ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

An Innovative PolSAR Image Classification Method Based on Non-Negative Constraints Stacked Sparse Autoencoder Network with Multi-Features Joint Representation Learning

Chen Ruixia[1], Zhu Shisong[1], Wang Zhensong[1], Wang Jianlong[1,*]

[1]*School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, 454003, China*
[*]*Corresponding author*

Abstract: *This paper proposed a framework based on joint multi-feature representation learning to reduce the inherent speckle phenomenon in Polarimetric Synthetic Aperture Radar (PolSAR) images interfere with the scattering characteristics of land objects. Firstly, the corresponding 6-dimensional real vector is obtained from the covariance matrix of PolSAR data and combined with the polarized feature vector obtained by the polarization decomposition method to improve the differentiation ability of similar features in images. Secondly, the stacked sparse autoencoder (SSAE) is employed, where the non-negative constraint method is incorporated to make the sparse features in the depth space robust by filtering the weights. Finally, a non-negative constrained SSAE model is proposed to effectively accomplish the classification task of PolSAR images. In the experiments, the proposed classification network is trained layer by layer using unlabeled data, the softmax classifier is trained with a small number of labeled pixels. The parameters obtained from the above steps are used as initial parameters to train the entire classification framework with labeled pixels, the resulting well-trained model is used to predict the labels corresponding to pixels in the datasets. Through experiments using the Flevoland and San Francisco Bay datasets, the results demonstrate that the proposed method yields superior classification results compared with traditional SVM, AE, and Gray Level Co-generation Matrix (GLCM) classification methods.*

Keywords: *Autoencoder (AE), polarized synthetic aperture radar (PolSAR), polarization decomposition, classification*

1. Introduction

Synthetic Aperture Radar (SAR) is a high-resolution active sensor that possesses the unique capability of electromagnetic wave imaging capability to work uninterrupted all-weather and all-day [1]. Due to the different polarization modes of the SAR sensors for transmitting and receiving echoes, these signals can connect the incident and scattered electric fields to obtain a backward scattering matrix reflecting the ground feature, which forms a PolSAR image [2]. PolSAR images have higher spatial resolution and better discrimination capabilities for land objects, which can reflect the features such as dielectric properties and geometric structures of real land objects. Therefore, PolSAR image interpretation is particularly important, and PolSAR image classification [3] is an important applied research content, whose primary objective is to classify all pixels in an image into a certain class according to specific rules.

At present, the feature extraction methods for PolSAR images can be broadly classified into two main categories: simple data transformations and polarimetric target decomposition. The methods based on simple data transformations involve the polarized scattering matrix and its corresponding vector forms [4-7], or mathematical transformations of the scattering matrix such as the polarimetric coherency matrix and polarimetric covariance matrix [8-11]. The method based on polarization target decomposition is to decompose the polarization coherence matrix or covariance matrix according to different physical mechanisms, and then use the decomposition parameters representing the target scattering or geometric structure to construct features. For example, entropy/alpha decomposition [12], Freeman decomposition [13], and Krogager decomposition [14]. Among them, the entropy/alpha decomposition proposed by Cloude and Pottier is one of the most widely used methods. This method decomposes the coherence matrix into three components with different scattering mechanisms, namely

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

internal reflection, external reflection, and polarized scattering components. Then, the pixels in the PolSAR image are classified according to the parameters derived from the H/α polarization decomposition. Although this method can only classify 8 types of ground objects, it is still widely welcomed due to its simplicity and clear physical meaning [15].

Considering the complexity of electromagnetic scattering from land objects and the interference caused by various factors such as imaging parameters and noise, it is crucial to select the optimal feature combination from a rich set of features to achieve effective land object classification. To adequately represent the complementary information between different land objects in PolSAR data, it is necessary to combine the real-valued vector with the polarized feature vector. Since different decomposition algorithms may lead to the weakening of some information features, this combination method can effectively exploit discriminative features of polarimetric SAR images, enhance the discrimination ability for similar land cover categories, and thus improve the robustness of classification results. Therefore, it is important to employ the approach for efficient and accurate classification of PolSAR images. To extract features from PolSAR images of data of interest, a great deal of manpower and experience is usually required. Therefore, to solve the specific problem of polarimetric SAR object classification, we need a novel and powerful strategy that can automatically learn features that will alleviate the dependence on human resources and provide higher accuracy.

In recent years, several studies have shown the effectiveness of deep learning for PolSAR image classification [16]. Some of the typical deep learning methods include Deep Belief Network (DBN) [17], Autoencoder (AE) [18], Convolutional Neural Network (CNN) [19], etc. As deep learning technology advances, studies in neuroscience have revealed the advantages of multilayer deep learning, which can automatically extract more abstract and invariant features from data. These automatically extracted deep features have stronger capabilities in representing images compared to manually extracted shallow features. The stacked sparse autoencoder (SSAE) is a classical deep learning model that has been successfully applied in the field of image classification. The model uses an unsupervised manner to automatically acquire a representation with coefficient abstraction features from the input data in a hierarchical manner by the process of layer-by-layer reconstruction [20, 21]. These deep features without human intervention can effectively distinguish between different categories. At the same time, it is also able to utilize data contextual information, thereby enhancing the accuracy of image classification. However, there are some challenges and problems with stacked autoencoders. First, due to the complex structure of the stacked autoencoder, it requires a lot of time and computational resources in the training process. Moreover, debugging the model can be difficult, requiring careful optimization and adjustment of parameters for each autoencoder within the stack.

