

Polarimetric Convolutional Network for PolSAR Image Classification

Xu Liu, *Student Member, IEEE*, Licheng Jiao, *Fellow, IEEE*, Xu Tang, *Member, IEEE*,
Qigong Sun, *Student Member, IEEE*, Dan Zhang

Abstract—The approaches for analyzing the polarimetric scattering matrix of polarimetric synthetic aperture radar (PolSAR) data have always been the focus of PolSAR image classification. Generally, the polarization coherent matrix and the covariance matrix obtained by the polarimetric scattering matrix are used as the main research object to extract features. In this paper, we focus on the original polarimetric scattering matrix and propose a polarimetric scattering coding way to deal with polarimetric scattering matrix and obtain a close complete feature. This encoding mode can also maintain polarimetric information of scattering matrix completely. At the same time, in view of this encoding way, we design a corresponding classification algorithm based on convolution network to combine this feature. Based on polarimetric scattering coding and convolution neural network, the polarimetric convolutional network is proposed to classify PolSAR images by making full use of polarimetric information. We perform the experiments on the PolSAR images acquired by AIRSAR and RADARSAT-2 to verify the proposed method. The experimental results demonstrate that the proposed method get better results and has huge potential for PolSAR data classification. Source code for polarimetric scattering coding is available at <https://github.com/liuxuvip/Polarimetric-Scattering-Coding>.

Index Terms—Classification, PolSAR, polarimetric scattering matrix, convolution network.

I. INTRODUCTION

Polarimetric synthetic aperture radar (PolSAR) images collected with airborne and satellite sensors are a wealthy source of information concerning the Earth's surface, and have been widely used in urban planning, agriculture assessment and environment monitoring [1], [2]. These applications require the fully understanding and interpretation of PolSAR images.

Hence, PolSAR image interpretation is of much significance in theory and application. Land use classification of PolSAR images are an important and indispensable research topic since these images contain rich character of the target (e.g., scattering properties, geometric shapes, and the direction of arrival). The land use classification is arranging the pixels

This work was supported in part by the Major Research Plan of the National Natural Science Foundation of China (No. 91438201 and No. 91438103), the National Natural Science Foundation of China (No. 61801351) the National Science Basic Research Plan in Shaanxi Province of China (No.2018JQ6018), the Fund for Foreign Scholars in University Research and Teaching Programs (the 111 Project) (No. B07048), the Fundamental Research Funds for the Central Universities (No. XJS17108) and the China Postdoctoral Fund (No. 2017M613081).

The authors are with the Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education of China, International Research Center of Intelligent Perception and Computation, School of Artificial Intelligence, Xidian University, Xi'an, Shaanxi Province 710071, China (e-mail: xuli361@163.com; lchjiao@mail.xidian.edu.cn).

to the different categories according to the certain rule. The common objects within the PolSAR images include land, buildings, water, sand, urban areas, vegetation, road, bridge and so on [3]. In order to distinguish them, the features of the pixels should be fully extracted and mined. With the development of the PolSAR image classification, many feature extraction algorithms based on physical scattering mechanisms have been introduced. The feature extraction techniques based on polarimetric characteristics can be divided into two kinds: coherent target decomposition and incoherent target decomposition. The former acts on the scattering matrix to characterize completely polarized scattered waves, which contains the fully polarimetric information. The latter acts only on the mueller matrix, covariance matrix, or coherency matrix in order to characterize partially polarized waves [4].

The coherent target decomposition algorithms include the Pauli decomposition, the sphere-diplane-helix (SDH) decomposition [5], the symmetric scattering characterization method (SSCM) [6], Cameron decomposition [7], Yamaguchi Four-component scattering power decomposition[8], General polarimetric model-based decomposition [9], [10], and some advances [11], [12]. The incoherent target decomposition algorithms include Huynen decomposition [13], Freeman-Durden decomposition [14], Yamguchi four-component decomposition [15], Cloude-Pottier decomposition [16], [17], and a number of approaches have been reported[18], [19], [20]. In addition to feature based on the polarization mechanism [21], [22], [23], [24], there are some traditional features of natural images, which have been utilized to analyze PolSAR image, such as color features [25], texture features [26], spatial relations [27], etc. Based on the above basic features, some multiple features of PolSAR data have been constructed to improve the classification performance [28], [25], [29].

For classification tasks, besides the feature extraction, classifier design is also a key point. According to the degree of data mark, the classification methods can be broadly divided into three groups, including unsupervised classification (without any labeled training data), semi-supervised classification (SSC) (with a small amount of labeled data and a large amount of unlabeled data) and supervised classification (with completely labeled training data) [30].

The unsupervised classification approaches design a function to describe hidden structure from unlabeled data. The traditional methods always make a decision rule to cluster PolSAR data into different groups, and the number of groups is also a hyper-parameter. There are a lot of unsupervised classification methods for PolSAR data, such as H/ α complex

Wishart classifier [31], polarimetric scattering characteristics preserved method [32], Fuzzy k-means cluster classifier [33], [34], the classification based on deep learning [35], [36], [37], etc. The SSC is a class of supervised learning tasks and techniques that also make use of a small amount of labeled data and a large amount of unlabeled data for training. Compared with the unsupervised method, SSC can improve classification performance so long as the target is to make full use of a smaller number of labeled samples [38]. There are many semi-supervised classification for PolSAR data, such as the classification based on hypergraph learning [39], the method based on parallel auction graph [40], spatial-anchor graph [38], etc. Unlike unsupervised approaches and semi-supervised approaches, the supervised classifications use enough labeled samples to train the classifiers which can be applied to determine the class of other samples. Lots of methods have been introduced, including maximum likelihood [41], support vector machines [42], [43], [44], sparse representation [45], deep learning [46], [47], [48].

Recently, deep learning has attracted considerable attention in the computer vision community [49], [50], [51], [52], as it provides an efficient way to learn image features and to represent certain function classes far more efficiently than shallow ones [53], [54], [55]. Deep learning has also been introduced into the geoscience and remote sensing (RS) community [35], [56], [57], [58], [59], [60], [61], [62], [63], [64]. Especially in the direction of PolSAR image classification, in [36], a specific deep model for polarimetric synthetic aperture radar (POLSTAR) image classification is proposed, which is named as Wishart deep stacking network (W-DSN). A fast implementation of Wishart distance is achieved by a special linear transformation, which speeds up the classification of POLSTAR image. In [35], a new type of restricted boltzmann machine (RBM) is specially defined, which we name the Wishart-Bernoulli RBM (WBRBM), and is used to form a deep network named as Wishart Deep Belief Networks (W-DBN). In [65], a new type of autoencoder (AE) and convolutional autoencoder (CAE) is specially defined, which we name them Wishart-AE (WAE) and Wishart-CAE (WCAE). In [66], a complex-valued CNN (CV-CNN) specifically for synthetic aperture radar (SAR) image interpretation. It utilizes both amplitude and phase information of complex SAR imagery. In [67], Si-Wei Chen, Xue-Song Wang and Motoyuki Sato proposed a polarimetric-feature-driven deep convolutional neural network (PFDCN) for PolSAR image classification. The core idea of which is to incorporate expert knowledge of target scattering mechanism interpretation and polarimetric feature mining to assist deep CNN classifier training and improve the final classification performance.

What is more, fully convolutional network (FCN) is successfully used for natural image semantic segmentation [68], [69], [70], [71], [72], [73], [74] and remote sense image classification based on one by one pixel [75], [76], [77], [78]. In [75], a fully convolutional neural network is trained to segment water on Landsat imagery. In [76], a novel deep model, i.e., a cascaded end-to-end convolutional neural network (CasNet), was proposed to simultaneously cope with the road detection and centerline extraction tasks. Specifically, CasNet

consists of two networks. One aims at the road detection task. The other is cascaded to the former one, making full use of the feature maps produced formerly, to obtain the good centerline extraction. In [77], a novel hyperspectral image classification (HSIC) framework, named deep multiscale spatial-spectral feature extraction algorithm, was proposed based on fully convolutional neural network. In [78], the authors presented a CNN-based system relying on a downsample-then-upsample architecture. Specifically, it first learns a rough spatial map of high-level representations by means of convolutions and then learns to upsample them back to the original resolution by deconvolution. By doing so, the CNN learns to densely label every pixel at the original resolution of the image.

Inspired by the previous research works, a supervised PolSAR image classification method based on polarimetric scattering coding and convolution network is proposed in this paper. Our goal is to solve the problems of PolSAR data coding, feature extraction, and land cover classification. Our work can be summarized into three main parts as follows.

- Firstly, a new encoding mode of polarimetric scattering matrix is proposed, which is called polarimetric scattering coding. It not only completely preserves the polarization information of data, but also facilitates to extract high-level features by deep learning, especially convolutional networks.
- Secondly, a novel PolSAR image classification algorithm based on polarimetric scattering coding and convolutional network is proposed, which is called polarimetric convolutional network and also an end-to-end learning framework.
- Thirdly, feature aggregation is designed to fuse the two kinds of feature and mine more advanced features.

The paper is organized as follows. In Section II, the representation of PolSAR images is described. In Section III, the proposed method named polarimetric convolutional network is given. The experimental setting is presented in Section IV. The results are presented in Section V, followed by the conclusions and discussions in Section VI.

II. PROPOSED METHOD

In the PolSAR image classification task, the land use classes are determined by different analysis include polarization of the target responses, scattering heterogeneity determination and determination of the polarization state for target discrimination, which need to be decided by different features. It is hard for researchers to consider all kinds of features. PolSAR data is a two-dimensional complex matrix. The traditional feature extraction method represents PolSAR data into a one-dimensional vector, which destroys data space structures. In order to solve this problem, the intuitive way is to express the original data directly. In this paper, the polarimetric scattering coding is proposed to express the original data directly, which can maintain structure information completely. Next, the polarimetric scattering coding matrix obtained by the encoding is fed into a classifier based on fully convolutional network. The following of this section consists of three parts. First, representation of PolSAR images is given. Second, the

polarimetric scattering coding for complex scattering matrix S is explained. Third, the proposed method called polarimetric convolutional network is presented.

A. Representation of PolSAR images

The fully PolSAR measures the amplitudes and phases of backscattering signals in four combinations: 1) HH; 2) HV; 3) VH; and 4) VV. Where H means horizontal mode, V means vertical mode. These signals form a 2×2 complex scattering matrix S to represent the information for one pixel, which relates the incident and the scattered electric fields. Scattering matrix S can be expressed as

$$S = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} = \begin{bmatrix} |S_{HH}| e^{i \times \phi_{HH}} & |S_{HV}| e^{i \times \phi_{HV}} \\ |S_{VH}| e^{i \times \phi_{VH}} & |S_{VV}| e^{i \times \phi_{VV}} \end{bmatrix} \quad (1)$$

where S_{HH}, S_{HV}, S_{VH} and S_{VV} are the complex scattering coefficients, S_{HV} is the scattering coefficient of the horizontal(H) transmitting and vertical(V) receiving polarization. Other elements have similar definitions. $|S_{HH}|, |S_{HV}|, |S_{VH}|, |S_{VV}|$ denote the amplitudes of the measured complex scattering coefficients, $\phi_{HH}, \phi_{HV}, \phi_{VH}$ and ϕ_{VV} are the value of phases. i is the complex unity.

The characteristics of the target can be specified by vectorizing the scattering matrix. Based on two important basis sets, lexicographic basis and Pauli spin matrix set, in the case of monostatic backscattering with reciprocal medium, the lexicographic scattering vector \vec{k}_L and Pauli scattering vector \vec{k}_p are defined as

$$\vec{k}_L = [S_{HH}, \sqrt{2}S_{HV}, S_{VV}]^T \quad (2)$$

$$\vec{k}_p = 1/\sqrt{2}[S_{HH} + S_{VV}, S_{HH} - S_{VV}, 2S_{HV}]^T \quad (3)$$

where superscript T denotes the transpose of vector.

