IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON INNOVATION AND APPLICATION OF INTELLIGENT PROCESSING OF DATA, INFORMATION AND KNOWLEDGE AS RESOURCES IN EDGE COMPUTING

Received March 11, 2020, accepted March 21, 2020, date of publication March 26, 2020, date of current version April 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2983478

A Robust Polarmetric SAR Terrain Classification Based on Sparse Deep Autoencoder Model Combined With Wavelet Kernel-Based Classifier

XIANGDONG CHEN[1] AND JIANGHONG DENG[2]
[1]School of Information Engineering, Huanghuai University, Zhumadian 463000, China
[2]School of Animation, Huanghuai University, Zhumadian 463000, China

Corresponding author: Xiangdong Chen (chenxiangdong@huanghuai.edu.cn)

This work was supported by the Henan Natural Science Foundation under Grant 162300410195.

ABSTRACT Since the existing terrain classification algorithm based on deep learning is not ideal for unbalanced PolSAR classification, a effective terrain classification algorithm based on wavelet kernel sparse deep coding network under unbalanced data set is proposed in this paper. The algorithm firstly adopts a structured sparse operation so as to enhance the accuracy of feature propagation and reduce the amount of stored data, where the unimportant parameter connections in each group are gradually reduced by dividing the network convolution kernel into multiple groups during the training process; The wavelet kernel-based classifier is used instead of the Sigmoid function to classify and identify features for different terrain, which has high generalization performance for small sample, nonlinear and high-dimensional mode classification problems. The experimental results show that our proposed classification algorithm can improve the classification performance of unbalanced samples, and improve the classification efficiency while ensuring the accuracy of classification.

INDEX TERMS Terrain classification, wavelet kernel, support vector machine, deep model, sparse coding, sigmoid function.

I. INTRODUCTION

Polarimetric synthetic aperture radar (PolSAR) has the advantages of all-weather, all-weather, high resolution, low degree of interference, and its unique phase data contains the information of target surface roughness, surface effect and position orientation, which is far superior to optical equipment, and widely used in geological exploration, regional planning, situation assessment and other fields [1]. In recent years, great progress has been made in polarimetric SAR image classification. However, due to the inherent problems of airborne/spaceborne radar and the influence of ground objects such as noise and radio-frequency interference in the process of data acquisition, it will lead to inaccurate image classification in the later stage, so polarimetric SAR image classification is still a challenging research direction [2].

Compared with the SAR single-polarization scalar imaging mechanism, polarimetric SAR data is characterized by a polarized scattering matrix, which has the advantage of being able to unify the scattering matrix data with the phase, energy, and polarization characteristics of the target's scattering. Polarization SAR image classification algorithms can be roughly divided into three categories: Bayesian classification methods based on statistical distribution models [3], classification methods based on polarization target parameter decomposition, and hybrid models of the two. The first type uses the statistical characteristics of the data for classification. The classic method is the Wishart classifier. The second type is the polarization characteristics of the data. The target's physical scattering characteristics are obtained by the method of target decomposition so as to achieve the purpose of classification. The last category is the organic combination of statistical characteristics and polarization characteristics, which is also the hottest direction of research.

Traditional terrain classification methods usually extract the radiation information, polarization information, subaperture decomposition information and other features manually according to the scattering features of the surface of the terrain object, and then input the single feature or multiple features into the appropriate classifier for classification. Because

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

single feature can not achieve satisfactory performance, even if these features are combined together, the accuracy of classification will not be improved significantly, and the computational complexity will increase. Therefore, the traditional methods can not make full use of the rich characteristics of polarimetric SAR data to improve the classification accuracy. In recent years, there are more and more researches on polarimetric SAR image classification based on machine learning algorithm. The polarimetric SAR image classification model including support vector machine classifier, random forest classifier and random subspace classifier is proposed. In addition, a robust, objective, repeatable, efficient and high accuracy polarimetric SAR image classification model combined with the characteristics of image spectrum, texture, shape and context information is established. However, the traditional shallow learning methods, such as machine learning and pattern recognition, are sensitive to noise and cannot effectively process and analyze large-scale polarimetric SAR terrain data.

Compared with the traditional low-level learning, the deep learning represented by convolutional neural network emphasizes the depth of the model structure and the importance of feature learning. Through layer by layer feature transformation, the feature representation of samples in the original space is transformed to a new deep feature space, so that the classification is easier. Compared with the manually constructing features, using big data to learn features can better represent the rich internal information of terrain data. Due to the limitation of samples and computing ability, shallow low-level learning lacks generalization ability to complex problems. Deep learning can realize the approximation of complex functions and represent the distributed representation of input data by learning a deep nonlinear network structure, and show a strong ability to learn the essential characteristics of data from the mass sample set. Many scholars at home and abroad have introduced deep learning frameworks to process rich polarimetric SAR feature information and have proposed many excellent classification models for different application backgrounds, such as deep restricted Boltzmann (DRB) [10], stacked autoencoder(SDA) [11] and deep convolution neural network (CNN) [12]. Wang *et al.* transformed polarimetric SAR image into scattering matrix, and then established a multichannel CNN classification model [13]. Tang *et al.* [14] transformed the complex matrix of polarimetric SAR image into the real matrix of multi-channel to adapt to the input of neural network, and designed a multi-cascade fully connected classifier classifier, which further improved the accuracy of polarimetric SAR image classification. According to the analysis of existing models, it can be found that polarimetric SAR image classification algorithm based on deep learning needs complex feature decomposition process [19]–[22].

Since most of the existing terrain classification algorithms introduce machine learning algorithm to realize the classification on the basis of polarimetric data decomposition. However, the classification results of these algorithms are too rough, which is mainly due to the fact that the existing feature extraction algorithms can not fully characterize the essential characteristics of PolSAR terrain and the classification algorithms can also not fully fit all the functions. Therefore, according to the stacked sparse autoencoder(SSA) model, this paper proposes a robust polarimetric SAR terrain classification algorithm, which uses the least square support vector machine based on Morlet wavelet kernel to replace the commonly used softmax classifier in the deep model. By combining with the SSA network, our proposed algorithmovercomes overcomes the shortcomings of traditional polarimetric SAR image classification method which is greatly affected by speckle noise and the result is too rough to a certain extent. It ensures the consistency of homogeneous region and in-homogeneous region in classification results. Qualitative and quantitative simulation results on two real polarimetric SAR image data show that our proposed method not only improves the classification accuracy, but also greatly shortens the time of terrain classification.