According to the above problems, many researchers have proposed different thoughts and solutions. Zhou et al. [22] proposed an architecture containing a four-layer convolutional neural network through a reasonable analysis of PolSAR data and successfully proved that deep neural networks have great potential in feature extraction and classification of PolSAR data. In 2018, to allow deep convolutional neural networks to have better generalization performance in limited training samples, researchers proposed a polarization feature-driven PolSAR image classification method [23]. Geng et al. [24] proposed a deep convolutional autoencoder (DCAE) with scale transformation before and after the sparse autoencoder (SAE). This method can effectively aggregate local image information, thus solving the problem of low classification progress caused by speckle noises to a large extent. Zhang et al. [25] proposed local spatial information to SSAE (SSAE-ls) to overcome speckle noises. They effectively alleviated the influence of neighbor pixels on the central pixel by controlling the spatial distances between the central pixel and adjacent pixels. Hou et al. [26] adopted the k-nearest neighbor algorithm based on superpixels and combined with SSAE to obtain the probability distributions of classes to reduce the influence of speckle noise. Wang et al. [27] defined the data error term in the loss function according to the data distribution and introduced a stacked MAE (SMAE) network, which successfully dealt with highly polluted speckle noise. Zhang et al. [28] proposed a method called multi-scale sparse autoencoder (MS-SSAE), which can simultaneously capture discriminative multi-scale features and preserve details in images. Hu et al. proposed an algorithm called Adaptive Nonlocal Stacked Sparse Autoencoder (ANSSAE). The algorithm realizes the adaptive extraction of non-local spatial information by calculating the weighted average value of each pixel in the non-local area. It can not only reduce the influence of speckle noise but also preserve edge details [29]. These studies collectively validate the strong classification ability of deep networks on PolSAR data. In Stacked Sparse Autoencoder (SSAE), the application of sparsity effectively eliminates redundant information and obtains complementary information that is more independent.

To more effectively improve the ability of polarized features to distinguish similar image classes, in

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

this paper, we first use the polarized scattering feature decomposition parameters to form a new feature vector and then use the original scattering matrix information to learn jointly with this feature vector to obtain a more effective feature representation. Then, a non-negative constraint method is introduced in SSAE to extract features better. Next, the weights are filtered by non-negative constraints to make the sparse features in the depth space robust, thus overcoming the scattering noise and improving the classification accuracy. Finally, the encoding process of the obtained feature representation network is combined with the softmax classifier to achieve the design of an end-to-end network architecture to accomplish the PolSAR image classification task.

The rest of this paper is structured as follows. In the second part, the data composition is first introduced, and then a non-negativity constraint-based SSAE classification framework is proposed. In Section III, experimental results and discussions are given to illustrate the effectiveness of the proposed method. Finally, the fourth section draws conclusions.

2. Methods

2.1. Data Representation

To fully extract the information in PolSAR images, the multi-look PolSAR data in covariance matrix format are used and transformed into normalized 6-dimensional real vectors. The polarization scattering eigenvalue decomposition is used to obtain the eigenvalues. The entropy, anisotropy, and alpha are extracted using the $H/A/\alpha$ decomposition. This information is combined into a new feature vector and used as the input to the proposed network for PolSAR image classification.

The scattering matrix is used to simply describe the individual pixels in PolSAR images [30]. The covariance matrix contains the full polarization information of the target obtained from radar measurement. The validity of the covariance matrix has been verified in [31]. According to the reciprocity theorem, the covariance matrix can be calculated by the Kronecker product of the $x = \left[S_{HH}\ \sqrt{2}S_{HV}\ \ S_{VV}\right]$ as follows:

$$\langle C\rangle = xx^{*T} = \begin{bmatrix} \langle |S_{HH}|^2\rangle & \sqrt{2}\langle S_{HH}S_{HV}{}^*\rangle & \langle S_{HH}S_{VV}{}^*\rangle \\ \sqrt{2}\langle S_{HV}S_{HH}{}^*\rangle & \sqrt{2}\langle |S_{HV}|^2\rangle & \sqrt{2}\langle S_{HV}S_{VV}{}^*\rangle \\ \langle S_{VV}S_{HH}{}^*\rangle & \sqrt{2}\langle S_{VV}S_{HV}{}^*\rangle & \langle |S_{VV}|^2\rangle \end{bmatrix} \tag{1}$$

where the superscript $*$ and T denote conjugation and transposition respectively. Introducing multi-look processing in covariance matrix to suppress speckle noise in PolSAR images.

$$\langle C\rangle = \frac{1}{L}xx^{*T} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \tag{2}$$

where L is the number of looks. According to the scattering property of PolSAR images, the principal diagonal of the covariance matrix $\langle C\rangle$ is composed of real values, while the rest are complex values. To obtain the real vector form of PolSAR images more effectively, the following representation is proposed:

$$\begin{aligned} AA &= 10\log_{10}（SPAN） \\ BB &= C_{22}/SPAN \\ CC &= C_{33}/SPAN \\ DD &= |C_{12}|/\sqrt{C_{11}\cdot C_{22}} \\ EE &= |C_{13}|/\sqrt{C_{11}\cdot C_{33}} \\ FF &= |C_{23}|/\sqrt{C_{33}\cdot C_{22}} \end{aligned} \tag{3}$$

where AA is the total decibel scattered power of each polarization channel BB and CC are the normalized ratio of C_{22} and C_{33}, respectively. DD, EE, and FF are the relative correlation

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

coefficients. By definition, except for AA, the remaining five parameters can be normalized in [0,1]. Therefore, any multi-look complex (MLC) PolSAR image can be converted to a six-dimensional real vector representation [32].