The scattering characteristics of a complex target are determined by different independent subscatterers and their interaction. The scattering characteristics described by a statistic method due to the randomness and depolarization. Moreover, the inherent speckle in the SAR data reduced by spatial averaging at the expense of lossing spatial resolution. Therefore, for the complex target, the scattering characteristics should be described by statistic coherence matrix or covariance matrix. Covariance and coherence matrices can be generated from the outer product of \vec{k}_L and \vec{k}_p respectively, with its conjugate transpose

$$\mathbf{C} = \langle \vec{k}_L \vec{k}_L^\mathbf{H} \rangle \quad (4)$$

$$\mathbf{T} = \langle \vec{k}_p \vec{k}_P^\mathbf{H} \rangle \quad (5)$$

where $\langle \cdot \rangle$ denotes the average value in the data processing stage, and the superscript \mathbf{H} stands for the complex conjugate and transpose of vector and matrix.

The covariance matrix \mathbf{C} has been proved to follow a complex Wishart distribution [79]. Moreover, the coherence matrix \mathbf{T} is used to express PolSAR data, which has a linear relation with covariance matrix \mathbf{C} . The PolSAR features always extracted indirectly from the PolSAR data, such as color features, texture features, and the decomposition features. The

color and texture features are extracted from the pseudo-color image which is comprised of the decomposition components. The spatial relation of the pixels could be obtained from such a PolSAR pseudo-color image. The decomposition features are made up of matrix \mathbf{C} or \mathbf{T} by polarimetric target decompositions, e.g., Pauli decomposition, Cloude decomposition, Freeman-Durden decomposition, etc. A number of works have included the computations of these features as shown in [80], [25].

B. Polarimetric scattering coding

Polarimetric data returned by polarimetric synthetic aperture radar is stored in the polarimetric scattering matrix. Polarimetric scattering matrix is used to show polarimetric information, in which the element is complex value. Inspired by the one-hot coding [81] and hash coding [82], we learned from the idea of position encoding and mapping relationship, proposed the polarimetric scattering coding for complex matrix encoding. We assume that $z = (x + yi)$ is a complex value, x and y are the real and imaginary parts of z respectively. Considering the sign of x and y , there are four possible for z , we give a complete encoding as follows, which is called sparse scattering coding φ by us.

$$\varphi(x + yi) = \begin{cases} \begin{bmatrix} x & 0 \\ 0 & |y| \end{bmatrix}, & \text{if } x \geq 0 \text{ and } y \leq 0 \\ \begin{bmatrix} x & y \\ 0 & 0 \end{bmatrix}, & \text{if } x \geq 0 \text{ and } y > 0 \\ \begin{bmatrix} 0 & y \\ |x| & 0 \end{bmatrix}, & \text{if } x < 0 \text{ and } y > 0 \\ \begin{bmatrix} 0 & 0 \\ |x| & |y| \end{bmatrix}, & \text{if } x < 0 \text{ and } y \leq 0 \end{cases} \quad (6)$$

$$\begin{array}{c} \text{Real} \quad \vdots \quad \text{Imag} \\ + \begin{bmatrix} x & \vdots & 0 \\ \dots & \ddots & \dots \\ 0 & \vdots & |y| \end{bmatrix} \\ - \end{array}$$

Fig. 1. The polarimetric scattering coding graphic of Eq. (6). When $x > 0, y < 0$.

Fig. 1 and Eq. (6) show the details of polarimetric scattering coding, the first column represents the real element, the second column represents the imaginary part of the element, the first line represents the positive element, and negative elements are expressed in the second row. The $|\cdot|$ is the absolute value operation. An example in Eq. (6) is given as follows

$$\varphi(x + yi) = \begin{bmatrix} x & 0 \\ 0 & |y| \end{bmatrix}, \text{ if } x \geq 0 \text{ and } y < 0 \quad (7)$$

φ represents the function of polarimetric scattering coding, when $x > 0, y < 0$. From Eq. (1), scattering matrix S can been written as

$$S = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \quad (8)$$

Because S is a complex matrix, we can write its elements as follows

$$S_{HH} = a + bi \quad (9)$$

$$S_{HV} = c + di \quad (10)$$

$$S_{VH} = e + fi \quad (11)$$

$$S_{VV} = g + hi \quad (12)$$

In order to facilitate the understanding and explanation, we give a general assumption, where $a, b, e, h > 0$, $c, d, f, g < 0$. This assumption can take into account the characteristics of the PolSAR data. For instance, some PolSAR data format is int16 (-32,768 to +32,767). Polarimetric scattering coding of the scattering matrix S can be given, which is called polarimetric scattering coding matrix.

$$\begin{aligned} \varphi(S) &= \varphi \left(\begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \right) \\ &= \varphi \left(\begin{bmatrix} a + bi & c + di \\ e + fi & g + hi \end{bmatrix} \right) \\ &= \begin{bmatrix} a & b & 0 & 0 \\ 0 & 0 & |c| & |d| \\ \hline e & 0 & 0 & h \\ 0 & |f| & |g| & 0 \end{bmatrix} \end{aligned} \quad (13)$$

Based on this new coding way, we can get the polarimetric scattering coding matrix, which is a 2-D sparse matrix. We also avoid transforming a complex matrix into a 1-D vector, as is shown in Fig 2. This will bring great convenience to the processing of PolSAR data. The proposed polarimetric coding could show more information and is easy to generate and restore polarimetric covariance matrix. Next to it, we proposed a new corresponding classification method named polarimetric convolutional network.

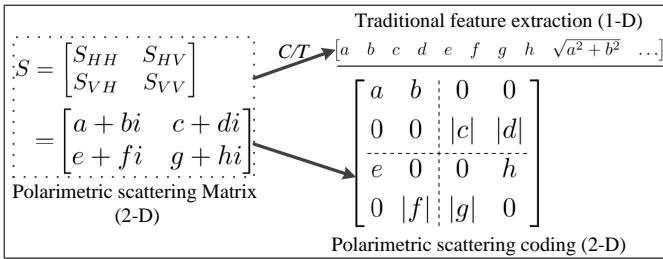


Fig. 2. The difference of traditional feature extraction and polarimetric scattering coding for one pixel in the PolSAR image.

Algorithm 1 Polarimetric scattering coding

Input: The raw polarimetric data

Output: Polarimetric scattering coding matrix;

- 1: Calculate polarimetric data and get scattering matrix S
- 2: Polarimetric scattering coding S and get polarimetric scattering coding matrix $\varphi(S)$
- 3: **return** polarimetric scattering coding matrix

Algorithm 2 Polarimetric convolutional network

Input: The raw polarimetric data I

Output: Classification map;

- 1: Calculate I and get S_I
 - 2: Polarimetric scattering coding S_I and get S_{Ir}
 - 3: Input S_{Ir} to a two layer convnet and get S_{Ir2}
 - 4: Input I and S_{Ir2} to two different FCNs individually, and get two feature maps.
 - 5: Stack and aggregate the two feature maps.
 - 6: Classify the aggregated feature by softmax classifier
 - 7: Update 3-6 by stochastic gradient descent
 - 8: **return** classification map
-

C. Polarimetric convolutional network

Recently, convolutional network performs well in the task of image semantic segmentation, especially fully convolutional network (FCN). Actually, compared to other models, FCN can not only predict class attributes, but also infer the spatial location and contour information, which are helpful to the pixel level classification task. Differences with conventional convolutional network, FCN replaces the fully connected layers with convolutional layers to output spatial maps instead of classification scores, those maps are upsampled using deconvolution to obtain dense per-pixel labeled outputs. The components of FCN mainly includes convolutional layer (encode) and deconvolutional layer (decode). In the forward propagation stage, the three-dimensional input data was down-sampled in the low convolutional layers with relatively convolutional filters. The intermediate feature maps were up-sampled in the high deconvolutional layers with corresponding filters. At last, a softmax classifier followed to predict pixel-wise labels for an output map which has the same size as the input image. The output of the softmax classifier is a C channels map of probabilities where C is the number of classes. In the back propagation learning stage, stochastic gradient descent was used to calculate parameters according to the difference between predictions and ground truth maps.

In this paper, a polarimetric convolutional network is proposed for polarimetric SAR data classification, as is shown in Fig. 3. First, the raw polarimetric data I can be equivalently converted into polarimetric scattering matrix S_I , the size of I is $s_1 \times s_2 \times 8$, the size of S_I is $(2 \times s_1) \times (2 \times s_2)$, which is a complex value matrix. Second, polarimetric scattering matrix can be transformed from complex value matrix to real value matrix $S_{Ir} = \varphi(S_I)$ by polarimetric scattering coding, the size of S_{Ir} is $4s_1 \times 4s_2$. Third, the sparse scattering matrix S_{Ir} is fed into a two layers convolutional networks and get feature maps S_{Ir2} with a size of $s_1 \times s_2$. Fourth, the feature maps generated from the two-layer convolutional networks can be fed into a fully convolutional network whose input size is same as the output size. Fifth, another pathway is also a fully convolutional network, the input data is the raw polarimetric data I . Sixth, the part of feature aggregation is designed to fuse the two kinds of the feature. The last layer feature maps of the two pathway are stacked together as new feature maps, the size of the new feature maps is $s_1 \times s_2 \times (2 \times M)$, M is the number of

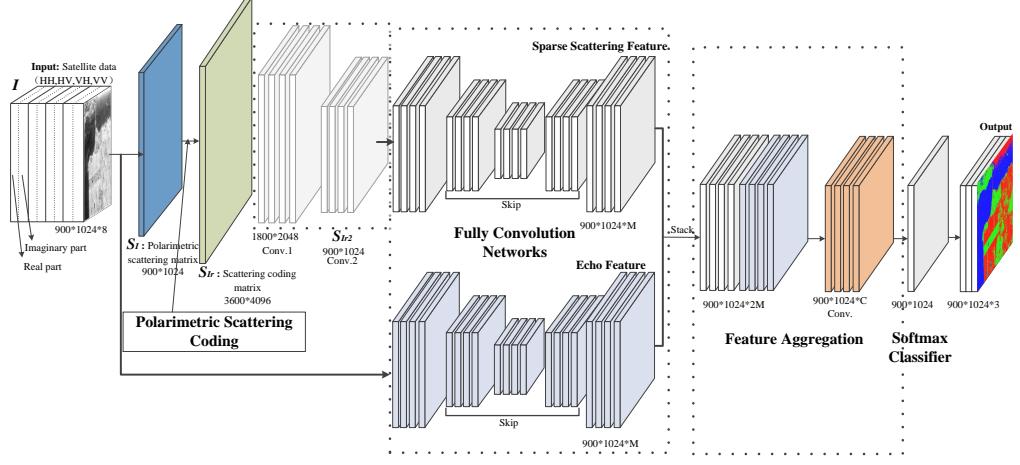


Fig. 3. The flowchart of polarimetric convolutional network. There are 8 channels in the original input data I , which contains four modes, and each has real and imaginary parts. For a pixel, eight channel elements can be represented by a-b in Formula. 1, and the input data I is encoded into a complex value matrix S_I , called polarimetric scattering matrix. By polarimetric scattering coding, polarimetric scattering coding matrix S_{Ir} can be obtained. The feature map S_{Ir2} is generated from the two layer convolutional networks. S_I and S_{Ir2} are fed into two fully convolutional networks to get sparse scattering feature and echo feature, respectively. Finally, two kinds of feature from the two networks are aggregated to get the classification results.

feature maps in each pathway. At last, in order to use softmax to classify data into C classes, we add a convolutional layer to reduce the dimension of the new feature maps from $(2 \times M)$ to C .

Above all, The whole network is an end-to-end network. For the entire network architecture, the loss function is a sum over the spatial dimensions of the final layer. We can use stochastic gradient descent to optimize it.