II. RELATED METHODS FOR OUR MODEL

A. AUTOENCODER NETWORK

A typical deep learning model usually refers to a multilayer perceptron network with multiple hidden layers, and the number of hidden layers is generally greater than 3, sometimes even more. The deeper the model (the greater the number of hidden layers), the more parameters such as the weight value W and threshold b in the corresponding neural network model, which directly reflects the stronger ability of the network model to automatically extract higher-level features. Generally, when designing a complex model, it will encounter low training efficiency and be easy to make the model overfit. However, with the rapid development of hardware technology, model parallelization, error back propagation (BP), fine-tuning and training techniques and other methods can greatly improve the efficiency of model training and reduce the risk of overfitting. In contrast, traditional machine learning algorithms are difficult to process the original data, and usually need to artificially extract features from the original data. This requires researchers to have a fairly professional understanding of the original data, and also to learn to build a strong classifier to combine the extracted features so as to achieve the desired classification effect. Deep learning can autonomously extract features at different levels. By effectively combining features at different levels, selecting a simple classifier can complete complex classification tasks and achieve better classification results.

With the development of deep learning theory, many deep network models have also been proposed. For example, relatively classic network models: Deep Belief Networks (DBN), Convolutional Neural Network (CNN), Stacked Auto Encoder (SAE), Recurrent Neural Network (RNN), and recently more popular Generative Adversarial Networks (GAN). Many algorithms are based on these models, and they have been widely used in various fields, especially on the classification of polarized SAR images, so that the accuracy of classification has been qualitatively improved.

Autoencoder network is a kind of artificial neural network which is used to learn efficient coding. By learning the essential characteristics of data set, the purpose of data deep coding can be achieved. The autoencoder network can transform the specific feature vector into the abstract one. The autoencoder network can well satisfy the two-way mapping nonlinear learning between high-dimensional data space and low-dimensional data space. It uses adaptive, multi-layer encoder network to transform high-dimensional original data into low-dimensional abstract data, and uses similar decoder network to reconstruct high-dimensional characteristics of original data from low-dimensional abstract data.

If a given set of samples data-set contains m samples, the overall cost function of the autocoder model can be defined as:

$$\mathrm{J}(W,b)=\frac{1}{m}\sum_{i=1}^{m}\frac{1}{2}\left\|h_{W,b}(x^i)-y^i\right\|^2+\frac{\lambda}{2}\sum_{l=1}^{\eta_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(W_{ji}^l)^2 \tag{1}$$

where m is the number of training samples; $h_{W,b}(x)$ is the corresponding output function; λ regularization parameter that adjusts the magnitude of the weight and bias changes.For the L-layer network, the traditional deep network has only L connections, while the improved sparse deep network has $L(L-1)/2$ connections. Although this can make full use of all the information of the previous layer, it will inevitably increase the complexity of the algorithm if the size of the convolution kernel is not limited. In order to reduce the dimension of the feature cascade between all layers, 1×1 convolutions are generally used to linearly combine on different channels to adjust the feature depth so as to complete the function of increasing or decreasing the network dimension [16]. Although the propagation and utilization of features are greatly improved, the network needs a lot of space to store intermediate variables. In order to reduce the redundant features of the network, this paper proposes an adaptive cropping strategy, where each layer of the unit structure does not receive the output features of all previous network layers, and its purpose is to retain the connections with important weights.

It is given a_i is the feature of the i th layer, where there are N unit structures, $N \leq i$. This shows there is K connections between a_i and the upper layer, and the connection weight is recorded as $b_{i1}, b_{i2}, \cdots, b_{iN}$. Thus, we can calculate the correlation coefficient between a_i and b_{in},

$$r_{ik}=\frac{E\left[a_i-u_{ai}\right]\left[b_i-u_{bik}\right]}{\sigma_{a_i}\sigma_{ik}} \tag{2}$$

where u and σ are the mean and variance obtained on the test data set. By sorting the weight vectors of positive correlation and random sampling the previous data of the first λ, we get $\lambda s N^+$ connections, where s is the corresponding sparsity. Random grouping assigns input channels to each group equally, which meets the precision requirements of sparse reconstruction. Therefore, by directly removing the connection weight less than a certain threshold value, the sparse network connection is finally obtained, reducing the complexity of the network. Because the basic level data of the improved deep network is cascaded by global pooling, the obtained parameters are sparse and distributed in a structured way [16]. In order to enhance the accuracy of feature propagation and reduce the amount of data stored, this paper adopts structured operation. By dividing the network convolution kernel into multiple groups, the connection of unimportant parameters in each group is gradually reduced in the training process. It can be seen that the pruning strategy adopted in this paper is obtained by self-learning of the network. The learned feature parameters can present a structured and sparse distribution, and the subsequent pruning of weights will not bring too much precision loss [23]–[26].

B. LEAST SQUARES SUPPORT VECTOR MACHINE

Least squares support vector machine (SVM) is an improved model of support vector machine. Compared with the traditional support vector machine, the least squares support vector machine can overcome the long training time, the randomness of training results, and the shortcomings of over-learning, so that the model classification accuracy meets the classification requirements, while reducing the time complexity and space complexity [30]–[36].

Given a linearly separable sample set: $\{(x_i, y_i)\,|x_i \in R^m\ , i = 1, 2, \cdots, n\}$, we can get the model for least squares support vector machine, which can be written as follows:

$$\begin{aligned}&\min\ (\frac{1}{2}\|\mathrm{w}\|^2+C\sum_{i=1}^{l}\varsigma_i^2)\\&s.t.\ y_i(\mathrm{w}\cdot\Phi(\mathrm{x}_i)+\mathrm{b})=1-\varsigma_i,\quad i=1,2,\cdots,l\end{aligned} \tag{3}$$

where, C is a regularization constant, which is used to balance differences between inter-class and inner-class. Since terrain classification is a linearly separable problem, it is necessary to map the training samples from a low-dimensional space to a high-dimensional feature space by introducing a kernel function. The problem to be solved is transformed into a linear classification problem in a high-dimensional feature space [27]–[29]. Therefore, it can be seen that the optimization problem of the least squares support vector machine can be obtained by solving linear equations, and its optimal classification surface is:

$$\min\left\{\frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_jK(x_i,y_i)+\sum_{i=1}^{n}\frac{\alpha_i^2}{2\gamma}-\sum_{i=1}^{n}\alpha_iy_i+b\sum_{i=1}^{n}\alpha_i\right\} \tag{4}$$

where the kernel function K() can select linear kernel function, polynomial kernel function and so on. The kernel function $\mathrm{K}(x_i, y_i)$ can be used to directly calculate the inner product $\mathrm{K}(x_i, y_i) = \varphi(x_i)\varphi(y_i)^{\mathrm{T}}$ in high dimensional space, but most kernel functions do not fit all the data. By adjusting the error term of the objective function in the SVM model into the quadratic, the least squares support vector machine can be transformed into solving a system of linear equations, which reduces the complexity of the problem and improves the

speed of solution, making it not only applicable to large-scale problems, but also to general problems. However, the solution obtained by the model can not be guaranteed to be the global optimal solution, and the solution lacks sparsity. It other words, it is easy to leads to over-fitting, and reduces the accuracy of terrain feature classification.