In general, the covariance matrix can be described further by the eigenvalues and eigenvectors. According to the reciprocity theorem, the covariance matrix can be decomposed into the sum of three independent covariance matrices. The eigenvalue decomposition of the covariance matrix is

$$[E] = \sum_{i=1}^{3} \lambda_i [E_i] = \lambda_1 e_1 e_1^* + \lambda_2 e_2 e_2^* + \lambda_3 e_3 e_3^* \tag{4}$$

where $[E_i]$ denotes the independent covariance matrix of rank 1. λ_i and e_i are the corresponding eigenvalues and eigenvectors, respectively. The eigenvalues indicate the strength of this mechanism, while the corresponding eigenvectors provide information on the direction of matrix change. $H/A/\alpha$ parameters are written as follows [33]:

$$H = \sum_{i=1}^{3} P_i \log_3(P_i) \tag{5}$$

$$P_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \lambda_3} \tag{6}$$

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \tag{7}$$

$$\alpha = \sum_{i=1}^{3} P_i \alpha_i \tag{8}$$

where λ_i denotes the ith eigenvalue of the covariance matrix, H is the randomness of the scattering medium from isotropic scattering (H =0) to completely random scattering (H =1), A indicates the relationship between the two weaker scattering components, α denotes the average scattering mechanism from surface scattering to double-bounce scattering.

2.2. Non-negative constraints stacked sparse Autoencoder Model

AE is a network that can copy input data to output, and it consists of two parts, an encoder and a decoder. The encoder maps input data into a low-dimensional space, while the decoder reconstructs the low-dimensional representation of the data into the original form. The process can be defined as

$$a = f_1(W_1 x + b_1) \tag{9}$$

$$z = f_2(W_2 a + b_2) \tag{10}$$

where W_1, b_1 and W_2, b_2 denote the weight matrix and bias of the encoding and decoding processes, respectively. f_1 and f_2 denote the sigmoid activation function used.

The AE utilizes the Mean Square Error (MSE) for network training and minimizes the objective function, which represents the reconstruction error between the input and output. It is shown in (11)

$$J(W,b) = \underset{W,\ b}{arg\min} \left(\frac{1}{2N} \sum_{i=1}^{N} (x_i - z_i)^2 \right) \tag{11}$$

where $1/2$ is set to simplify the calculation of the subsequent backpropagation process and N represents the total number of training samples. With the random gradient descent algorithm, the optimal solution of the parameters W and b can be obtained, significantly reducing the error between the output and input layers.

The non-negatively constrained sparse AE is an advancement over AE. In non-negatively constrained sparse AE, the cost function is a sparse constraint term and a weight decay term. The sparse

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

constraint term is usually implemented using the KullbackLeibler (KL) scattering function, which restricts the probability distribution of the activation of the hidden unit. It is expressed as

$$\Omega_s = \sum_{m=1}^{m} \rho \log\left(\frac{\rho}{\overline{\rho}_i}\right) + (1-\rho)\log\left(\frac{1-\rho}{1-\overline{\rho}_i}\right) \tag{12}$$

where ρ is the sparsity parameter, $\overline{\rho}_i$ is the average activation value of all input signals in ith hidden node, and m is the number of hidden nodes. During the training process, Ω_s tends to be 0 if the constraint $\overline{\rho}_i = \rho$ is satisfied and when the difference is very large, Ω_s is big.

The AE cost function incorporates a weight decay term to better capture the non-negative features in input data. We employ the asymmetric piecewise linear decay function.

$$\Phi_w = \frac{1}{2}\sum_{l=1}^{L}\sum_{m=1}^{M}\sum_{n=1}^{N}\phi^2(W_{mn}^{(l)}) \tag{13}$$

$$\phi(W_{mn}^{(l)}) = \begin{cases} -W_{mn}^{(l)} & W_{mn} < 0 \\ 0 & W_{mn} \geqslant 0 \end{cases} \tag{14}$$

where $\phi(W_{mn}^{(l)})$ is denoted as (14), W_{mn} is the weight coefficient, l is the number of layers. Therefore, the cost function of the non-negative constraints stacked sparse autoencoder network is given by

$$E = J(W,b) + \gamma\Omega_s + \beta\Phi_w \tag{15}$$

where γ is the weight of the sparse penalty term and β is the weight of the weight decay term. The non-negatively constrained sparse AE contains only one hidden layer, it is difficult to complete the classification task when the input data is too complex. In addition, shallow networks are prone to information loss, so non-negative constraints stacked sparse AE network is proposed to achieve a deeper representation of input data by adding multiple hidden layers to the existing shallow network.

A non-negative constraint stacked sparse AE network is a network model that consists of a stack of multiple hidden layers of AE. The model is trained using a layer-by-layer training approach, where the first AE is trained with a given training set and the activation of the hidden layer of the first AE is used as input to the second AE. The architecture of the proposed network is shown in Fig. 1.