In the proposed method, the model mainly includes polarimetric scattering coding layer, convolutional layer, and deconvolutional layer. The details of polarimetric scattering coding layer are in Section. II-B. The convolutional layer can be defined as below.

$$O_{ij} = F \left[\sum_{m_i, m_j=0}^k G(X_{s_i+m_i, s_j+m_j}) \right] \quad (14)$$

where O_{ij} is the output map of the convolutional layer in i row and j column, F denotes the normalized function, rectified linear unit (ReLU) is a good choice and used in this article. G is the convolutional function, k is the size of the kernel, X denotes the input pixels from the form layer, s denotes the sampling stride.

The deconvolution layer is composed of upsampling and convolutional layers, upsampling layer corresponds to a max-pooling one in the downsampling stage. Those layers upsample feature maps using the max-pooling indices from their corresponding feature maps. The upsampled maps are then convolved with a set of trainable convolutional kernels to produce dense feature maps.

At last, we compute the energy function by a pixel-wise softmax over the final feature maps and the cross entropy loss function. The softmax is defined as

$$p_c(x) = \frac{\exp(a_c(x))}{\sum_{c'=1}^C \exp(a_{c'}(x))} \quad (15)$$

where $a_{c'}(x)$ denotes the activation in the c' th feature channel at the pixel position $x \in \Omega$ with $\Omega \subset \mathbb{Z}^2$, C is the number of land cover categories, $p_c(x)$ is the approximated maximum function, which represents the probability of each pixel for each category. The cross entropy calculates the deviation of $p_{\iota(x)}(x)$ at each position.

$$E = \sum_{x \in \Omega} \log(p_{\iota(x)}(x)) \quad (16)$$

where $\iota : \Omega \rightarrow \{1, \dots, C\}$ is the ground truth of each pixel, $\iota(x)$ is the label in the location of x .

III. EXPERIMENTS

In this section, four PolSAR images are described in detail, which are used to verify the performance of the proposed algorithm. The four images have significant representativeness, obtaining from two airborne systems and two cities, the details are listed in Table I-V. The parameter settings of the proposed method are discussed. Meanwhile, evaluation metrics are given.

A. Data set description

For our experiments and evaluations, we select four PolSAR images from a spaceborne system (Canadian Space Agency RADARSAT-2) and an airborne system (NASA/JPL-Caltech AIRSAR). The AIRSAR supports full polarimetric modes for C, L, and P-bands where we focus on the L and C-bands. RADARSAT-2 works in C-band with the support of full polarimetric mode. The four selected PolSAR pseudo images are from two different areas, including Flevoland,



Fig. 4. Pauli RGB image of the San Francisco big figure. The whole data of this image downloaded from the RADARSAT-2 web [83], and the area in box is often used. The area coordinate is (7326:9125,661:2040).



Fig. 5. Pauli RGB image of the Flevoland big figure. The whole data of this image downloaded from the RADARSAT-2 web [83], and the area in box is often used. The area coordinate is (4061:6435,97:1731).

Netherlands, and the San Francisco bay area, California, USA. In the four PolSAR images, the first two images come from Canadian Space Agency RADARSAT-2 [83], as shown in the Fig. 4 and Fig. 5. The latter two are from NASA/JPL-Caltech AIRSAR[84]. We consider that this setup can fully test the performance of the algorithm over a variety of PolSAR images in terms of the system (AIRSAR and RADARSAT-2), the underlying classification problem and the operative band (C and L). The information about the four images is shown below.

1) San Francisco, RADARSAT-2, C-Band

The area around the bay of San Francisco with the golden gate bridge is probably one of the most used scenes in PolSAR image classification over the past decades. It provides a good coverage of both natural (e.g., water, vegetation) and man-made targets (e.g., high-density, low-density and developed).

This RADARSAT-2 fully PolSAR image at fine quad-pol mode (8-m spatial resolution) was taken in April 2, 2008. The selected scene is an 1380×1800 -pixel subregion. The Pauli-coded pseudo color image, the used ground truth data and the color code are shown in Fig. 6. In the ground truth map Fig. 6(b), there are five kinds of objects including developed, high-density, low-density, water and vegetation, which can be simply written as c1-c5. The size of original image is 2820×14416 , which is a relatively large scene with great research significance. Sub area coordinate is (7326:9125,661:2040), which was shown in Fig. 4. This coordinate can help researchers easily find the area and know the source of the data. Table I shows the numbers of the train and test samples.

2) Flevoland, Radarsat-2, C-Band

This RADARSAT-2 fully PolSAR image at one quad-pol mode (8-m spatial resolution) of Flevoland, the Netherlands, was taken in April 2, 2008. The selected scene is a 1635×2375 -pixel subregion, which mainly contains four terrain

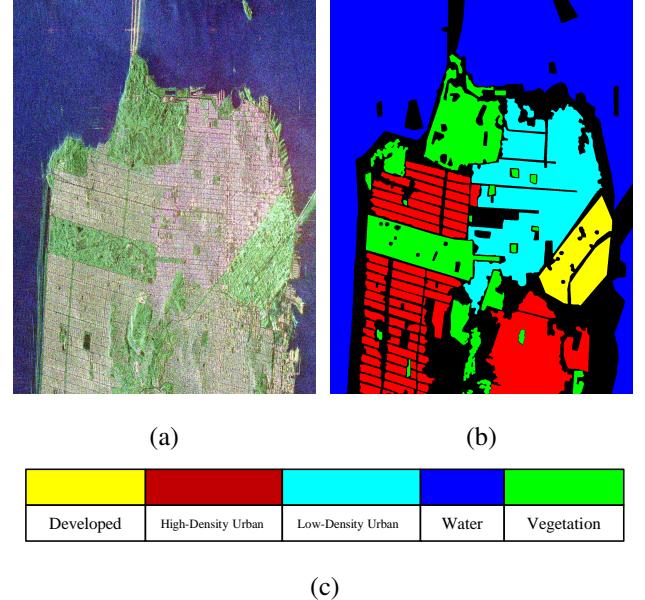


Fig. 6. San Francisco pseudo image and ground truth, Radarsat-2. (a) SanFrancisco image. (b) Ground truth image. (c) The color code.

classes: 1) woodland/forest; 2) cropland; 3) water; and 4) urban area. The Pauli color coded image, the ground truth data and the color code are shown in Fig. 7. The size of original image is 2820×12944 . Sub area coordinate is (4061:6435,97:1731), as was shown in Fig. 5. Table II shows the number of the train and test samples.

3) San Francisco, AIRSAR, L-Band

This PolSAR image of San Francisco bay has been used in many literature works, Fig. 8 show the Pauli RGB image,

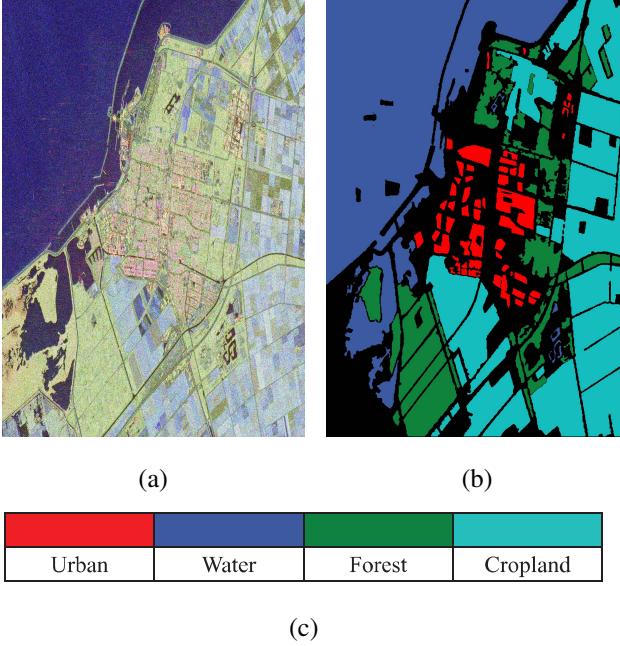


Fig. 7. Flevoland pseudo image and ground truth, Radarsat-2. (a) Flevoland image. (b) Ground truth image. (c) The color code.

the ground truth map and the color code. The size of this image is 900×1024 . The spatial resolution is 10 m for 20 MHz. Pixels in this image can be classified into five categories, abbreviated letter c1-c5 indicate the categories of mountain, ocean, urban, vegetation and bare soil, respectively. Table III shows the number of the train and test samples.

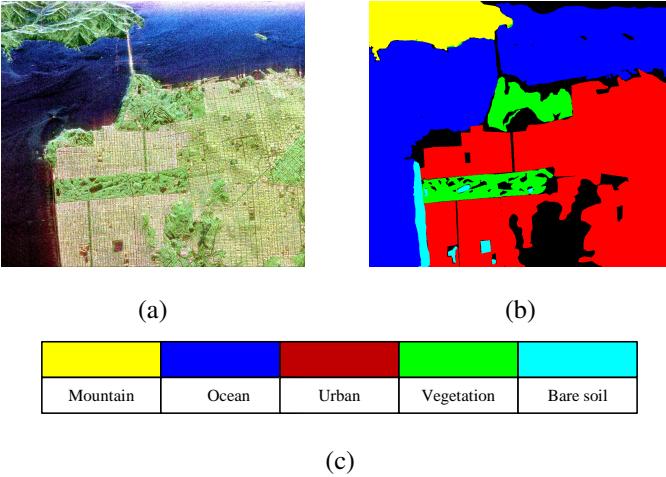


Fig. 8. Sanfrancisco pseudo image and ground truth, AIRSAR. (a) Sanfrancisco image. (b) Ground truth image. (c) The color code.

4) Flevoland, AIRSAR, L-Band

The PolSAR image of Flevoland is shown in Fig. 9(a), there are 15 categories in the ground truth map Fig. 9(b), and the color code is shown in Fig. 9(c). It is the picture with the most kinds of objects in the public PolSAR data collection at present. The spatial resolution is 10 m for 20 MHz. The size of this PolSAR data is 750×1024 . There are 15 kinds

of objects to be identified, including stem beans, rapeseed, bare soil, potatoes, beet, wheat2, peas, wheat3, lucerne, barley, wheat, grasses, forest, water and building. These objects are simply written as c1-c15 in this paper. The numbers of the train and test samples are shown in Table IV.

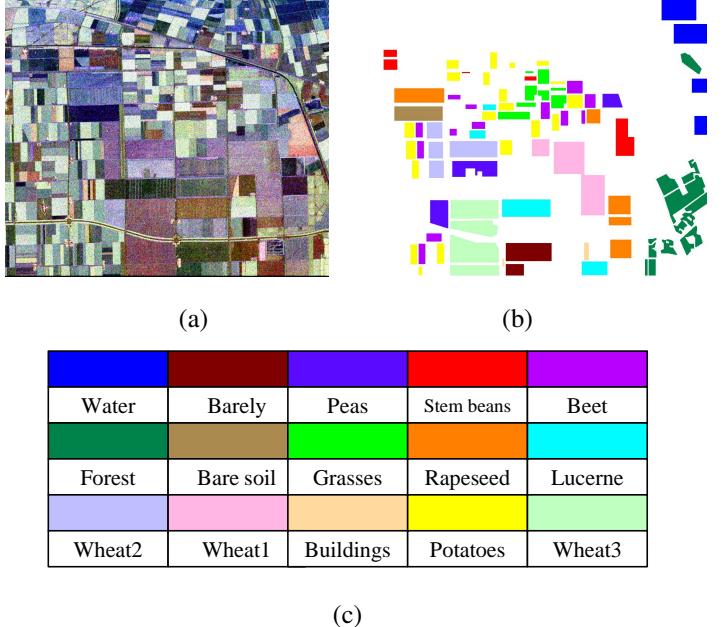


Fig. 9. Flevoland pseudo image and ground truth, AIRSAR. (a) Flevoland image. (b) Ground truth image. (c) The color code.

