199.
[37] P. Gehler and S. Nowozin, "On feature combination for multiclass object classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Kyoto, Japan, Sep./Oct. 2009, pp. 221–228.
[38] L. D. Lathauwer, B. De Moor, and J. Vandewalle, "On the best rank-1 and rank-(R_1, R_2, ... R_N) approximation of higher-order tensors," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1324–1342, Apr. 2000.
[39] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, 2011.
[40] X. He and P. Niyogi, "Locality preserving projections," in *Proc. Conf. Adv. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, 2003, pp. 1–8.
[41] X. Tang and L. Jiao, "Fusion similarity-based reranking for SAR image retrieval," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 2, pp. 242–246, Feb. 2017.
[42] J. An, X. Zhang, H. Zhou, and L. Jiao, "Tensor-based low-rank graph with multimanifold regularization for dimensionality reduction of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4731–4746, Aug. 2018.

XIAOYING WU received the B.S. degree in computer science and technology from Henan Normal University and the M.S. degree from the School of Computer and Information Technology, Beijing Jiaotong University, China. She is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, Tianjin University of Technology. Her research interests include image interpretation and machine learning.

XIANBIN WEN received the Ph.D. degree from Northwestern Polytechnical University, Xi'an, China, in 2005. He is currently a Professor with the School of Computer Science and Engineering, Tianjin University of Technology, Tianjin, China. His research interests include image interpretation, machine learning, and information hiding.

LIMING YUAN received the Ph.D. degree in computer science and technology from the Harbin Institute of Technology, China, in 2014. He is currently a Lecturer with the School of Computer Science and Engineering, Tianjin University of Technology, China. His research interests mainly include machine learning and image processing.



CHANGLUN GUO received the B.S. degree in computer science and technology from the Shandong University of Technology, in 2018. He is currently pursuing the master's degree with the School of Computer Science and Engineering, Tianjin University of Technology. His research interests include image processing, neural networks, and machine learning.

HAIXIA XU received the M.S. degree in applied mathematics and the Ph.D. degree in computer science and technology from Northwestern Polytechnical University, China, in 2006 and 2009, respectively. She is currently an Associate Professor with the School of Computer Science and Engineering, Tianjin University of Technology, China. Her main research interests include image analysis, signal processing, and pattern recognition.