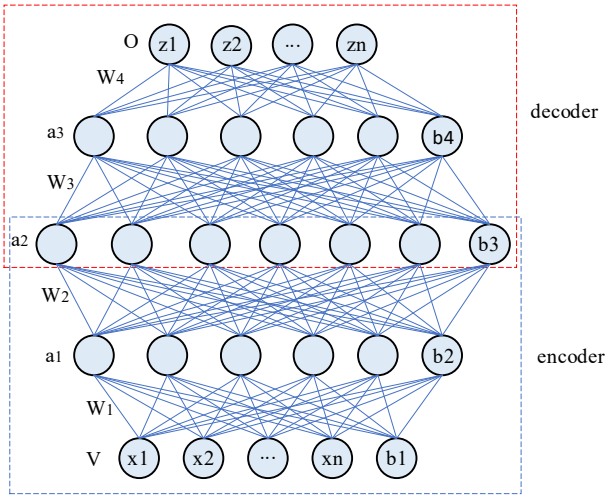


Figure 1: Non-negative constraints stacked sparse Autoencoder structure.

A non-negative constraint stacked sparse AE network is a deep learning model. The network obtains the training sample set V as its input vector, where each sample set V= {x1, x2, ..., xn}, with dimension n equal to the number of neurons in the input layer and also equal to the number of neurons in the output layer. First, by encoding, the proposed network encodes each input vector V into a feature vector a1. Then, the encoding part maps feature vector a1 to feature vector a2. Subsequently, the decoding part decodes feature vector a2 to feature vector a3. Finally, the decoding part reconstructs feature vector a3

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

to the original vector to output the output vector O= {z1, z2, ..., zn}. During this process, b1, b2, b3, and b4 are used as bias terms to adjust the weight between different features in the network.

2.3. Optimization Process

Similar to the autoencoder (AE), the non-negative constraints stacked sparse AE also utilizes a backpropagation algorithm to calculate the overall cost function. To minimize the cost function, each parameter needs to be initialized randomly and the optimization algorithm is applied to the training data. In this process, the L-BFGS optimization algorithm can be used to adjust the parameters in the network.

First, update decoding weight and biases using the stochastic gradient descent algorithm

$$W_2 = W_2 - \eta \frac{\partial E}{\partial W_2} \tag{16}$$

$$b_2 = b_2 - \eta \frac{\partial E}{\partial b_2} \tag{17}$$

where η denotes the learning rate.

Second, the backpropagation algorithm is utilized to calculate the partial derivatives in the overall cost function. The calculation of the overall cost function on the weight in the decoder can be expressed as

$$\frac{\partial E}{\partial W_2} = \frac{\partial J(W,b)}{\partial W_2} + \gamma \frac{\partial \Omega_s}{\partial W_2} + \beta \frac{\partial \Phi_w}{\partial W_2} \tag{18}$$

During weight and bias updates, the parameters of the network are adjusted based on the changes in the cost function to better fit the training data. To specify the operations of the above formula, let us define the variable δ_2 as

$$\delta_2 = \frac{1}{N} \sum_{i=1}^{N} (z_i - x_i) f_2'(W_2 a + b_2) \tag{19}$$

where f_2' is the partial derivative of f_2. By substituting (19) into (18), the partial derivatives of the loss function for the weight and biases in the decoding process can be expressed as follows:

$$\frac{\partial E}{\partial W_2} = \delta_2 a^T + \beta \phi(W_2) sign(W_2) \tag{20}$$

$$\frac{\partial E}{\partial b_2} = \delta_2 \tag{21}$$

where the $sign$ is defined as the symbol function.

Finally, the update of the weight and biases of the encoding process is similar to the decoding process and will not be explained in detail. The result is directly given as follows:

$$W_1 = W_1 - \eta(\delta_1 x^T + \beta \phi(W_1) sign(W_1)) \tag{22}$$

$$b_1 = b_1 - \eta \delta_1 \tag{23}$$

$$\delta_1 = \frac{1}{N} \left(W_2 \delta_2 + \gamma \left(-\frac{\rho}{\overline{\rho}_i} + \frac{1-\rho}{1-\overline{\rho}_i} \right) \right) f'(W_1 x + b_1) \tag{24}$$

2.4. Classification Tasks

The non-negative constraints stacked sparse AE network can extract the features of PolSAR images. The information of interest is contained in the activation of the hidden layer. Thus, the features of the image can be represented by the values of the hidden nodes. To accomplish the classification task of

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

PolSAR images, a small number of labeled pixels can be selected as training samples for the classifier. Therefore, the features obtained after network training and the corresponding label information are selected to train the classifier. In the experiments, the softmax classifier after training is chosen to be connected with the non-negative constraints stacked sparse AE network. The structure is shown in Fig. 2.

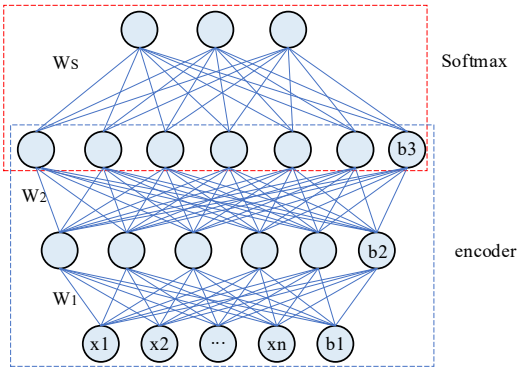


Figure 2: Non-negative constraints stacked sparse Autoencoder classification structure.

The softmax classifier is a commonly used classification algorithm that generates a k-dimensional real-value vector in the range of 0 to 1 to represent the classification probability distribution, which represents the probability of each class, with the maximum probability corresponding to the final classification. The softmax function represents how to predict the probability of a class given a sample vector X and can be denoted as

$$a_w(X) = P(z = j|X;W) = \frac{e^{X^T W_j}}{\sum_{k=1}^{K} e^{X^T W_k}} \tag{25}$$

where k is the total number of categories, W_k denotes the weight vector of category k and W denotes the weight matrix of the softmax classifier.