TABLE I
LAND CLASSES AND NUMBERS OF PIXELS IN THE FIRST DATA SET.

| class code | name | No. of training samples | No. of testing samples |
|------------|--------------------|-------------------------|------------------------|
| 1 | Water | 1000 | 852078 |
| 2 | Vegetation | 1000 | 237237 |
| 3 | High-Density Urban | 1000 | 351181 |
| 4 | Low-Density Urban | 1000 | 282975 |
| 5 | Developed | 1000 | 80616 |

TABLE II
LAND CLASSES AND NUMBERS OF PIXELS IN THE SECOND DATA SET.

| class code | name | No. of training samples | No. of testing samples |
|------------|----------|-------------------------|------------------------|
| 1 | Urban | 1000 | 136561 |
| 2 | Water | 1000 | 1050726 |
| 3 | Forest | 1000 | 403584 |
| 4 | Cropland | 1000 | 935821 |

TABLE III
LAND CLASSES AND NUMBERS OF PIXELS IN THE THIRD DATA SET.

| class code | name | No. of training samples | No. of testing samples |
|------------|------------|-------------------------|------------------------|
| 1 | Mountain | 1000 | 13701 |
| 2 | Ocean | 1000 | 62731 |
| 3 | Urban | 1000 | 329566 |
| 4 | Vegetation | 1000 | 342795 |
| 5 | Bare soil | 1000 | 53509 |

TABLE IV
LAND CLASSES AND NUMBERS OF PIXELS IN THE FOURTH DATA SET.

| class code | name | No. of training samples | No. of testing samples |
|------------|------------|-------------------------|------------------------|
| 1 | Water | 1000 | 12232 |
| 2 | Barely | 1000 | 6595 |
| 3 | Peas | 1000 | 8582 |
| 4 | Stem beans | 1000 | 5338 |
| 5 | Beet | 1000 | 9033 |
| 6 | Forest | 1000 | 17044 |
| 7 | Bare soil | 1000 | 4109 |
| 8 | Grasses | 1000 | 6058 |
| 9 | Rapeseed | 1000 | 12863 |
| 10 | Lucerne | 1000 | 9181 |
| 11 | Wheat2 | 1000 | 10159 |
| 12 | Wheat1 | 1000 | 15386 |
| 13 | Buildings | 200 | 535 |
| 14 | Potatoes | 1000 | 15156 |
| 15 | Wheat3 | 1000 | 21241 |

TABLE V
OVERVIEW OF POLARIMETRIC IMAGE DATA USED WITHIN EXPERIMENTS

| Name | System+Band | Abbr. | Date | Spatial Resolution | Dimensions | class |
|---------------|--------------|---------|----------|--------------------|-------------|-------|
| San Francisco | RADARSAT-2 C | SF-L | Apr 2008 | 8 (m) | 1380 × 1800 | 5 |
| Flevoland | RADARSAT-2 C | Flevo-C | Apr 2008 | 8 (m) | 1635 × 2375 | 4 |
| San Francisco | AIRSAR L | SF-L | Aug 1989 | 10 (m) | 1024 × 900 | 5 |
| Flevoland | AIRSAR L | Flevo-L | Aug 1989 | 10 (m) | 1024 × 750 | 15 |

B. Experimental setup

Traditional polarimetric features are extracted to compare with the polarimetric scattering coding in the proposed method and contrast algorithms. In this paper, we used a 22-dimensional feature vector as the traditional polarimetric feature, including the upper right element's absolute value of the 3×3 polarimetric coherency matrix \mathbf{T} , the upper right element's absolute value of the 3×3 polarimetric covariance matrix \mathbf{C} , three component of Pauli decomposition, three components of Freeman decomposition [14], and four components of Yamaguchi decomposition [8], expressed as PF22. The feature extracted by polarimetric scattering coding can be written as SSCF. In the experiment, we consider two aspects for comparison experiments: feature extraction and classifier design. For feature extraction, the commonly feature PF22 and the proposed SSCF are adopted. For classifier design, the traditional algorithms maximum likelihood (MLD) [85] and support vector machine (SVM) [86] are used to classify the image data, the deep learning algorithms also be adopted, such as convolutional neural network (CNN) [49], PFDCN [67] and full convolutional network (FCN) [87]. Based on these considerations, the comparison methods include PF22-MLD, PF22-SVM, PF22-CNN, SSC-CNN, PF22-FCN and PFDCN. The proposed method can be represented as PCN.

In contrast algorithms, in order to make as fair a comparison as possible, The standard deviation has been added in the result through ten random experiments, the parameters of CNN and FCN are set to the same as possible. The CNN is structured as follows, the first convolutional layer filters the input image patch with 64 kernels of size 5×5 . The second convolutional layer takes the output of the first convolutional layer as the input and filters it with 128 kernels of size 5×5 . The third, fourth, and fifth convolutional layers are connected to one another without any intervening pooling or normalization layers. The third convolutional layer has 256 kernels of size

5×5 connected to the (normalized, pooled) outputs of the second convolutional layer. The fourth convolutional layer has 256 kernels of size 5×5 , and the fifth convolutional layer has 128 kernels of size 5×5 . At last, The fully-connected layer has 100 neurons.

The structure of FCN is as follows, the relevant setting of the first five layers is the same as that of the CNN, the latter four layers is corresponding to the first five layers and the original image size. Skip layer also been adopted at the FCN's sixth to eighth layers to keep image edge information. In Fig. 3, the last feature maps of the two FCN are stacked. And then, the output layer then performs a 1×1 convolution to produce the same number of feature maps as the number of classes in each data set. In the case that the input is polarimetric scattering coding matrix, we added a specially designed two layer networks in front of CNN and FCN, the first layer is 32 kernels of size 4×4 with a stride of 1 pixel, the second layer is 64 kernels of size 4×4 with a stride of 2 pixels. Through this network, we can get the classification map with the same size as the original image. We train the model in a single step of optimization, and the weight is initialized by *Xavier* [88]. The stochastic gradient descent with momentum 0.9 is used to train the weights of the model. The initial learning rate is set to 0.01, the train batch size is 1. In the experiments, we randomly selected 1000 points per class for training, and the remaining labeled samples were used for testing. In addition to training data, other samples do not participate in training and learning. In SVM, the kernel function is the Radial Basis Function (RBF). The multiclass strategy is one-versus-rest. In the contrast algorithm, the size of image patch is 32×32 . In the proposed method, the training pixel is activated, others is set to 0 in the label, which is the ignore label in the training stage.

All the experiments are running on a HP Z840 workstation with an Intel Xeon(R) CPU, a GeForce GTX TITAN X GPU, and 64G RAM under Ubuntu 16.04 LTS. All of these methods are implemented using the deep learning framework of TensorFlow.

C. Evaluation metrics

For evaluating the classification performance, the experiment results were assessed by single class recall rate, overall accuracy (OA), average accuracy (AA) and kappa coefficient (Kappa). Overall accuracy can be defined as follows,

$$OA = \frac{M}{N} \quad (17)$$

where M is the number of classified correctly, and N is the total number of samples. Average accuracy can also be write as follows,

$$AA = \frac{1}{C} \sum_{i=1}^C \frac{M_i}{N_i} \quad (18)$$

where C is the number of categories, i category index, M_i is the number of correct samples for the i -th categories, N_i is

the number of samples in the i -th class. $\frac{M_i}{N_i}$ is the i -th class recall rate. The formula of kappa coefficient can be written as,

$$Kappa = \frac{OA - P}{1 - P}, P = \frac{1}{N^2} \sum_{i=1}^C \bar{Z}(i,:) * \bar{Z}(:,i) \quad (19)$$

where OA is the overall accuracy, Z is the confusion matrix, $\bar{Z}(i,:)$ is the sum of the i -th row elements, and $\bar{Z}(:,i)$ is the sum of the i -th column elements. N is the total number of samples.

IV. RESULT AND DISCUSSIONS

Our experimentation will be separated into three parts. In the first part, we demonstrate the proposed algorithm using the two data sets from the RADARSAT-2. In the second part, we evaluate the proposed algorithm using the other two data sets from the AIRSAR. In the third part, we give a significance analysis of the results. Finally, we show the computation times of the proposed algorithm and comparison algorithms in part four. The results are shown in the Fig. 10 - 15 and TABLE VI - IX.

A. Experimental with the dataset from the RADARSAT-2

The classification results from different algorithms are demonstrated in Fig. 10(a)-(g) and Fig. 11(a)-(g), and the accuracies for each class are listed in Table VI and Table VII, respectively. Fig. 14(a)-(b) show a clear contrast of classification accuracy, the trend line is generated by the data in Table VI and Table VII. It can be seen that the performance of the proposed method is better than others. The classification accuracies are higher than the contrast algorithms. The classification maps are closer to the ground truth. Trend chart of accuracy in Fig. 14(a)-(b) show that the PCN is better than others.

Fig. 10 shows that PF22-MLD and PF22-SVM cannot distinguish high-density urban well, and mistake much high-density urban for low-density urban, which is due to the fact that the two objects are similar, and the method is difficult to distinguish its characteristics. The result of FCN is better than the CNN's, which can be seen from the contiguous area of two class, such as coastal area. The classification map of PF22-CNN is worse than that of SSC-CNN in recognizing mixed terrain. Meanwhile, the result of PCN is better than that of PF22-FCN and PFDCN, the proposed algorithm PFDCN and PCN almost detects all terrain. Especially, the low-density urban misclassified in high-density urban is almost completely corrected. Table VI shows that the accuracy of deep learning methods are higher than the conventional methods, with a difference of two to six percentage points. It can be seen clearly that high-density urban is misclassified. The classification accuracy of high-density urban is 80.54%.

In Fig. 11, there are only four types of objects that need to be distinguished, namely urban, water, forest and cropland. Water is the most easily recognizable relatively, inland water can also be judged right. It is difficult to judge fake forest in urban and cropland. Some urban and cropland are mistakenly divided into forests in the classification maps obtained by

PF22-MLD and PF22-SVM. The classification methods based on convolution get the better classification maps, some homogeneous areas can be correctly obtained. Further more, the performance of the proposed method is more outstanding. The urban and cropland can be seen completely, there is almost no noise. Meanwhile, Table VII gives a quantitative result, PF22-MLD and PF22-SVM get a low accuracy, especially in urban and cropland areas, other methods show an upward trend. PCN has the highest accuracy.

TABLE VI
CLASSIFICATION ACCURACY (%) OF THE FIRST SCENE IMAGE

| Method | c1 | c2 | c3 | c4 | c5 | AA | OA | Kappa |
|----------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| PF22-MLD | 86.90 \pm 0.16 | 82.83 \pm 0.14 | 80.54 \pm 0.16 | 82.98 \pm 0.09 | 84.24 \pm 0.14 | 83.49 \pm 0.17 | 86.23 \pm 0.15 | 83.42 \pm 0.13 |
| PF22-SVM | 89.14 \pm 0.18 | 86.48 \pm 0.16 | 88.24 \pm 0.16 | 89.1 \pm 0.12 | 82.31 \pm 0.09 | 87.05 \pm 0.1 | 89.13 \pm 0.1 | 83.99 \pm 0.14 |
| PF22-CNN | 92.82 \pm 0.15 | 92.93 \pm 0.15 | 93.86 \pm 0.14 | 94.13 \pm 0.13 | 85.93 \pm 0.14 | 91.93 \pm 0.11 | 93.05 \pm 0.07 | 88.54 \pm 0.14 |
| SSC-CNN | 86.27 \pm 0.09 | 94.07 \pm 0.09 | 94.56 \pm 0.08 | 93.54 \pm 0.15 | 87.24 \pm 0.12 | 91.14 \pm 0.12 | 92.14 \pm 0.08 | 89.29 \pm 0.08 |
| PF22-FCN | 93.46 \pm 0.06 | 97.39 \pm 0.13 | 92.07 \pm 0.09 | 96.43 \pm 0.14 | 92.42 \pm 0.12 | 94.35 \pm 0.11 | 95.05 \pm 0.14 | 90.54 \pm 0.09 |
| PFDCN | 94.68 \pm 0.12 | 97.64 \pm 0.07 | 92.54 \pm 0.06 | 96.92 \pm 0.09 | 92.82 \pm 0.1 | 95.88 \pm 0.08 | 96.05 \pm 0.12 | 93.54 \pm 0.15 |
| PCN | 96.2 \pm 0.15 | 94.83 \pm 0.15 | 98.56 \pm 0.14 | 98.5 \pm 0.12 | 98.79 \pm 0.08 | 97.44 \pm 0.14 | 98.24 \pm 0.07 | 95.27 \pm 0.1 |