C. CHARACTERISTICS OF RAW POLARIMETRIC SAR DATA

As we all know, the observed data of polarimetric SAR is a polarimetric scattering complex matrix with the size of 2×2, which represents the scattering phenomenon of the terrain to the incident radar wave. Its electromagnetic scattering model is shown as follows:

$$\mathrm{E}^{sc} = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} E^{tr} = [S]\, E^{tr} \tag{5}$$

where h and v horizontal and vertical polarization directions, respectively; E^{sc} and E^{tr} correspond to the electric field vector received and transmitted by the polarization radar antenna. S_{hh}/S_{vv} is denoted as co-polarization component, S_{hv}/S_{vh} is cross-polarization component. In order to facilitate the subsequent description, Pauli orthogonal basis is used to vectorize the polarization scattering matrix, and the scattering vector can be written as follows:

$$\overline{\mathrm{K}} = \frac{1}{\sqrt{2}}(S_{hh} + S_{vv}, S_{hh} - S_{vv}, 2S_{hv})^{\mathrm{T}} \tag{6}$$

In order to suppress speckle noise and reduce storage space in polarized images for engineering applications, multi-look coherent processing of polarimetric data is usually needed. Therefore, the corresponding polarization coherence matrix can be obtained by coherent transformation as (7), as shown at the bottom of this page.

It can be seen that the polarization coherence matrix is a conjugate symmetric matrix; therefore, for the convenience of analysis, the three real values at the diagonal are respectively recorded as T_{11}, T_{22} and T_{33}; the three complex values in the upper right of [T] are recorded as T_{12}, T_{13} and T_{23}.

Since the input data of the autoencoder network model is the pixel matrix, the deep feature data is obtained by encoding the image spatial information. In order to realize the data description of terrain feature information in polarimetric SAR image, it is necessary to extract 9 elements of polarimetric coherence matrix T from each pixel of polarimetric SAR image and combine them into the original polarimetric SAR data features, which can be described as the following matrix form:

$$\begin{bmatrix} T_{11} & T_{22} & T_{33} \\ real(T_{12}) & imag(T_{12}) & real(T_{13}) \\ imag(T_{13}) & real(T_{23}) & imag(T_{23}) \end{bmatrix} \tag{8}$$

where $real()$, $imag()$ represents the real part and the imaginary part of the polarization value, respectively. Since there is correlation between neighboring pixels in polarimetric image and the feature spaces of different dimensions are overlapped, The $(3N)^2$-dimensional feature is obtained by N-dimensional expansion of the polarization matrix, which can describe the characteristics of polarization data more accurately.

III. OUR PROPOSED TERRAIN CLASSIFICATION ALGORITHM FOR POLARIMETRIC SAR

The traditional polarimetric SAR terrain classification algorithm mainly uses the polarimetric target feature decomposition method to obtain the classification features, and then uses the classifier to achieve the classification result. Their feature extraction model is complex in calculation, poor in generalization ability and rough in classification results. In recent years, with the rapid development of hardware and the use of complex mathematical tools, deep learning represented by convolutional neural network has made great progress. Deep learning model encodes the low-level features to form a more abstract high-level attribute categories or features so as to discover the more essential feature representation of data. In this paper, a polarimetric SAR terrain classification algorithm based on sparse stacked autoencoder neural network is proposed by using the excellent feature learning ability of deep learning network. More essential feature information is obtained by directly extracting the deep feature of the original polarimetric SAR covariance data. In addition, in order to improve the accuracy of terrain feature classification of polarimetric SAR image, this paper uses the least square support vector machine based on Morlet wavelet kernel to replace the commonly used softmax classifier in deep model. By combining with sparse stacked autoencoder network, our algorithm overcomes the shortcomings of traditional polarimetric SAR image classification method which is greatly affected by speckle noise and the result is too rough to a certain extent. It ensures the consistency of homogeneous region and inhomogeneous region in classification results.

A. SPARSE STACKED AUTOENCODER NETWORK

A stack autoencoder network is a multi-layer deep autoencoder neural network composed of multiple autoencoder cascaded together. The output of the auto-encoder in the previous layer is used as the input of the auto-encoder in the subsequent layer. Therefore, the encoding process of the stack autoencoder network is to perform the encoding of each layer of the auto-encoder in order from front to back. Due to the large number of self-similar features in polarized SAR images, the introduction of sparse constraints in stacked autoencoder networks can make tasks perform better in network learning.

$$[\mathrm{T}] = \left\langle \overline{\mathrm{K}}, \overline{\mathrm{K}}^{\mathrm{T}} \right\rangle = \frac{1}{2} \begin{vmatrix} |S_{hh} + S_{vv}|^2 & \langle (S_{hh} + S_{vv})(S_{hh} - S_{vv})^* \rangle & 2\langle (S_{hh} + S_{vv}) S_{hv}^* \rangle \\ \langle (S_{hh} - S_{vv})(S_{hh} + S_{vv})^* \rangle & |S_{hh} - S_{vv}|^2 & 2\langle (S_{hh} - S_{vv}) S_{hv}^* \rangle \\ 2\langle S_{hv}(S_{hh} + S_{vv})^* \rangle & 2\langle S_{hv}(S_{hh} - S_{vv})^* \rangle & 4|S_{hv}|^2 \end{vmatrix} \tag{7}$$

Given $a_i(x)$ as the activation degree of hidden layer i, we have $a_i(x) = \mathrm{s}(\sum_{n=0}^{N} w_{ni}x_n + b)$. This equation shows the activation degree of the i-th neuron corresponding to x, N is the input vector dimension, and w_{ni} is the connection weight between the neuron n and i. In order to increase the sparsity constraint and inhibit the activity of neuron, the average activity of neurons proposed in [17], [37] is introduced in this paper:

$$\overline{\rho}_i = \frac{1}{m}\sum_{i=1}^{m} a_i^{(k)} x^{(i)} \tag{9}$$

where ρ is the sparsity parameter. If $\overline{\rho}$ is approximately equal to ρ, we can get the sparsity of neurons i; $a_i^{(k)}$ represents the polarization value of layer K; $x^{(i)}$ is the input of the *i-th* sample. In order to limit the average activity of the stack autoencoder networks cascaded by multiple autoencoder, it is necessary to limit the neurons with large differences. The most common method is the relative entropy deviation method, whose expression is written as follows:

$$\sum_{j=1}^{M} KL(\rho \,\|\overline{\rho}) = \sum_{j=1}^{M} \rho \log(\rho/\overline{\rho}) + (1-\rho)\log\frac{1-\rho}{1-\overline{\rho}} \tag{10}$$

where M is the number of hidden neurons, which can measure the difference between the two distributions. Therefore, the total cost function can be expressedas follows:

$$\mathrm{J}_s(W, b) = J(W, b) + \beta(KL(\rho \,\|\overline{\rho})) \tag{11}$$

where β is a regularization parameter to control the sparsity penalty factor, and the appropriate parameters (W, b) can be obtained by the random gradient descent algorithm.