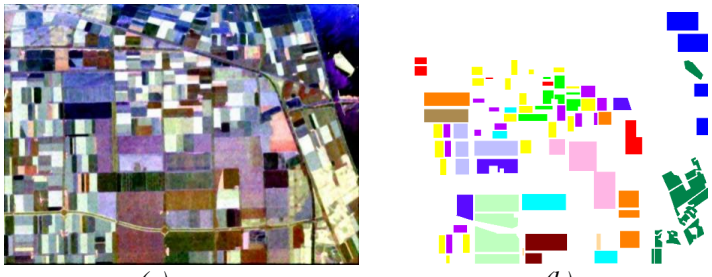
In summary, the features of each pixel point in the PolSAR image extracted by non-negative constraints stacked sparse AE are fed into the trained softmax classifier to obtain the classification results for each category.

3. Experimental Analysis and Results

In our experiments, Considering that PolSAR images are seriously affected by speckle noise, a refined Lee filter [34] in the preprocessing, and the window size of the filter is set to 7 by experience. The effectiveness of the proposed method is verified by SVM, AE, NC-AE, NC-SSAE and GLCM (contrast, correlation, energy, and homogeneity) are spatial-based methods. The classification performance is evaluated by the overall accuracy (OA). Because ground truth labeled samples are few, randomly picked 5% of the labeled samples as training data, and use the whole datasets as test data. This experiment uses the Flevoland and San Francisco Bay datasets, which are described below.

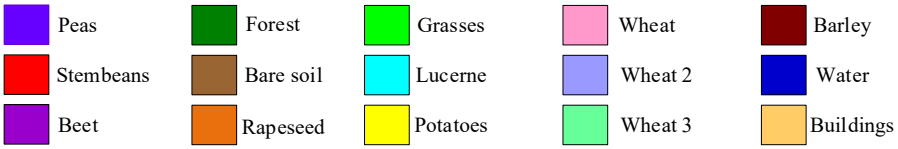
3.1. Flevoland Image

1) Dataset:



(a) *(b)*

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815



(c)

Figure 3: (a) Pauli RGB image of Flevoland. (b) Corresponding ground-truth map. (c) Terrains.

The Flevoland image is obtained from a subset of an L-band, multilook PolSAR data, acquired by the AIR-SAR airborne platform on August 16, 1989. The image size is 750 x 1024, and the total number of labeled pixels in the ground truth is 167712. The Pauli RGB image, the corresponding ground-truth map, and the terrains are shown in Fig. 3. The ground-truth map contains 15 classes.

2) Hyperparameters Selection:

Table 1: OAs (%) of Two Classifiers with Different Values of Hidden Size (N) and Sparsity (ρ) over the Flevoland Image

Sparsity	0.01		0.05		0.10		0.50		1.00	
Hidden size	AE	NC-AE	AE	NC-AE	AE	NC-AE	AE	NC-AE	AE	NC-AE
30	72.56	90.11	73.95	91.21	74.70	91.45	82.77	91.01	74.56	91.77
80	85.46	92.36	87.33	92.04	86.85	92.15	85.78	92.31	85.15	92.33
100	86.83	92.30	85.83	92.73	86.37	92.78	86.17	92.86	85.73	92.54
150	85.23	91.80	85.53	92.68	85.02	92.94	85.65	92.93	84.52	93.1
300	85.43	91.52	85.52	91.36	85.52	91.23	85.48	91.34	85.41	91.38

There are 2 proposed algorithms in our experiments. We do experiments to choose the parameters of the AE and NC-AE models, including hidden node numbers of 50, 80, 100, 150, and 300, and the sparsity, including 0.01, 0.05, 0.1, 0.5, and 1. By setting the single one and keeping the other one to be constants when discussing how these two parameters affect the classification performance, which is presented in Table. It is evident from Table 1 that the AE model is more sensitive to these parameters than the NC-AE model. To balance the classification accuracy, we set the number of hidden nodes to 100 and the sparsity is 0.05. Additionally, the weights of the regularization term and the softmax classifier are also set as $1e^{-6}$.

(a) *(b)*

(c) *(d)*

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

(e) *(f)*

Figure 4: Classification results of the Flevoland image. (a) SVM. (b) AE. (c) NC-AE. (d) NC-SSAE. (e) Texture features+softmax. (f) Proposed.

Table 2: Classification Accuracy of Flevoland

Class	SVM	AE	NC-AE	NC-SSAE	Texture	Proposed
Water	97.45	97.32	98.32	99.80	92.45	99.93
Stem beans	95.69	79.93	95.29	99.26	75.10	99.38
Potatoes	87.78	83.37	85.43	98.29	78.14	99.64
Forest	88.91	92.98	91.86	98.06	89.39	99.54
Grasses	80.29	73.90	91.75	95.71	72.36	97.44
Beet	95.41	84.22	90.51	99.13	83.78	99.35
Rapeseed	79.35	78.73	88.93	94.25	72.25	95.40
Peas	97.13	88.53	93.18	98.28	84.59	98.80
Lucerne	95.76	94.20	95.27	98.45	90.41	98.62
Bare Soil	90.88	97.93	97.98	99.88	93.71	99.88
Wheat2	81.25	80.85	86.10	94.24	75.88	94.96
Wheat	92.19	84.18	91.26	96.34	80.85	97.27
Wheat3	95.35	92.39	96.02	98.57	89.73	99.25
Buildings	4.9	56.73	81.59	93.47	51.69	94.42
Barley	96.84	79.08	97.53	98.68	65.64	99.28
OA	90.57	86.82	92.34	97.66	82.53	98.45