TABLE VII
CLASSIFICATION ACCURACY (%) OF THE SECOND SCENE IMAGE

| Method | c1 | c2 | c3 | c4 | AA | OA | Kappa |
|----------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| PF22-MLD | 88.63 \pm 0.07 | 81.91 \pm 0.16 | 89.90 \pm 0.18 | 86.46 \pm 0.14 | 86.71 \pm 0.12 | 87.98 \pm 0.06 | 85.36 \pm 0.08 |
| PF22-SVM | 86.47 \pm 0.09 | 86.4 \pm 0.12 | 85.16 \pm 0.08 | 88.62 \pm 0.17 | 86.66 \pm 0.16 | 88.42 \pm 0.08 | 84.09 \pm 0.07 |
| PF22-CNN | 90.95 \pm 0.17 | 85.12 \pm 0.06 | 94.39 \pm 0.06 | 87.05 \pm 0.13 | 89.38 \pm 0.15 | 90.09 \pm 0.09 | 88.88 \pm 0.08 |
| SSC-CNN | 89.65 \pm 0.12 | 91.95 \pm 0.09 | 92.79 \pm 0.12 | 91.45 \pm 0.16 | 91.45 \pm 0.08 | 93.4 \pm 0.08 | 89.68 \pm 0.08 |
| PF22-FCN | 91.14 \pm 0.10 | 95.64 \pm 0.08 | 93.12 \pm 0.11 | 98.52 \pm 0.11 | 94.6 \pm 0.08 | 95.38 \pm 0.09 | 91.95 \pm 0.07 |
| PF22-FCN | 94.33 \pm 0.10 | 96.08 \pm 0.16 | 94.16 \pm 0.15 | 97.66 \pm 0.13 | 95.6 \pm 0.13 | 96.21 \pm 0.14 | 93.54 \pm 0.06 |
| PCN | 98.86 \pm 0.11 | 98.9 \pm 0.13 | 99.31 \pm 0.13 | 95.76 \pm 0.08 | 98.23 \pm 0.08 | 98.41 \pm 0.08 | 95.84 \pm 0.13 |

TABLE VIII
CLASSIFICATION ACCURACY (%) OF THE THIRD SCENE IMAGE

| Method | c1 | c2 | c3 | c4 | c5 | AA | OA | Kappa |
|----------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| PF22-MLD | 88.27 \pm 0.11 | 86.83 \pm 0.22 | 85.52 \pm 0.18 | 89.93 \pm 0.13 | 86.97 \pm 0.06 | 87.50 \pm 0.09 | 88.64 \pm 0.11 | 84.91 \pm 0.07 |
| PF22-SVM | 84.75 \pm 0.17 | 89.73 \pm 0.11 | 87.94 \pm 0.15 | 82.79 \pm 0.14 | 90.98 \pm 0.14 | 87.24 \pm 0.11 | 88.35 \pm 0.14 | 87.01 \pm 0.09 |
| PF22-CNN | 91.32 \pm 0.12 | 92.10 \pm 0.11 | 91.74 \pm 0.09 | 86.56 \pm 0.15 | 94.08 \pm 0.12 | 91.16 \pm 0.12 | 91.87 \pm 0.13 | 89.81 \pm 0.10 |
| SSC-CNN | 91.27 \pm 0.15 | 94.01 \pm 0.15 | 90.90 \pm 0.09 | 93.37 \pm 0.12 | 91.59 \pm 0.14 | 92.22 \pm 0.07 | 92.74 \pm 0.06 | 91.12 \pm 0.06 |
| PF22-FCN | 96.45 \pm 0.07 | 94.39 \pm 0.06 | 91.22 \pm 0.13 | 94.28 \pm 0.16 | 94.09 \pm 0.08 | 94.08 \pm 0.06 | 95.22 \pm 0.21 | 91.89 \pm 0.08 |
| PFDCN | 97.08 \pm 0.09 | 95.04 \pm 0.10 | 91.46 \pm 0.09 | 95.21 \pm 0.11 | 94.89 \pm 0.14 | 94.88 \pm 0.08 | 96.21 \pm 0.09 | 92.34 \pm 0.08 |
| PCN | 93.54 \pm 0.07 | 94.23 \pm 0.08 | 97.19 \pm 0.09 | 97.59 \pm 0.11 | 96.45 \pm 0.10 | 95.82 \pm 0.10 | 97.73 \pm 0.08 | 94.65 \pm 0.07 |

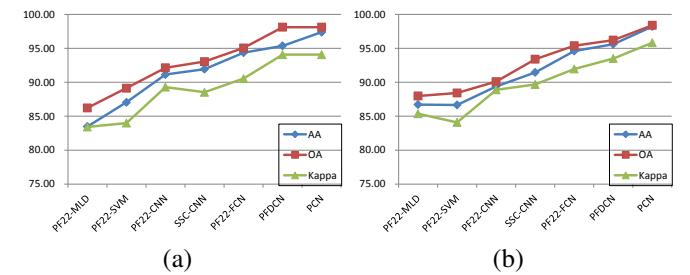


Fig. 14. Classification accuracies on Fig. 6 and Fig. 7 by different methods.
(a) Classification accuracies on Fig. 6. (b) Classification accuracies on Fig. 7.

B. Experimental with the dataset from the AIRSAR

In this section, Fig. 12(a)-(g) and Fig. 13(a)-(g) show the classification maps from the proposed method and the contrast algorithms. Table VIII and Table IX list the statistical classification accuracy from the proposed method and the compared algorithms. A clear exhibition of classification accuracy is shown in Fig 15(a)-(b).

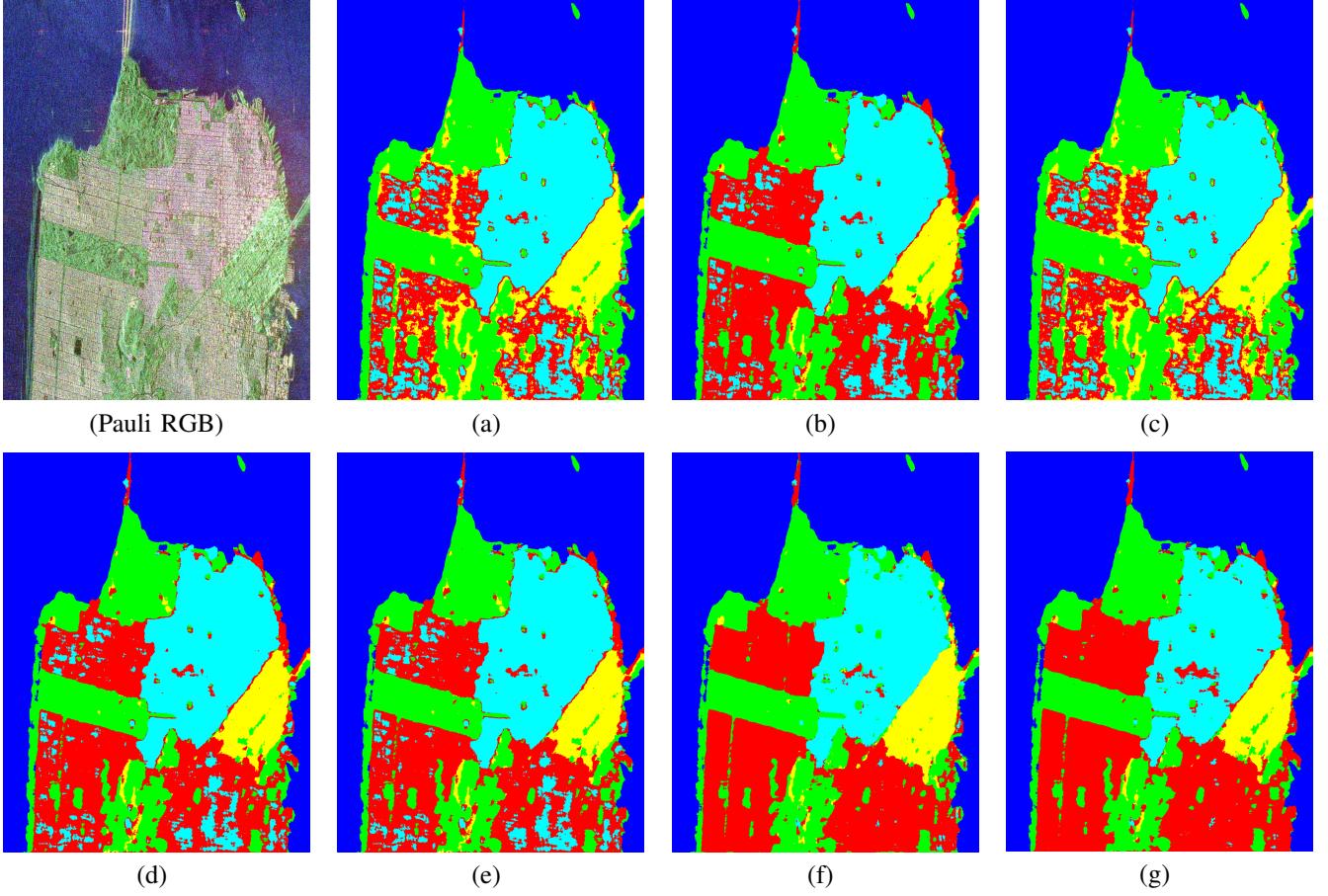


Fig. 10. The classification maps of different methods on the first scene image. (a)-(g) : PF22-MLD, PF22-SVM, PF22-CNN, SSC-CNN, PF22-FCN, PFDCN and PCN