B. LEAST SQUARE SUPPORT VECTOR MACHINE BASED ON MORLET WAVELET KERNEL

Support vector machine transforms low-dimensional data into high-dimensional space by kernel mapping. Its purpose is to find an optimal hyperplane for linear classification. However, the approximation of SVM to any function in high-dimensional space is not always so accurate, which makes the actual classification results too rough. The main reason is that the existing kernel function can not generate a set of complete orthogonal functions, so it is necessary to introduce a basic function with complete space transformation.

Since the kernel function of traditional SVM can not fit all the data, it directly affects the classification effect. In recent years, scholars at home and abroad are studying the improved support vector machine algorithm. LV *et al.* constructed the wavelet kernel function by using the wavelet analysis theory, which can achieve a good approximation to the image and better reconstruct the original image [15], [38]. Therefore, on the basis of least square support vector machine, wavelet kernel function is introduced in this paper. Sparse kernel function is helpful to improve the classification accuracy of the model and the convergence speed of iteration. According to the advantages of wavelet kernel function, Morlet wavelet function is used as the kernel of SVM.

It can be seen from literature [15] that for any multidimensional wavelet, it can be separated into one-dimensional product form by tensor product theory, and the expression form of wavelet kernel function is as follows:

$$\mathrm{K}(\mathrm{x}_i, \mathrm{y}_i) = \prod_{i=1}^{N} h(\frac{x_i - y_i}{a}) \tag{12}$$

where $h(x)$ is wavelet mother function, a is the corresponding scale factor. In this paper, we choose the wavelet function Morlet which is widely used and can satisfy the condition of support vector kernel function. The wavelet function can be denoted as:

$$h(x) = \cos(1.75x)\exp(-\frac{x^2}{2}) \tag{13}$$

Therefore, the wavelet kernel function based on Morlet can be written as follows

$$\begin{aligned}\mathrm{K}(\mathrm{x}_i, \mathrm{y}_i) &= \prod_{i=1}^{N} h(\frac{x_i - y_i}{a}) \\ &= \prod_{i=1}^{N} (\cos(1.75(\frac{x_i - y_i}{a})\exp(-\frac{\|x_i - y_i\|^2}{2a^2}) \end{aligned} \tag{14}$$

Then the expression of the classification hyperplane of the wavelet kernel LSSVM is rewritten as follows:

$$\min\left\{\frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j\prod_{i=1}^{N}h(\frac{x_i - y_i}{a}) + \sum_{i=1}^{n}\frac{\alpha_i^2}{2\gamma} - \sum_{i=1}^{n}\alpha_i y_i + b\sum_{i=1}^{n}\alpha_i\right\} \tag{15}$$

C. MODEL FRAMEWORK FOR TERRAIN CLASSIFICATION

In this paper, our proposed terrain classification takes advantage of the excellent feature learning ability of stacked sparse autoencoder model. By learning a deep nonlinear network structure, it can approximate complex functions and represent the most essential feature information of input data. The complex steps of obtaining the classification features through the polarization data feature decomposition are avoided; and the Morlet wavelet kernel-based least squares support vector machine is used to replace the traditional Softmax classifier to achieve accurate classification of terrain information. The main implementation steps of our proposed algorithm are: firstly, preprocess the original image information to reduce the influence of speckle noise on feature extraction and classification in post-processing; and then, the stacked sparse autoencoder model is adopted to process the original polarimetric SAR coherent vector data so as to extract the sparse deep features of training samples and get the most essential feature information; finally, the least square support vector machine classifier based on Morlet wavelet kernel is used to normalize and classify the sparse depth features so as to obtain the classification results. The brief flow of the proposed algorithm in this paper is shown as follows: preprocess the original image information, reduce the coherent

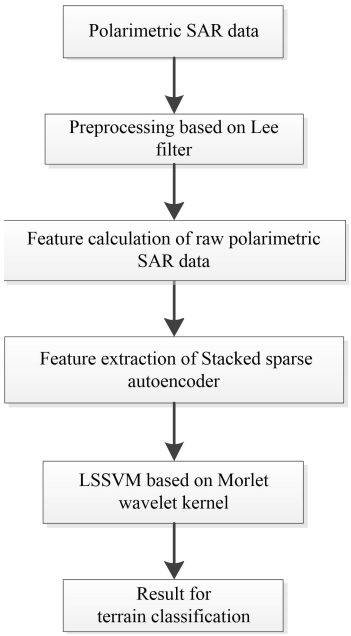


FIGURE 1. Model framework for terrain classification.

speckle noise to extract and classify the image for post-processing features; then, the stack sparse self-coding learning and extraction on the original polarized SAR coherent vector data The sparse depth features of the training samples are used to obtain the most essential feature information. Finally, a Morlet wavelet kernel least squares support vector machine classifier is used. The brief flow of the algorithm proposed in this article is shown in the following figure:

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In order to facilitate the analysis of the effectiveness of the terrain classification algorithm, this paper uses the polarimetric SAR data set acquired by China's independently developed high-resolution Gaofen-3 satellite for training, and also use the general international benchmark polarimetric SAR data for qualitative and quantitative analysis. The polarimetric SAR image classification method based on Wishart and SVM (WSVM) [3], the terrain classification algorithm based on autoencoder network (AE) [12], the terrain classification algorithm based on Deeplab-V1 [18], and terrain classification algorithm based on stack based on experiments Feature (SAE) of the autoencoder network [17]. Deeplab-V1 is a pixel-based multi-objective classification network that uses the fully connected conditional random fields to fine-tune the classification results in detail. The improved deep feature model proposed in this paper uses the wavelet kernel based LSSVM instead of the Sigmoid function. For the sake of analysis, the improved algorithm using the wavelet-kernel-based LSSVM is denoted as SDL_SVM, while the model using the traditional Sigmoid function is denoted as SDL_Sig.