The classification accuracy results are displayed in Table 2, and the classification images are shown in Fig. 4. From the results, the classification accuracy of the proposed method is 98.45%, which is about 7.88%, 11.63%, 6.11%, 0.79%, and 15.92% higher than that of the others, respectively. It is worth noting that not all the categories can be classified effectively by SVM. For instance, the accuracy for buildings is lower than 10%, Rapeseed and Grasses both are approximately 80%. This indicates that the results of SVM classifiers are not satisfactory. The classification networks based on NC-AE is effective to distinguish all categories, and for most classification result, the accuracy of NC-AE approaches 90%. This shows features learned through the NC-AE network can be effectively classified. It should be noted that NC-AE classifies buildings better compared to SVM. As shown in the classification result diagram in Fig. 4(c), even though the data are filtered already using a window size of 7 Lee filter, the classification map of NC-AE is affected by the speckle noise. This may be because the pixel count of buildings is relatively low, causing the classification network to be unable to fully recognize the characteristics of buildings, thus leading to inaccurate classification results. Compare to NC-AE, the OA of NC-SSAE improved by 5.32%, the speckle noise is significantly reduced, and it can effectively predict the class of the pixels in the whole image. A comparison with the texture features obtained using the same softmax classifier shows that the discriminatory power of the texture features extracted from the GLCM is relatively weak compared to using the proposed feature combinations. From the resulting image, it can observe that the visual effect of the proposed method is better than each method, and the misclassified speckle noise is reduced significantly.

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

3.2. San Francisco Bay Image

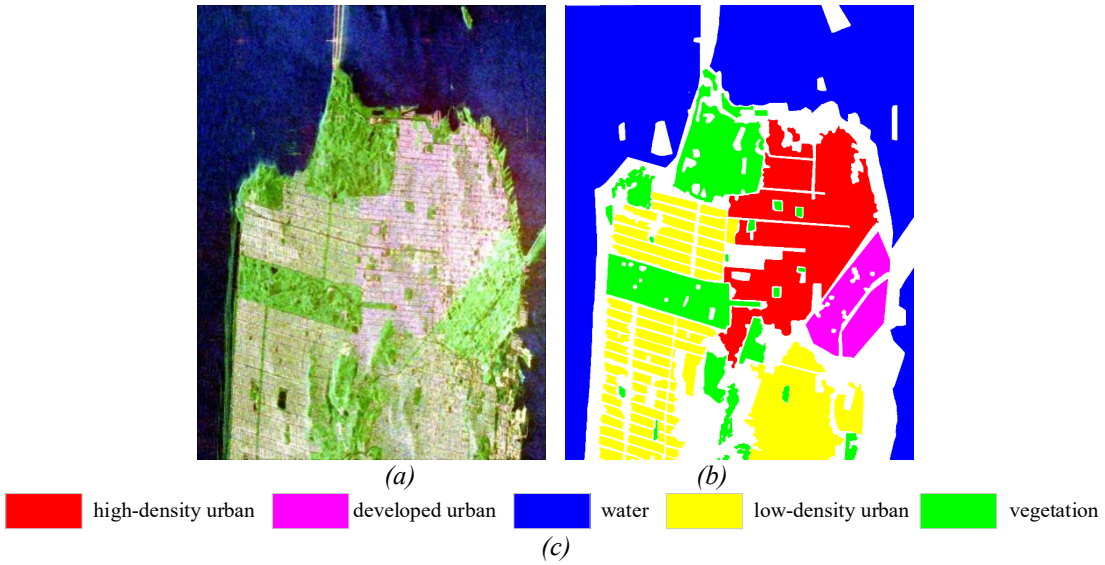


Figure 5: (a) Pauli RGB image of San Francisco Bay. (b) Corresponding ground-truth map. (c) Terrains.

The San Francisco Bay data is obtained from RADARSAT-2 C-band observation. The image size of the selected scene is 1380 x 1800 which contains five recognized crop classes: High-density urban, Water, Vegetation, Developed urban, and Low-density urban. The Pauli RGB image, the ground-truth map, and the terrains are given in Fig. 5. The total number of labeled pixels in the ground truth is 1804087. The Hyperparameters of the AE, NC-AE, and NC-SSAE models are selected in the same way as they are in the Flevoland experiments.