TABLE IX
CLASSIFICATION ACCURACY (%) OF THE FOURTH SCENE IMAGE

| Method | c1 | c2 | c3 | c4 | c5 | c6 | c7 | c8 | c9 | c10 | c11 | c12 | c13 | c14 | c15 | AA | OA | Kappa |
|----------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|----------------------------------|----------------------------------|----------------------------------|
| PF22-MLD | 88.32 \pm 0.11 | 88.61 \pm 0.11 | 84.61 \pm 0.08 | 85.31 \pm 0.14 | 83.83 \pm 0.12 | 82.58 \pm 0.07 | 86.08 \pm 0.12 | 85.86 \pm 0.15 | 82.29 \pm 0.12 | 81.97 \pm 0.12 | 88.29 \pm 0.15 | 84.61 \pm 0.13 | 88.52 \pm 0.13 | 88.81 \pm 0.09 | 85.06 \pm 0.11 | 85.65 \pm 0.10 | 87.46 \pm 0.15 | 83.85 \pm 0.06 |
| PF22-SVM | 82.47 \pm 0.11 | 83.89 \pm 0.08 | 88.37 \pm 0.09 | 89.20 \pm 0.11 | 83.13 \pm 0.11 | 82.7 \pm 0.09 | 82.08 \pm 0.08 | 87.49 \pm 0.10 | 86.24 \pm 0.10 | 86.78 \pm 0.13 | 89.57 \pm 0.11 | 81.15 \pm 0.08 | 88.15 \pm 0.11 | 87.3 \pm 0.15 | 82.81 \pm 0.06 | 85.42 \pm 0.16 | 85.8 \pm 0.07 | 83.04 \pm 0.16 |
| PF22-CNN | 86.45 \pm 0.14 | 86.19 \pm 0.13 | 85.02 \pm 0.12 | 88.82 \pm 0.11 | 92.62 \pm 0.12 | 94.72 \pm 0.11 | 93.78 \pm 0.13 | 89.53 \pm 0.11 | 89.98 \pm 0.14 | 92.52 \pm 0.08 | 85.56 \pm 0.13 | 93.41 \pm 0.10 | 93.36 \pm 0.11 | 87.98 \pm 0.10 | 88.22 \pm 0.14 | 89.88 \pm 0.08 | 91.71 \pm 0.07 | 87.24 \pm 0.06 |
| SSC-CNN | 89.05 \pm 0.08 | 87.07 \pm 0.09 | 88.33 \pm 0.07 | 89.39 \pm 0.13 | 93.21 \pm 0.07 | 95.87 \pm 0.11 | 89.36 \pm 0.11 | 95.94 \pm 0.10 | 92.83 \pm 0.08 | 92.64 \pm 0.08 | 92.41 \pm 0.14 | 87.04 \pm 0.11 | 92.23 \pm 0.06 | 93.52 \pm 0.07 | 94.63 \pm 0.13 | 91.56 \pm 0.10 | 91.78 \pm 0.11 | 90.38 \pm 0.07 |
| PF22-FCN | 97.97 \pm 0.13 | 93.78 \pm 0.09 | 92.78 \pm 0.14 | 95.84 \pm 0.10 | 96.25 \pm 0.07 | 92.41 \pm 0.08 | 97.71 \pm 0.08 | 93.60 \pm 0.09 | 90.25 \pm 0.12 | 90.10 \pm 0.16 | 92.19 \pm 0.14 | 93.37 \pm 0.10 | 93.76 \pm 0.12 | 91.31 \pm 0.10 | 90.66 \pm 0.12 | 94.65 \pm 0.16 | 96.34 \pm 0.11 | 93.64 \pm 0.14 |
| PFDCN | 98.33 \pm 0.10 | 94.32 \pm 0.12 | 93.55 \pm 0.11 | 96.34 \pm 0.09 | 96.55 \pm 0.10 | 93.87 \pm 0.13 | 98.51 \pm 0.16 | 97.96 \pm 0.13 | 92.54 \pm 0.10 | 91.13 \pm 0.08 | 97.35 \pm 0.12 | 94.54 \pm 0.08 | 94.05 \pm 0.08 | 98.06 \pm 0.13 | 91.33 \pm 0.07 | 95.02 \pm 0.13 | 96.67 \pm 0.07 | 93.88 \pm 0.12 |
| PCN | 99.50 \pm 0.15 | 94.51 \pm 0.09 | 95.29 \pm 0.10 | 96.44 \pm 0.15 | 97.59 \pm 0.10 | 95.05 \pm 0.11 | 96.45 \pm 0.10 | 94.33 \pm 0.08 | 93.98 \pm 0.07 | 93.51 \pm 0.16 | 95.03 \pm 0.10 | 95.2 \pm 0.10 | 95.58 \pm 0.13 | 95.89 \pm 0.08 | 95.43 \pm 0.08 | 95.59\pm0.11 | 96.94\pm0.13 | 94.08\pm0.07 |

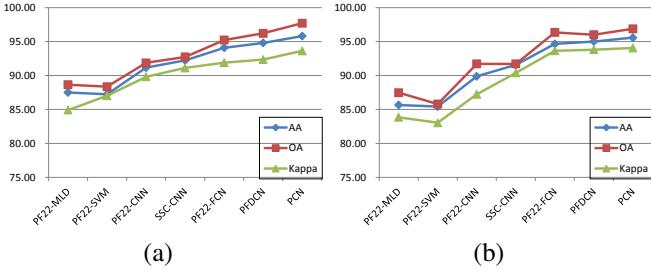


Fig. 15. Classification accuracies on Fig. 8 and Fig. 9 by different methods.
(a) Classification accuracies on Fig. 8. (b) Classification accuracies on Fig. 9.

From the above results, we can get that the proposed method have the best performances. The misclassified pixels are much less in the classification maps. The classification accuracy are

95.82 and 95.59, respectively. It is 2-8 percentage points higher than other algorithms.

In Fig. 12, there are five kinds of objects to be determined, including mountain, ocean, urban, vegetation, bare soil. There is an island on the upper right of the image, called Alcatraz Island, it did not appear in Fig 12(a). On contrast, the island was distinguished as a mountain in Fig 12(b)-(e). Although the island has not been marked, it is still detected. Fig 12(g) shows a clearer outline. What's more, in the bare soil detection task, the difference of the algorithm performance can also be shown. The shape and area of the soil can be described more accurately, which can be seen in Fig 12(g). The accuracy are listed in Table VIII, the overall accuracy of PF22-MLD and PF22-SVM are only 87.50% and 87.24%, other algorithms are larger than 0.90 and the proposed algorithm is as high as 95.82.

Fig. 13 shows the classification results of the fourth data

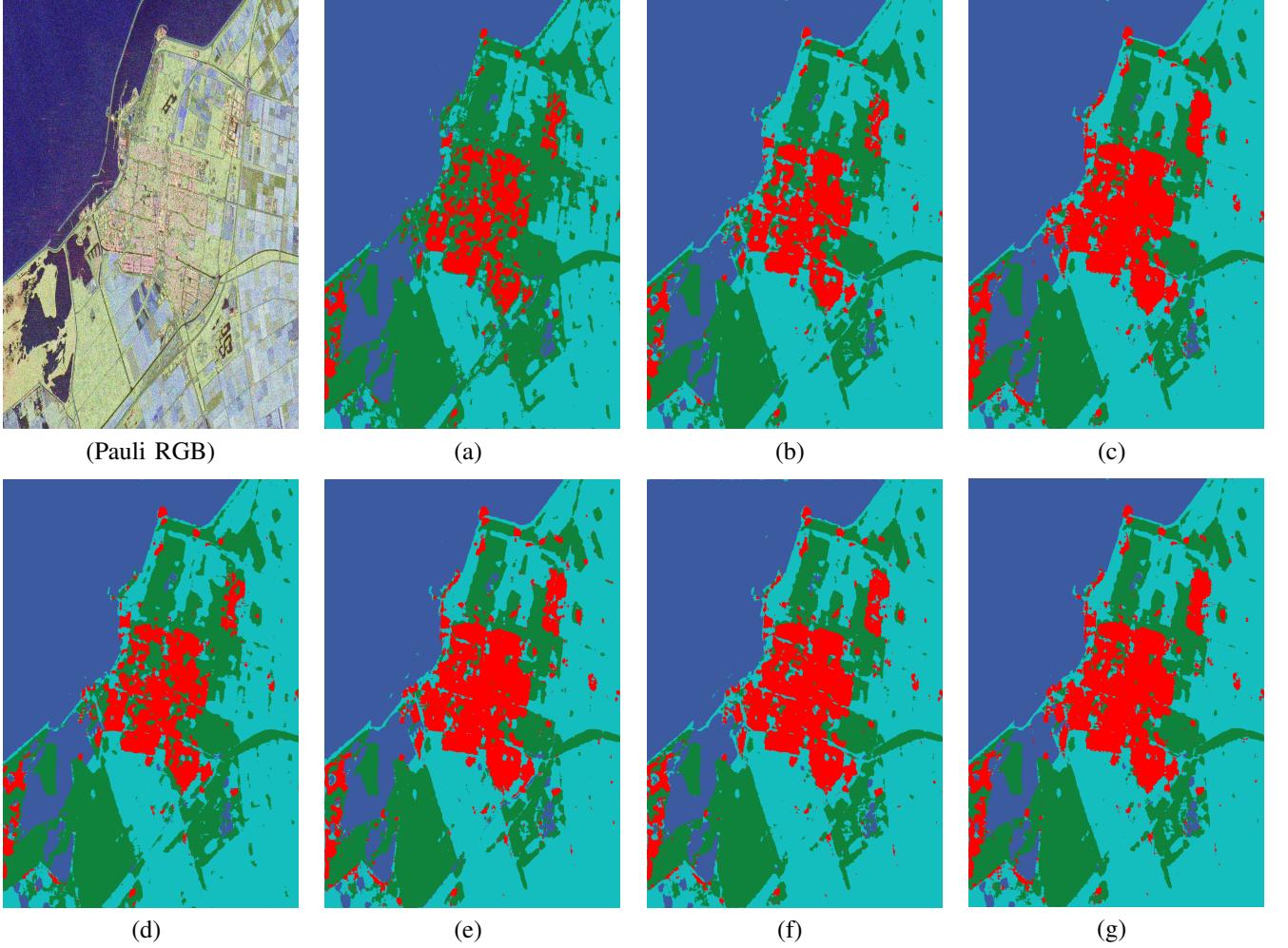


Fig. 11. The classification maps of different methods on the second scene image. (a)-(g) : PF22-MLD, PF22-SVM, PF22-CNN, SSC-CNN, PF22-FCN, PFDCN and PCN

set. This data set is shown in Fig. 9 and contains 15 kinds of objects, so many kinds of labeled data are extremely rare. Naturally, the problem will increase in difficulty. For instance, it is difficult to distinguish wheat1 wheat2 and wheat3, some wheat3 was wrongly divided into wheat1 and wheat2 in Fig. 13(a)-(c), Fig. 13(a) is the largest, Fig. 13(b) and Fig. 13(c) are relatively few. The proposed algorithm can judge almost all pixels correctly, which is shown in Fig. 13(g). Similarly, some potatoes are wrongly classified as peas, but the proposed method can solve this problem. In Table IX, the above experimental phenomena can be seen accurately through numbers, such as the classification accuracy of wheat1, wheat2 and wheat3 get a 5 percentage point increase.

We can see that the proposed approach outperforms the compared methods. It indicates that the encoded data through polarimetric scattering coding is easier to be identified and distinguished. At the same time, we can find that the modified fully convolutional network has a better classification performance than conventional convolutional network.

C. Significance analysis

In this section, we give a significance analysis about the experiment results by T-test score. The analysis results are shown in TABLE X. In the table, low values are good and significant. If the value is less than 0.05, it means the result is significant. It can be seen that the proposed approach has a significant advantage.

TABLE X
THE T-TEST SCORES BETWEEN PCN AND THE COMPARED METHODS.

| Method | AA | OA | Kappa |
|----------|--------|--------|--------|
| PF22-MLD | 0.0028 | 0.0006 | 0.0002 |
| PF22-SVM | 0.0005 | 0.0207 | 0.0016 |
| PF22-CNN | 0.0067 | 0.0036 | 0.0011 |
| SSC-CNN | 0.0074 | 0.0002 | 0.0065 |
| PF22-FCN | 0.0316 | 0.0296 | 0.0502 |
| PFDCN | 0.0508 | 0.0423 | 0.0464 |

¹ The significant value is less than 0.05.