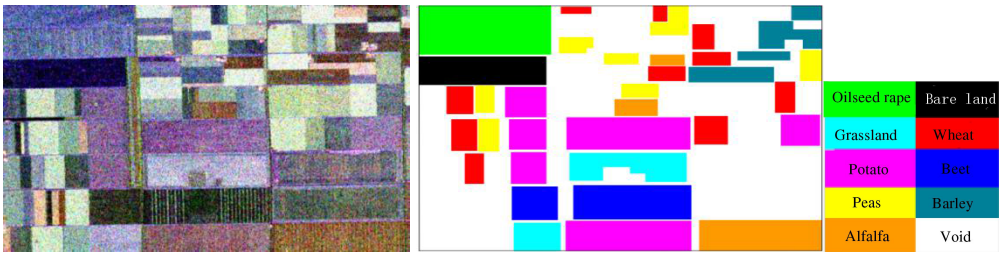
In order to avoid the difference in results caused by different training samples, the number of samples selected for all experiments is the same, where 5% of the samples are used as training samples, and the remaining samples are used as test samples. Different terrain in SAR image are displayed in different colors. The classification results of all experiments are the average of 20 experimental results. Because the test images for quantitative analysis have benchmark results, this paper uses the percentage of classification accuracy in the benchmark sample as the final terrain classification accuracy. In the experiment, Tensorflow deep learning framework was selected.

As we all know, the classification results are greatly affected by speckle noise. In order to enhance the accuracy of classification and reduce the impact of noise, the same preprocessing method is used in all experiments. Firstly, the classified polarimetric SAR data is filtered by Lee with a window size of 7×7 to remove the speckle noise. Then, the polarimetric coherence matrix of the polarimetric SAR data after Lee filtering is obtained and it is decomposed by Cloude. The scattering entropy **H**, scattering angle **alpha** and total scattering power **span** of each pixel point are calculated, and the feature set $\mathbf{F} = \mathbf{[H\ alpha\ span]}$ and the feature set **T** of the coherence matrix are constructed. In addition, the similarity matrix W^F and W^T of feature set F and T are calculated, and the subsequent experiments will be carried out on the basis of data decomposition.

A. QUALITATIVE AND QUANTITATIVE ANALYSIS OF FLEVOLAND POLARIMETRIC SAR IMAGE

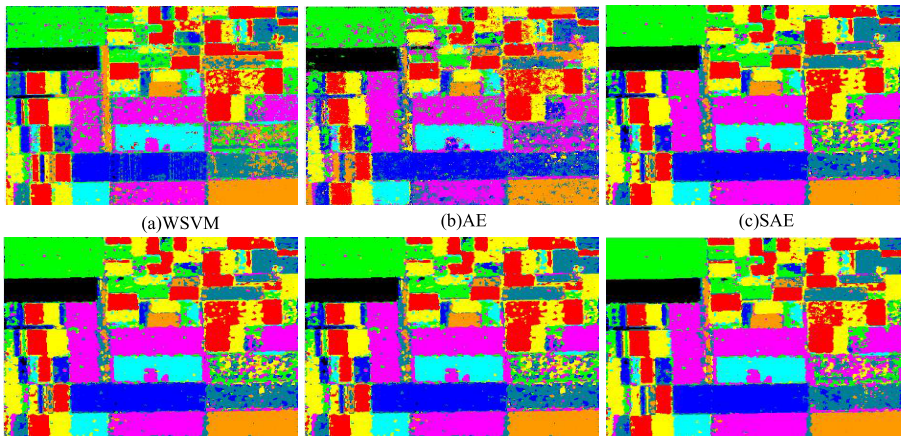
The full polarization data of Flevoland region in the Netherlands is the L-band image obtained by the The NASA/JET Propulsion Laboratory Airborne Synthetic Aperture Radar system. In the experiment, the sub regional polarization SAR image with the size of 380×420 is selected for processing. The image includes 9 kinds of terrain with the resolution of 12.1m $\times$ 6.7m. The Pauli pseudo-color and its benchmark data are shown in **FIGURE 2**. It can be seen that there are a lot of speckle noise in the original image, which seriously interferes with the accuracy of feature extraction and classifier training in the later stage.

The AE terrain classification algorithm is the earliest proposed to transfer the maximum pooling index to the decoder module, generate a terrain probability map through a convolution layer and some skip-connections, and then gradually terrain it to eventually improve the precision of feature classification. However, AE directly uses the Softmax loss function judges the results, the result is relatively rough and the spatial consistency is poor both in the processing of the boundary and the processing of small objects. From the qualitative classification results in **FIGURE 3**(b), it can be seen that there exists at the boundary mis-classification and incomplete classification results. The structure of SAE is very similar to that of AE, but the network uses a fully connected layer as a relay between the encoder and decoder, and uses multiple network modules in a stack connection at the same time; as shown in **FIGURE 3**(c), the classification results have been improved, especially it is the same terrain with good



(a)Pauli pseudo-color (b)Benchmark mark

FIGURE 2. Flevoland.



(a)WSVM (b)AE (c)SAE

(d)Deeplab v1 (d)SDL_Sig (e)SDL_SVM

FIGURE 3. Classification results of *Flevoland* for different comparison algorithms.

TABLE 1. Comparison of average classification accuracy of 20 experiments of flevoland PolSARimage(%).

Category	WSVM	AE	SAE	Deeplab v1	SDL_Sig	SDL_SVM
Grassland	74.16	80.01	78.00	83.50	87.25	**90.63**
alfalfa	57.24	65.34	69.00	73.21	79.22	**81.08**
Beet	71.35	82.39	88.47	86.60	**90.10**	**90.10**
Oilseed rape	71.04	74.27	76.60	82.78	88.96	**90.63**
Bare land	67.49	68.35	72.47	80.01	82.95	**82.94**
Potato	70.11	73.14	74.29	77.29	78.98	**80.36**
Peas	76.83	82.13	86.64	84.48	**91.44**	91.40
Wheat	62.50	64.37	68.56	75.83	79.55	**79.56**
Barley	54.58	65.20	78.39	81.33	86.12	**86.17**
Mean	66.14	72.81	76.05	77.34	*84.68*	**86.07**

consistency and relatively smooth vision; the noise of the Deeplab v1 result in **FIGURE 3**(d) is significantly reduced; SDL_SVM and SDL_Sig are the two algorithms proposed in this paper. From the qualitative results of **FIGURE 3**(e) and 3(f), it can be seen that the internal terrain of SDL_SVM are better than SDL_Sig. From the quantitative results in **TABLE 2**, it can also be seen that the classification accuracy of SDL_SVM is 1.5% higher than that of SDL_Sig, which is mainly attributed to the ability of the wavelet kernel to fit complex ground structures.