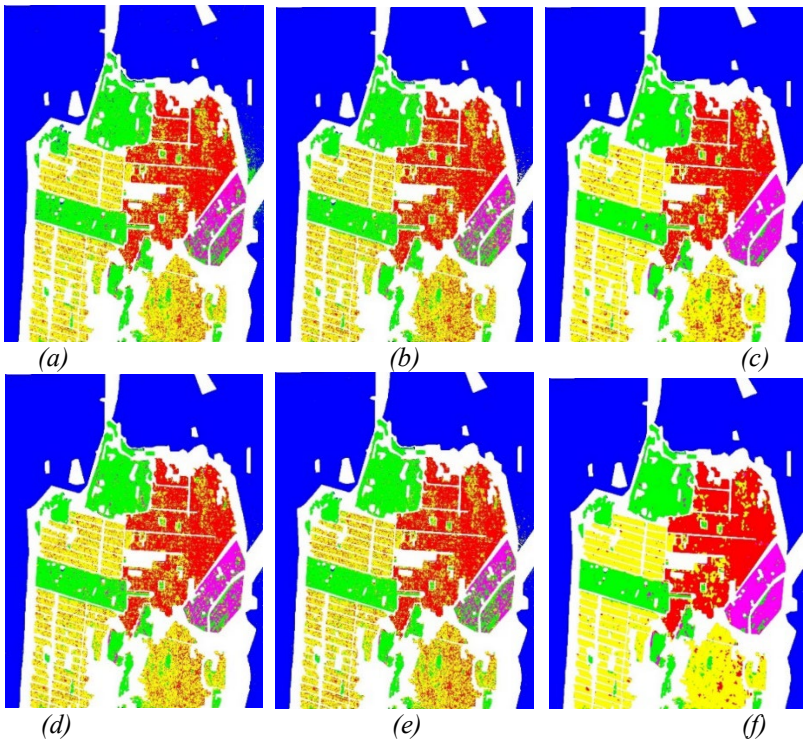


Figure 6: Classification results of the San Francisco Bay image. (a) SVM. (b) AE. (c) NC-AE. (d) NC-SSAE. (e) Texture features+softmax. (f) Proposed.

The classification accuracy results are shown in Table 3, with the classified images shown in Fig. 6. The OA of the proposed method is 95.63%, and compare with other methods improved by 6.8%, 7.63%, 6.29%, 2.8%, and 7.8%, respectively. The accuracy performance of various methods across different categories is listed in Table 3. It can be seen that the proposed method performs best in the

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

Water category, reaching 100% of OA. Additionally, the proposed method also achieves high classification accuracy in other categories. From Fig. 6, it can be seen that the proposed method achieves the most accurate and smooth classification image for the Water category, with lower speckle noise. For the Vegetation category, the OA of NC-SSAE reaches 91.23%, significantly improving on the OA of NC-AE. It can also be seen from the classification image that compare to the ground-truth map, the corresponding speckle noise is significantly reduced. This indicates that using NC-SSAE for feature dimensionality reduction can effectively extract deep features of PolSAR data. Compare to other classification methods, the OA of the Developed urban is almost lower than 70%. The classification performance of the texture features is almost the poorest among all methods, with a value of only 52.01% for the Developed urban, significantly lower than the OA of the proposed method. This indicates that the proposed feature combinations have a better feature expression. Additionally, both SVM, AE, and Texture performed poorly on OA in the urban area, with values lower than 80%. Possible reasons for this include a lack of sufficient training data, which is insufficient to satisfy the structure parameters of the network. However, after the proposed method is improved, the OA significantly improves.

Table 3: Classification Accuracy of San Francisco Bay

Class	SVM	AE	NC-AE	NC-SSAE	Texture	Proposed
Water	99.25	99.81	99.93	99.97	99.81	100.00
Vegetation	88.23	87.00	87.73	91.23	88.58	92.25
High-density urban	73.79	73.11	73.99	80.12	74.68	86.80
Low-density urban	82.18	79.52	81.82	88.37	77.07	95.35
Developed urban	62.32	55.22	68.75	86.05	52.01	91.74
OA	88.83	88.00	89.34	92.83	87.83	95.63

4. Conclusion

A non-negative constraint stacked sparse autoencoder classification framework is constructed to fully extract the information from PolSAR images. The framework combines the 6-dimensional real vector derived from the covariance matrix with the polarization feature vector obtained by the polarization decomposition method as the input. Through a detailed analysis of the combined vectors, a non-negative constraint method is introduced into SSAE to achieve the screening of weights and improve the discrimination of image classes with similar polarization features. The framework is evaluated using the Flevoland and San Francisco Bay image and compared with the traditional SVM, AE, and Gray Level Coevolution Matrix (GLCM) for classification. The results show that the proposed classification framework has a significant improvement in classification performance and has better robustness.

Acknowledgements

This research was funded partly by the National Natural Science Foundation of China under Grant 62201201; the Doctoral Foundation of Henan Polytechnic University under Grant B2022-15.

References

[1] Moreira A., Prats-Iraola, et al. A tutorial on synthetic aperture radar [J]. Geoscience & Remote Sensing Magazine IEEE, 2013.
[2] Yin Q, Hong W, Zhang F, et al. Optimal Combination of Polarimetric Features for Vegetation Classification in PolSAR Image[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, PP (99):1-13.
[3] West R D, Riley R M. Polarimetric Interferometric SAR Change Detection Discrimination [J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(6).
[4] Wang H, Xu F, Jin Y Q. A Review of Polsar Image Classification: from Polarimetry to Deep Learning: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 2019.
[5] Lim Y X, Burgin M S, Zyl J J V. An Optimal Nonnegative Eigenvalue Decomposition for the Freeman and Durden Three-Component Scattering Model [J]. IEEE Transactions on Geoscience & Remote Sensing, 2017, 55(4):2167-2176.
[6] Lemme A ,René Felix Reinhart, Steil J J .Online learning and generalization of parts-based image representations by non-negative sparse autoencoders[J].Neural Networks, 2012, 33(9):194-203.DOI:10.1016/j.neunet.2012.05.003.
[7] Hua W, Wang S, Liu H, et al. Semisupervised PolSAR Image Classification Based on Improved