D. Computation times

In Table XI, we show the computation times achieved by different methods on the data sets. Factors that affect the

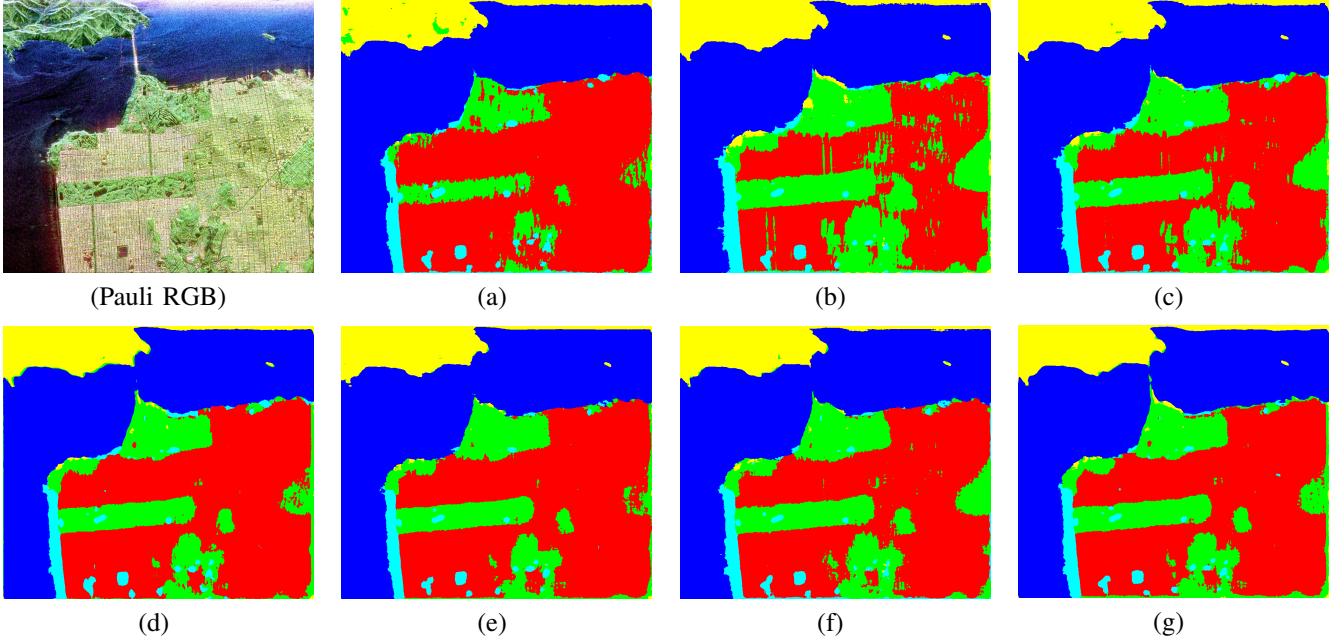


Fig. 12. The classification maps of different methods on the third scene image. (a)-(g) : PF22-MLD, PF22-SVM, PF22-CNN, SSC-CNN, PF22-FCN PFDCN and PCN

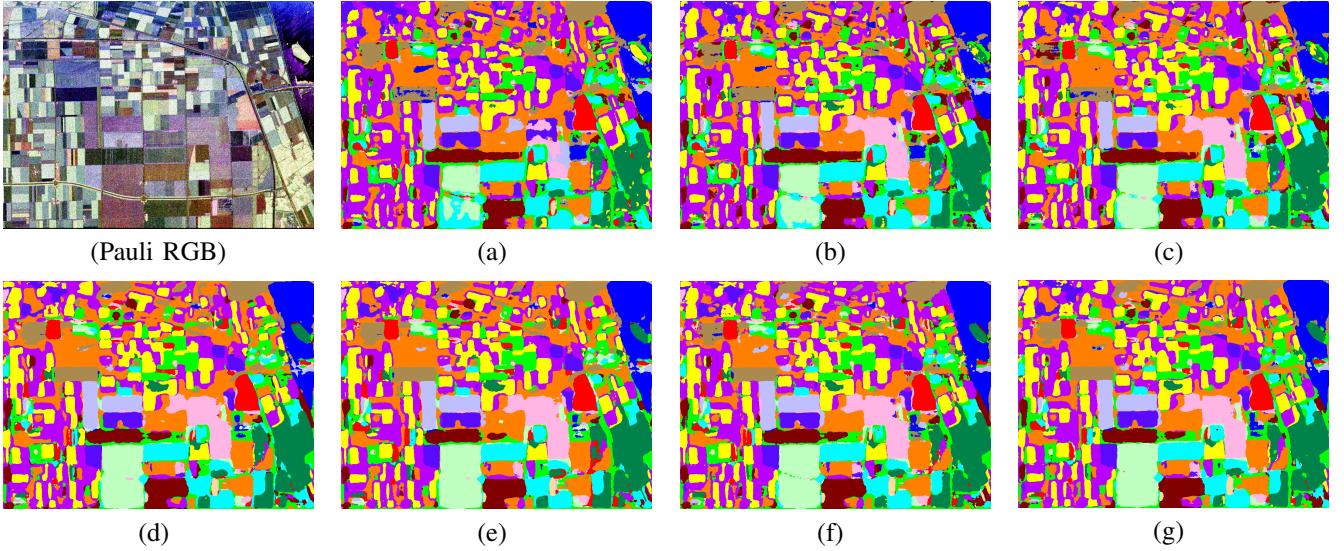


Fig. 13. The classification maps of different methods on the fourth scene image. (a)-(g) : PF22-MLD, PF22-SVM, PF22-CNN, SSC-CNN, PF22-FCN, PFDCN and PCN

computation times include the size of image, the complexity of the data sets, and the methods. The computation times then increase with the size of image and the complexity of the data sets. When compared Fig. 6 and Fig. 8, the size of image is import factor that impacts the computational efficiency. From 7 and Fig. 9, it can easily be seen that the complexity of the data sets is an import factor for the computational efficiency. Since there are fifteen categories in Fig. 9. At last, the speed of PF22-FCN and PCN is faster than others, the main reason is that this algorithm does not need sliding window to calculate the pixels.

TABLE XI
COMPUTATION TIMES WITH DIFFERENT METHODS ON DATA SETS.

| Methods | Data Sets | | Fig. 6 | | Fig. 7 | | Fig. 8 | | Fig. 9 | |
|----------|-------------|------------|-------------|------------|-------------|------------|-------------|------------|-------------|------------|
| | training(s) | testing(s) |
| PF22-MLD | 380 | 131 | 355 | 118 | 435 | 155 | 335 | 117 | | |
| PF22-SVM | 432 | 156 | 420 | 137 | 487 | 178 | 430 | 125 | | |
| PF22-CNN | 300 | 60 | 289 | 57 | 354 | 86 | 300 | 55 | | |
| SSC-CNN | 283 | 56 | 276 | 54 | 348 | 75 | 289 | 52 | | |
| PF22-FCN | 321 | 40 | 315 | 38 | 361 | 55 | 325 | 33 | | |
| PFDCN | 285 | 58 | 284 | 54 | 351 | 76 | 279 | 51 | | |
| PCN | 350 | 42 | 328 | 40 | 383 | 57 | 341 | 35 | | |

V. CONCLUSIONS

In this paper, we proposed a new PolSAR image classification method named polarimetric convolutional network, which

based on polarimetric scattering coding and fully convolutional network. polarimetric scattering coding can keep the structure information of scattering matrix, and avoid breaking the matrix into a one-dimensional vector.

Coincidentally, convolution network needs a 2-D input, where polarimetric scattering coding matrix meets this condition. We design an improved full convolutional network to classify data encoded by polarimetric scattering coding. In order to make the experiment more fully and effectively, the experimental data sets consist of four data sets from two satellites, the contrast algorithms include traditional methods and latest methods. Experimental results show that the proposed algorithm PCN is robust and get a better results, the classification maps of the proposed method are very close to the ground truth maps, and the classification accuracies are higher than contrast algorithms. These results are mainly due to the fact that the proposed algorithm can preserve the structural semantic information of the image in the raw data. In contrast algorithms, PF22-MLD, PF22-SVM and PF22-CNN do not give full consideration to the structural semantic information, SSC-CNN only thinks about the structural semantic information of the raw data, and PF22-FCN only considers the structural semantic information of model. The results of PFDCN outperform other comparison algorithms. Experimental results can also confirm the above inference. From comparative experiments, we also found that this polarimetric scattering coding is effective. For this coding, the performance of designed classification network is better.

ACKNOWLEDGMENT

The authors would like to thank the NASA/JPL-Caltech and Canadian Space Agency for kindly providing the polarimetric AIRSAR/RADARSAT2 data used in this paper. The authors would also like to thank the anonymous reviewers for their helpful comments.

REFERENCES

- [1] H. Wang, R. Magagi, and K. Goita, "Comparison of different polarimetric decompositions for soil moisture retrieval over vegetation covered agricultural area," *Remote Sens. Environ.*, vol. 199, pp. 120–136, 2017.
- [2] K. Voormansik, T. Jagdhuber, K. Zalite, M. Noorma, and I. Hajnsek, "Observations of cutting practices in agricultural grasslands using polarimetric sar," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 9, no. 4, pp. 1382–1396, 2016.
- [3] F. Liu, J. Shi, L. Jiao, H. Liu, S. Yang, J. Wu, H. Hao, and J. Yuan, "Hierarchical semantic model and scattering mechanism based polsar image classification," *Pattern Recognition*, vol. 59, pp. 325–342, 2016.
- [4] C. Dickinson, P. Siqueira, D. Clewley, and R. Lucas, "Classification of forest composition using polarimetric decomposition in multiple landscapes," *Remote Sens. Environ.*, vol. 131, pp. 206–214, 2013.
- [5] E. Krogager, "New decomposition of the radar target scattering matrix," *Electronics Letters*, vol. 26, no. 18, pp. 1525–1527, 1990.
- [6] R. Touzi and F. Charbonneau, "Characterization of target symmetric scattering using polarimetric sars," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2507–2516, 2002.
- [7] W. L. Cameron and L. K. Leung, "Feature motivated polarization scattering matrix decomposition," in *Record of the IEEE 1990 International Radar Conference*, 1990, pp. 549–557.
- [8] Y. Yamaguchi, A. Sato, W. M. Boerner, R. Sato, and H. Yamada, "Four-component scattering power decomposition with rotation of coherency matrix," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2251–2258, 2011.
- [9] S. W. Chen, X. S. Wang, S. P. Xiao, and M. Sato, "General polarimetric model-based decomposition for coherency matrix," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1843–1855, 2013.
- [10] S. W. Chen, X. S. Wang, Y. Z. Li, and M. Sato, "Adaptive model-based polarimetric decomposition using polinsar coherence," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1705–1718, 2013.
- [11] S.-W. Chen, X.-S. Wang, S.-P. Xiao, and M. Sato, "Advanced polarimetric target decomposition," in *Target Scattering Mechanism in Polarimetric Synthetic Aperture Radar*. Springer, Singapore, 2018, pp. 43–106.
- [12] M. Mahdianpari, B. Salehi, F. Mohammadimanesh, B. Brisco, S. Mahdavi, M. Amani, and J. E. Granger, "Fisher linear discriminant analysis of coherency matrix for wetland classification using polsar imagery," *Remote Sens. of Environment*, vol. 206, no. 1 March 2018, pp. 300–317, 2018.
- [13] J. R. Huynen, "Phenomenological theory of radar targets," *Electromagnetic Scattering*, pp. 653–712, 1978.
- [14] A. Freeman and S. L. Durden, "Three-component scattering model to describe polarimetric SAR data," in *Radar Polarimetry*, vol. 1748, 1993, pp. 213–225.
- [15] Y. Yamaguchi, T. Moriyama, M. Ishido, and H. Yamada, "Four-component scattering model for polarimetric SAR image decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 8, pp. 1699–1706, 2005.
- [16] S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 498–518, 1996.
- [17] S. R. Cloude and Pottier, "An entropy based classification scheme for land applications of polarimetric sar," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 68–78, 1997.
- [18] J. S. Lee, T. L. Ainsworth, and Y. Wang, "Generalized polarimetric model-based decompositions using incoherent scattering models," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2474–2491, 2014.
- [19] N. Besic, G. Vasile, J. Chanussot, and S. Stankovic, "Polarimetric incoherent target decomposition by means of independent component analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1236–1247, 2014.
- [20] H. Aghababaee and M. R. Sahebi, "Incoherent target scattering decomposition of polarimetric SAR data based on vector model roll-invariant parameters," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4392–4401, 2016.
- [21] S. W. Chen, Y. Z. Li, X. S. Wang, and S. P. Xiao, "Modeling and interpretation of scattering mechanisms in polarimetric synthetic aperture radar: Advances and perspectives," *IEEE Signal Processing Mag.*, vol. 31, no. 4, pp. 79–89, 2014.
- [22] S. W. Chen, X. S. Wang, and M. Sato, "Uniform polarimetric matrix rotation theory and its applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4756–4770, 2014.
- [23] F. Xu, Q. Song, and Y. Q. Jin, "Polarimetric SAR image factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. PP, no. 99, pp. 1–16, 2017.
- [24] C. Tao, S. Chen, Y. Li, and S. Xiao, "Polsar land cover classification based on roll-invariant and selected hidden polarimetric features in the rotation domain," *Remote Sens.*, vol. 9, no. 7, 2017.
- [25] S. Uhlmann and S. Kiranyaz, "Integrating color features in polarimetric SAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 4, pp. 2197–2216, 2014.
- [26] G. D. D. Grandi, J. S. Lee, and D. L. Schuler, "Target detection and texture segmentation in polarimetric SAR images using a wavelet frame: Theoretical aspects," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3437–3453, 2007.
- [27] X. Ma, H. Shen, J. Yang, L. Zhang, and P. Li, "Polarimetric-spatial classification of SAR images based on the fusion of multiple classifiers," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, no. 3, pp. 961–971, 2014.
- [28] T. Zou, W. Yang, D. Dai, and H. Sun, "Polarimetric SAR image classification using multifeatures combination and extremely randomized clustering forests," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, p. 4, 2010.
- [29] B. Ren, B. Hou, J. Zhao, and L. Jiao, "Unsupervised classification of polarimetric SAR image via improved manifold regularized low-rank representation with multiple features," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 10, no. 2, pp. 580–595, 2017.
- [30] H. Liu, D. Zhu, S. Yang, B. Hou, S. Gou, T. Xiong, and L. Jiao, "Semisupervised feature extraction with neighborhood constraints for polarimetric SAR classification," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 9, no. 7, pp. 3001–3015, 2016.