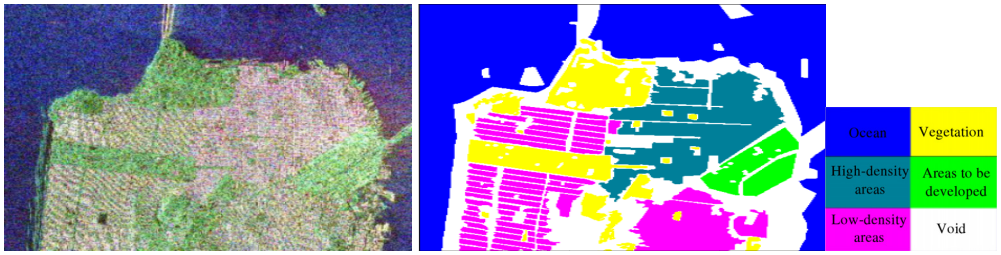
B. QUALITATIVE AND QUANTITATIVE ANALYSIS OF SAN FRANCISCO BAY POLARIMETRIC SAR IMAGE

FIGURE 3 is the L-band polarization SAR data of the San Francisco Bay Area in the United States. The image size is 900 × 1024 and the number of look is 4. The data contains relatively rich terrain, mainly sea surface, vegetation and urban, of which urban areas are also divided into high-density areas, low-density areas and areas to be developed.

FIGURE 5 shows the terrain classification results of different models. **FIGURE 5**(a) is the result of the traditional

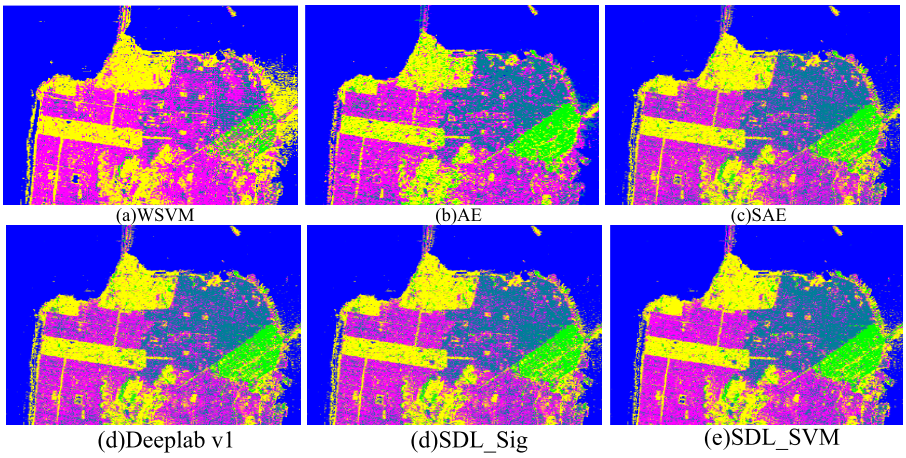
TABLE 2. Comparison of average classification accuracy in San francisco SAR image(%).

Category	WSVM	AE	SAE	Deeplab v1	SDL_Sig	SDL_SVM
Ocean	86.58%	86.70%	88.59%	89.55%	96.97%	**97.01%**
Vegetation	77.50%	78.57%	84.56%	85.85%	89.56%	**90.32%**
High-density areas	76.85%	77.95%	81.66%	86.68%	91.60%	**92.14%**
Low-density areas	76.96%	78.09%	78.71%	85.50%	90.65%	**90.85%**
Areas to be developed	67.76%	75.56%	79.00%	85.79%	**89.08%**	88.97%
Mean	77.13%	79.37%	82.50%	*86.67%*	**91.57%**	90.83%



(a)Pauli pseudo-color (b)Benchmark mark

FIGURE 4. San francisco.



(a)WSVM (b)AE (c)SAE

(d)Deeplab v1 (d)SDL_Sig (e)SDL_SVM

FIGURE 5. Classification results of San francisco for different comparison algorithms.

method using Wishart for distribution fitting and SVM for feature classification. The rest of the comparison algorithms are based on the improvement of the deep model. It can be seen from the results that these deep learning models have achieved very good terrain classification effects, and the qualitative and quantitative indicators are better than traditional low-level feature terrain classification algorithms, but there are also their own problems. Basically all algorithms have sawtooth phenomenon at the boundary. In the result of **FIGURE 5**(b), AE has a large number of misclassifications; SAE fails to accurately classify the vegetation in the middle of the high-density area; Deeplab v1 does not segment the area to be developed in the ocean, and the contour is not accurate enough. It can be seen from the experimental results that the SDL_SVM and SDL_Sig models proposed in this paper can learn the hidden structural features in the image so as to improve the performance of the terrain classification model. **TABLE 2** shows the average of the overall accuracy (Overal Accuary, OA), average accuracy (AA), Kappa coefficient, and running time (Times) of the classification results in 20 experiments of different methods. From the data results in **TABLE 2** and **3**, it can be seen that the overall index of the proposed algorithm in this paper is higher, especially SDL_SVM is more accurate than SDL_Sig. This is because the ocean category is basically not misclassified, and other categories are basically not misclassified into the ocean category. This indicates that the difference between the ocean and other terrain features is large, and the correlation between the

TABLE 3. Performance comparison of different models for San Francisco terrain classification.

Index	WSVM	AE	SAE	Deeplab v1	SDL_Sig	SDL_SVM
OA(%)	85.85	85.86	86.27	89.90	92.20	**96.76**
AA(%)	72.34	75.49	76.85	77.21	84.87	**95.85**
KAPPA	0.8051	0.8245	0.8298	0.8356	0.8831	**0.9417**
Time(s)	15.12	12.11	24.86	19.10	**7.12**	8.33

categories is small; in other words, the classification effect using the Morlet wavelet kernel LSSVM has been significantly improved, which is at least 6 percentage higher than the direct classification based on Sigmoid. The classification effect on urban areas, high-density urban areas, and low-density urban areas has been improved, which fully illustrates the effectiveness of the wavelet kernel LSSVM.