ISSN 2616-5775 Vol. 6, Issue 8: 119-130, DOI: 10.25236/AJCIS.2023.060815

Cotraining [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, PP (11):1-16.
[8] Yahia M, Khalfa F, Chabir M, et al. Bias compensation in h/a/A polarimetric sar decomposition and its implication for the classification[J]. Progress in Electromagnetics Research B, 2016, 68(1):105-121.
[9] Manickam S, Bhattacharya A, Singh G, et al. Estimation of Snow Surface Dielectric Constant From Polarimetric SAR Data [J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2016.
[10] Yan L L, Zhang J X, Gao J X, et al. Four-Component Model-Based Decomposition of Polarimetric SAR Data for Oriented Urban Buildings[J]. Acta Electronica Sinica, 2015, 43(1):203-208.
[11] Wang Y, Liu H, Jiu B. PolSAR Coherency Matrix Decomposition Based on Constrained Sparse Representation[J]. IEEE Transactions on Geoscience & Remote Sensing, 2014, 52(9):5906-5922.
[12] Cloude S R, Pottier E. An entropy based classification scheme for land applications of polarimetric SAR [J]. IEEE Trans. Geo. Remote Sens, 1997, 35(1):68-78.
[13] Freeman A, Durden S L. A three-component scattering model for polarimetric SAR data [J]. IEEE Transactions on Geoscience & Remote Sensing, 1998, 36(3):963-973.
[14] Krogager E. New decomposition of the radar target scattering matrix [J]. Electronics Letters, 2002, 26(18):1525-1527.
[15] Van Zyl J J. Unsupervised classification of scattering behavior using radar polarimetry data [J]. IEEE Trans. geosci. remote Sens, 1989, 27(1):36-45.
[16] Zhou Y, Wang H, Xu F, et al. Polarimetric SAR Image Classification Using Deep Convolutional Neural Networks [J]. IEEE Geoscience & Remote Sensing Letters, 2017, 13(12):1935-1939.
[17] Hinton G E. Learning multiple layers of representation [J]. Trends in Cognitive Sciences, 2007, 11(10): 428-434.
[18] Baldi P, Guyon G, Dror V, et al. Autoencoders, unsupervised learning and deep architectures: UTLW'11 Proceedings of the 2011 International Conference on Unsupervised and Transfer Learning workshop - Volume 27, 2011.
[19] Yin Y, Gelenbe E. Nonnegative autoencoder with simplified random neural network[J]. 2016. DOI:10.48550/arXiv.1609.08151.
[20] Bengio Y. Learning Deep Architectures for AI [J]. Foundations & Trends in Machine Learning, 2009, 2(1):1-127.
[21] Xu J, Xiang L, Hang R, et al. Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. [J]. IEEE, 2014.
[22] Zhou Y, Wang H, Xu F, et al. Polarimetric SAR Image Classification Using Deep Convolutional Neural Networks[J]. IEEE Geoscience & Remote Sensing Letters, 2017, 13(12):1935-1939.
[23] Chen S W, Tao C S. PolSAR Image Classification Using Polarimetric-Feature-Driven Deep Convolutional Neural Network [J]. IEEE Geoscience and Remote Sensing Letters, 2018.
[24] Geng J, Fan J, Wang H, et al. High-Resolution SAR Image Classification via Deep Convolutional Autoencoders [J]. IEEE Geoscience & Remote Sensing Letters, 2015, 12(11):2351-2355.
[25] Zhang L, Ma W, Zhang D. Stacked Sparse Autoencoder in PolSAR Data Classification Using Local Spatial Information[J]. IEEE Geoscience & Remote Sensing Letters, 2017, 13(9):1359-1363.
[26] Hou B, Kou H, Jiao L. Classification of Polarimetric SAR Images Using Multilayer Autoencoders and Superpixels [J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017, 9(7):3072 -3081.
[27] Wang J, Hou B, Jiao L, et al. POL-SAR Image Classification Based on Modified Stacked Autoencoder Network and Data Distribution[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, PP(99):1-18.
[28] Zhang L, Jiao L, Ma W, et al. PolSAR Image Classification Based on Multi-scale Stacked Sparse Autoencoder [J]. Neurocomputing, 2019, 351(JUL. 25):167-179.
[29] Hu Y, Fan J, Wang J. Classification of PolSAR Images Based on Adaptive Nonlocal Stacked Sparse Autoencoder [J]. IEEE Geoence & Remote Sensing Letters, 2018:1-5.
[30] Li Y, Chen Y, Liu G, et al. A Novel Deep Fully Convolutional Network for PolSAR Image Classification [J]. Remote Sensing, 2018, 10(12):1984.
[31] Chen S W, Wang X S, Sato M. PolInSAR Complex Coherence Estimation Based on Covariance Matrix Similarity Test [J]. IEEE Transactions on Geoscience and Remote Sensing, 2012, 50(11):4699-4710.
[32] Zhou Y, Wang H, Xu F, et al. Polarimetric SAR Image Classification Using Deep Convolutional Neural Networks[J]. IEEE Geoscience & Remote Sensing Letters, 2017, 13(12):1935-1939.
[33] Cloude S R, Pottier E. A review of target decomposition theorems in radar polarimetry [J]. IEEE Transactions on Geoscience & Remote Sensing, 1996, 34(2):498-518.
[34] Lee J S, Grunes M R, De Grandi G. Polarimetric SAR speckle filtering and its implication for classification [J]. IEEE Transactions on Geoscience & Remote Sensing, 1999, 37(5):2363-2373.