The comparison algorithm used in this paper is based on the source code or executable file provided by the author, and its parameters and initialization values are default values provided by the original literature. Because of the difference between the programming language and the programming style of the comparison algorithm, it is difficult to evaluate the operation time of our proposed algorithm in this paper. Therefore, this paper only analyzes the same hardware environment, and its average classification time for the same image is shown in **TABLE 2**. It can be seen that although our proposed terrain classification algorithm adds wavelet kernel least square support vector machine for feature classification, but because the least square SVM has linear solution performance, this does not increase the time complexity of the algorithm. The experimental results also show that the single frame computing time of our proposed model is the lowest. If multi-channel GPU parallel programming can be adopted, the algorithm in this paper will achieve real-time classification effect.

V. CONCLUSION

The existing terrain classification algorithms based on SVM are too rough, mainly because support vector machines cannot fit all functions accurately, and traditional feature classification algorithms have high computational complexity, long training time, and low efficiency. Based on the stack sparse autoencoder model, this paper proposes a robust polarimetric SAR terrain classification algorithm. Through experimental verification, the following conclusions are obtained:

(1) When the Morlet wavelet kernel least squares support vector machine is used as a classifier, the classification boundaries of each type of terrain are relatively clear, and the internal differences are small. Softmax has more stray points in some complex areas, and there are serious misclassifications.

(2) By combining with the stack sparse autoencoder network, our proposed algorithm overcomes the shortcomings of traditional polarimetric SAR image classification method which is greatly affected by speckle noise and the result is too rough to a certain extent. It ensures the consistency of homogeneous region and inhomogeneous region in classification results.

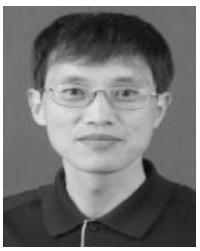
(3) Compared with other low-level feature learning or deep feature learning, the multi-feature cascade coding module used in this paper can extract higher-level and more essential features, and the efficient and improved Morlet wavelet kernel least squares SVM classifier is adopted to improve classification accuracy.

The polarimetric SAR terrain classification algorithm proposed in this paper can improve the accuracy of classification and is suitable for engineering applications. However, due to the limitation of equipment conditions, polarimetric SAR image usually contains more speckle noise. Whether traditional machine learning method or deep learning method is used, these noises will affect the learning model to learn enough accurate features for classification. In addition to relying on the development of radar system, how to effectively remove the noise or offset the impact of these noises from the aspect of algorithm are also worthy of in-depth study in next steps.

REFERENCES

[1] W. Song, M. Li, P. Zhang, Y. Wu, L. Jia, and G. Liu, "A classification method of PolSAR image based on weighted composite kernel and triplet Markov field," *Chin. J. Electron.*, vol. 44, no. 3, pp. 520–526, Mar. 2016.
[2] J. Zhang, B. Zhu, P. Zhang, M. Wang, and T. Liu, "Polarimetric SAR imagery target CFAR detection analytical algorithm with Wishart distribution," *Acta Electronica Sinica*, vol. 46, no. 2, pp. 433–439, 2018.
[3] L. Zhang, B. Duan, and B. Zou, "Research development on target decomposition method of polarimetric SAR image," *JEIT*, vol. 38, no. 12, pp. 3289–3297, 2016.
[4] B. Chen, S. Wang, L. Jiao, F. Li, S. Mao, and S. Zhang, "Polarimetric SAR image classification via weighted ensemble based on 0-1 matrix decomposition," *JEIT*, vol. 37, no. 6, pp. 1495–1501, 2015.
[5] X. Li, "Polarimetric SAR image classification based on Wishart classifier and spectral clustering," *J. Shanxi Normal Univ. (Natural Sci. Ed.)*, vol. 29, no. 1, pp. 35–39, 2015.
[6] C. Li, W. Wang, and P. Wang, "Current situatioand development trends of spaceborne SAR technology," *J. Electron. Inf. Technol.*, vol. 38, no. 1, pp. 229–240, 2016.
[7] W.-M. Boerner, "The development of multi-band equatorial orbiting POLSAR satellite sensors," in *Proc. IEEE Int. Conf. Aerosp. Electron. Remote Sens. Technol.*, Yogyakarta, Indonesia, Nov. 2014, pp. 127–131.
[8] J. D. Ballester and M. Lopezj, "Applying the Freeman–Durden decomposition concept to polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 466–479, 2010.
[9] N. P. Minh, B. Zou, and H. Cai, "Forest height estimation from mountain forest areas using general model-based decomposition for PolInSAR image," *J. Appl. Remote Sens.*, vol. 8, no. 1, 2014.
[10] J. Han, D. Zhang, G. Cheng, L. Guo, and J. Ren, "Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3325–3337, Jun. 2015.

[11] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, pp. 3371–3408, Dec. 2010.
[12] M. Sun, D. Zhang, J. Ren, Z. Wang, and J. S. Jin, "Brushstroke based sparse hybrid convolutional neural networks for author classification of chinese ink-wash paintings," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Montreal, QC, Canada, Sep. 2015, pp. 626–630.
[13] Y. Y. Wang, G. H. Wang, and Y. H. Lan, "PolSAR image classification based on deep convolutional neural network," *Metall. Mining Ind.*, vol. 8, pp. 366–371, Aug. 2015.
[14] Z. Tang, S. Wang, J. Huo, H. Guo, H. Zhao, and Y. Mei, "Bayesian framework with non-local and low-rank constraint for image reconstruction," *J. Phys., Conf. Ser.*, vol. 787, Jan. 2017, Art. no. 012008.
[15] W. Lv, Y. Chen, and S. Zhang, "Image denoising filtering algorithm based on wavelet support vector machine," *Remote Sens. Inf.*, vol. 3, no. 15, pp. 673–680, 2008.
[16] H. Song, W. Yang, Y. Bai, and X. Xu, "Unsupervised classification of polarimetric SAR imagery using large-scale spectral clustering with spatial constraints," *Int. J. Remote Sens.*, vol. 36, no. 11, pp. 2816–2830, Jun. 2015.
[17] L. Jiao, L. Bo, and L. Wang, "Fast sparse approximation for least squares support vector machine," *IEEE Trans. Neural Netw.*, vol. 18, no. 3, pp. 685–697, May 2007.
[18] H. Gao, W. Huang, and X. Yang, "Applying probabilistic model checking to path planning in an intelligent transportation system using mobility trajectories and their statistical data," *Intell. Automat. Soft Comput.*, vol. 25, no. 3, pp. 547–559, 2019.
[19] H. Gao, W. Huang, Y. Duan, X. Yang, and Q. Zou, "Research on cost-driven services composition in an uncertain environment," *J. Internet Technol.*, vol. 20, no. 3, pp. 755–769, 2019.
[20] H. Gao, Y. Duan, L. Shao, and X. Sun, "Transformation-based processing of typed resources for multimedia sources in the IoT environment," *Wireless Netw.*, vol. S1, no. 12, pp. 1–17, 2019.
[21] H. Gao, Y. Xu, Y. Yin, W. Zhang, R. Li, and X. Wang, "Context-aware QoS prediction with neural collaborative filtering for Internet-of-Things services," *IEEE Internet Things J.*, no. S1, pp. 259–267, Dec. 2019.
[22] X. Ma, H. Gao, H. Xu, and M. Bian, "An IoT-based task scheduling optimization scheme considering the deadline and cost-aware scientific workflow for cloud computing," *EURASIP J. Wireless Commun. Netw.*, vol. 249, no. 1, pp. 1241–1260, Dec. 2019.
[23] J. Yu, J. Li, Z. Yu, and Q. Huang, "Multimodal transformer with multi-view visual representation for image captioning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 37, no. 9, pp. 1–12, Oct. 2019.
[24] J. Yu, M. Tan, H. Zhang, D. Tao, and Y. Rui, "Hierarchical deep click feature prediction for fine-grained image recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 9, pp. 69–77, Jul. 2019.
[25] J. Yu, C. Zhu, J. Zhang, Q. Huang, and D. Tao, "Spatial pyramid-enhanced NetVLAD with weighted triplet loss for place recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 2, pp. 661–674, Feb. 2020, doi: 10.1109/TNNLS.2019.2908982.
[26] Y. Yin, J. Xia, Y. Li, Y. Xu, W. Xu, and L. Yu, "Group-wise itinerary planning in temporary mobile social network," *IEEE Access*, vol. 7, pp. 83682–83693, 2019.
[27] Y. Yin, L. Chen, Y. Xu, J. Wan, H. Zhang, and Z. Mai, "QoS prediction for service recommendation with deep feature learning in edge computing environment," *Mobile Netw. Appl.*, vol. 12, no.8, pp. 1–11, Apr. 2019.
[28] Y. Chen, S. Deng, and H. Ma, "Deploying data-intensive applications with multiple services components on edge," *Mobile Netw. Appl.*, vol. 12, no. 2, pp. 1–16, Apr. 2019.
[29] C Zhang, H Zhao, and S Deng, "A density-based offloading strategy for IoT devices in edge computing systems," *IEEE Access*, vol. 6, pp. 73520–73530, 2018.
[30] L. Kuang, X. Yan, X. Tan, S. Li, and X. Yang, "Predicting taxi demand based on 3D convolutional neural network and multi-task learning," *Remote Sens.*, vol. 11, no. 11, p. 1265, 2019.
[31] K. Xia, H. Yin, P. Qian, Y. Jiang, and S. Wang, "Liver semantic segmentation algorithm based on improved deep adversarial networks in combination of weighted loss function on abdominal CT images," *IEEE Access*, vol. 7, pp. 96349–96358, 2019.
[32] Y. Jiang, K. Zhao, K. Xia, J. Xue, L. Zhou, Y. Ding, and P. Qian, "A novel distributed multitask fuzzy clustering algorithm for automatic MR brain image segmentation," *J. Med. Syst.*, vol. 43, no. 5, pp. 118:1–118:9, May 2019.
[33] P. Qian, J. Zhou, Y. Jiang, F. Liang, K. Zhao, S. Wang, K.-H. Su, and R. F. Muzic, "Multi-view maximum entropy clustering by jointly leveraging inter-view collaborations and intra-view-weighted attributes," *IEEE Access*, vol. 6, pp. 28594–28610, 2018.
[34] P. Qian, C. Xi, M. Xu, Y. Jiang, K.-H. Su, S. Wang, and R. F. Muzic, "SSC-EKE: Semi-supervised classification with extensive knowledge exploitation," *Inf. Sci.*, vol. 422, pp. 51–76, Jan. 2018.
[35] Y. Jiang, Z. Deng, F.-L. Chung, G. Wang, P. Qian, K.-S. Choi, and S. Wang, "Recognition of epileptic EEG signals using a novel multiview TSK fuzzy system," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 1, pp. 3–20, Feb. 2017.
[36] K.-J. Xia, H.-S. Yin, and Y.-D. Zhang, "Deep semantic segmentation of kidney and space-occupying lesion area based on SCNN and ResNet models combined with SIFT-flow algorithm," *J. Med. Syst.*, vol. 43, no. 1, pp. 2:1–2:12, Jan. 2019.
[37] K. Xia, X. Zhong, L. Zhang, and J. Wang, "Optimization of diagnosis and treatment of chronic diseases based on association analysis under the background of regional integration," *J. Med. Syst.*, vol. 43, no. 3, pp. 46:1–46:8, Mar. 2019.
[38] Y. Jiang, D. Wu, Z. Deng, P. Qian, J. Wang, G. Wang, F.-L. Chung, K.-S. Choi, and S. Wang, "Seizure classification from EEG signals using transfer learning, semi-supervised learning and TSK fuzzy system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 12, pp. 2270–2284, Dec. 2017.
[39] P. Qian, Y. Chen, J.-W. Kuo, Y.-D. Zhang, Y. Jiang, K. Zhao, R. A. Helo, H. Friel, A. Baydoun, F. Zhou, J. U. Heo, N. Avril, K. Herrmann, R. Ellis, B. Traughber, R. S. Jones, S. Wang, K.-H. Su, and R. F. Muzic, "MDixon-based synthetic CT generation for PET attenuation correction on abdomen and pelvis jointly using transfer fuzzy clustering and active learning-based classification," *IEEE Trans. Med. Imag.*, early access, Aug. 16, 2019, doi: 10.1109/TMI.2019.2935916.
[40] C. Liu, J. Yin, and J. Yang, "Classification of multi-frequency polarimetric SAR images based on multi-linear subspace learning of tensor objects," *Remote Sens.*, vol. 7, no. 7, pp. 9253–9268, 2015.
[41] L. Qi, R. Wang, C. Hu, S. Li, Q. He, and X. Xu, "Time-aware distributed service recommendation with privacy-preservation," *Inf. Sci.*, vol. 480, pp. 354–364, Apr. 2019.



XIANGDONG CHEN received the degree from the Department of Space Physics, Wuhan University, in 1991. He works with the School of Information Engineering, Huanghuai University. His research interest includes signal and image processing.

JIANGHONG DENG received the degree from the Department of Physics, Henan Normal University, in 2001. She works with the School of Animation, Huanghuai University. Her research interests include computer aided education and processing of image and graphics.