- [31] J. S. Lee, M. R. Grunes, T. L. Ainsworth, L. J. Du, D. L. Schuler, and S. R. Cloude, "Unsupervised classification using polarimetric decomposition and the complex wishart classifier," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2249–2258, 2002.
- [32] J. S. Lee, M. R. Grunes, E. Pottier, and L. Ferro-Famil, "Unsupervised terrain classification preserving polarimetric scattering characteristics," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 4, pp. 722–731, 2004.
- [33] L. DU and J. S. LEE, "Fuzzy classification of earth terrain covers using complex polarimetric SAR data," *International Journal of Remote Sens.*, vol. 17, no. 4, pp. 809–826, 1996.
- [34] P. R. Kersten, J. S. Lee, and T. L. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and em clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 519–527, 2005.
- [35] F. Liu, L. Jiao, B. Hou, and S. Yang, "Pol-sar image classification based on wishart dbn and local spatial information," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3292–3308, 2016.
- [36] L. Jiao and F. Liu, "Wishart deep stacking network for fast polsar image classification," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3273–3286, 2016.
- [37] L. Jiao, J. Zhao, S. Yang, and F. Liu, *Deep Learning, Optimization and Recognition*. Tsinghua University Press, 2017.
- [38] H. Liu, Y. Wang, S. Yang, S. Wang, J. Feng, and L. Jiao, "Large polarimetric SAR data semi-supervised classification with spatial-anchor graph," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 9, no. 4, pp. 1439–1458, 2016.
- [39] B. Wei, J. Yu, C. Wang, H. Wu, and J. Li, "Polsar image classification using a semi-supervised classifier based on hypergraph learning," *Remote Sens. Letters*, vol. 5, no. 4, pp. 386–395, 2014.
- [40] H. Liu, X. Xing, S. Wang, Z. Feng, E. Zhang, S. Yang, B. Hou, and L. Jiao, "Fast semi-supervised classification based on parallel auction graph for polarimetric SAR data," in *IGARSS*, 2016, pp. 1528–1531.
- [41] O. Harant, L. Bombrun, G. Vasile, M. Gay, L. Ferro-Famil, R. Fallourd, E. Trouvé, J.-M. Nicolas, and F. Tupin, "Fisher PDF formaximum likelihood texture tracking with high resolution polsar data," in *8th European Conference on Synthetic Aperture Radar*, 2010, pp. 1–4.
- [42] S. Fukuda and H. Hirosawa, "Polarimetric SAR image classification using support vector machines," *IEICE Trans. Electronics*, vol. 84, no. 12, pp. 1939–1945, 2001.
- [43] H. Li, H. Gu, Y. Han, and J. Yang, "Object-oriented classification of polarimetric SAR imagery based on statistical region merging and support vector machine," in *Earth Observation and Remote Sens. Applications, 2008. EORSA 2008. International Workshop on*. IEEE, 2008, pp. 1–6.
- [44] H. Aghababaei, "Contextual polsar image classification using fractal dimension and support vector machines," *European Journal of Remote Sens.*, vol. 46, no. 1, pp. 317–332, 2013.
- [45] L. Zhang, L. Sun, B. Zou, and W. M. Moon, "Fully polarimetric SAR image classification via sparse representation and polarimetric features," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 8, no. 8, pp. 3923–3932, 2015.
- [46] H. Li, Y. He, and W. Wang, "Improving ship detection with polarimetric SAR based on convolution between co-polarization channels," *Sensors*, vol. 9, no. 2, pp. 1221–1236, 2009.
- [47] Y. Zhou, H. Wang, F. Xu, and Y. Q. Jin, "Polarimetric SAR image classification using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1935–1939, 2016.
- [48] H. Liu, S. Yang, S. Gou, D. Zhu, R. Wang, and L. Jiao, "Polarimetric SAR feature extraction with neighborhood preservation-based deep learning," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 10, no. 4, pp. 1456–1466, 2017.
- [49] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *NIPS*, vol. 25, no. 2, 2012, pp. 1097–1105.
- [50] Y. Bengio et al., "Learning deep architectures for ai," *Foundations and trends® in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009.
- [51] W. Yang, X. Yin, and G. S. Xia, "Learning high-level features for satellite image classification with limited labeled samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4472–4482, 2015.
- [52] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Tensor discriminative locality alignment for hyperspectral image hyperspectral feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 242–256, 2013.
- [53] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 7, no. 6, pp. 2094–2107, 2014.
- [54] J. Han, D. Zhang, G. Cheng, L. Guo, and J. Ren, "Object detection in optical remote sens. images based on weakly supervised learning and high-level feature learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3325–3337, 2015.
- [55] J. Yu, X. Yang, F. Gao, and D. Tao, "Deep multimodal distance metric learning using click constraints for image ranking," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4014–4024, Dec 2017.
- [56] M. Gong, J. Zhao, J. Liu, Q. Miao, and L. Jiao, "Change detection in synthetic aperture radar images based on deep neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 1, pp. 125–138, 2016.
- [57] W. Zhao and S. Du, "Spectralspatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 1–11, 2016.
- [58] H. Petersson, D. Gustafsson, and D. Bergstrom, "Hyperspectral image analysis using deep learninga review," in *Image Processing Theory Tools and Applications (IPTA), 2016 6th International Conference on*. IEEE, 2016, pp. 1–6.
- [59] S. Yu, S. Jia, and C. Xu, "Convolutional neural networks for hyperspectral image classification," *Neurocomputing*, vol. 219, pp. 88–98, 2017.
- [60] X. Tang, L. Jiao, W. J. Emery, F. Liu, and D. Zhang, "Two-stage reranking for remote sens. image retrieval," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5798–5817, 2017.
- [61] X. Liu, L. Jiao, J. Zhao, D. Zhang, F. Liu, S. Yang, and X. Tang, "Deep multiple instance learning-based spatial-spectral classification for pan and ms imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. PP, no. 99, pp. 1–13, 2017.
- [62] X. Lu, X. Zheng, and Y. Yuan, "Remote sens. scene classification by unsupervised representation learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5148–5157, 2017.
- [63] X. Lu, X. Li, and L. Mou, "Semi-supervised multitask learning for scene recognition," *IEEE Trans. Cybern.*, vol. 45, no. 9, pp. 1967–1976, 2015.
- [64] S. Yang, M. Wang, Z. Feng, Z. Liu, and R. Li, "Deep sparse tensor filtering network for synthetic aperture radar images classification," *IEEE Trans. Neural Netw. Learn. Syst.*, 2018.
- [65] W. Xie, L. Jiao, B. Hou, W. Ma, J. Zhao, S. Zhang, and F. Liu, "Polsar image classification via wishart-ae model or wishart-cae model," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 10, no. 8, pp. 3604–3615, 2017.
- [66] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric SAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7177–7188, 2017.
- [67] S. W. Chen and C. S. Tao, "Polsar image classification using polarimetric-feature-driven deep convolutional neural network," *IEEE Geosci. Remote Sens. Letters*, vol. PP, no. 99, pp. 1–5, 2018.
- [68] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for scene segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PP, no. 99, pp. 1–1, 2017.
- [69] N. Audebert, B. Le Saux, and S. Lefèvre, "Semantic segmentation of earth observation data using multimodal and multi-scale deep networks," in *Asian Conference on Computer Vision*. Springer, Cham, 2016, pp. 180–196.
- [70] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 39, no. 4, pp. 640–651, 2017.
- [71] M. Siam, M. Gamal, M. Abdel-Razek, S. Yogamani, and M. Jagersand, "Rtseg: Real-time semantic segmentation comparative study," *arXiv preprint arXiv:1803.02758*, 2018.
- [72] Y.-H. Tsai, W.-C. Hung, S. Schulter, K. Sohn, M.-H. Yang, and M. Chandraker, "Learning to adapt structured output space for semantic segmentation," *arXiv preprint arXiv:1802.10349*, 2018.
- [73] X. Liang, H. Zhou, and E. Xing, "Dynamic-structured semantic propagation network," *arXiv preprint arXiv:1803.06067*, 2018.
- [74] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 40, no. 4, pp. 834–848, 2018.
- [75] F. Isikdogan, A. C. Bovik, and P. Passalacqua, "Surface water mapping by deep learning," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 10, no. 11, pp. 4909–4918, 2017.
- [76] G. Cheng, Y. Wang, S. Xu, H. Wang, S. Xiang, and C. Pan, "Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3322–3337, 2017.
- [77] L. Jiao, M. Liang, H. Chen, S. Yang, H. Liu, and X. Cao, "Deep fully convolutional network-based spatial distribution prediction for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5585–5599, 2017.

- [78] M. Volpi and D. Tuia, "Dense semantic labeling of subdecimeter resolution images with convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 881–893, 2017.
- [79] J.-S. Lee and E. Pottier, *Polarimetric radar imaging: from basics to applications*. CRC press, 2009.
- [80] G. D. De Grandi, J. S. Lee, and D. L. Schuler, "Target detection and texture segmentation in polarimetric SAR images using a wavelet frame: Theoretical aspects," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3437–3453, 2007.
- [81] D. Harris and S. Harris, *Digital design and computer architecture*. Morgan Kaufmann, 2010.
- [82] J. Leskovec, A. Rajaraman, and J. D. Ullman, *Mining of massive datasets*. Cambridge university press, 2014.
- [83] RADARSAT2 data set, the polarimetric data provided by Canadian Space Agency. Online: <http://www.rsi.ca/rs2/prod/xml/schemas>.
- [84] AIRSAR data set, the polarimetric data provided by the NASA/JPL-Caltech. Online: <https://earth.esa.int/web/polsarpro/data-sources/>.
- [85] A. H. Strahler, "The use of prior probabilities in maximum likelihood classification of remotely sensed data," *Remote Sens. Environ.*, vol. 10, no. 2, pp. 135–163, 1980.
- [86] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst.*, vol. 13, no. 4, pp. 18–28, 1998.
- [87] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *CVPR*, 2015, pp. 3431–3440.
- [88] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceed. Thirteenth Intern. Confer. Artif. Intell. Stati.*, 2010, pp. 249–256.



Qigong Sun (S'15) received the B.Eng. degrees in intelligence science and technology from Xidian University, Xi'an, China in 2015. He is currently pursuing the Ph.D. degree in circuit and system from Xidian University, Xi'an China. Currently, he is a member of Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, and International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, China. His research interests include deep learning and image processing.



Xu Liu (S'15) received the B.Sc. degrees in Mathematics and applied mathematics from North University of China, Taiyuan, China in 2013. He is currently pursuing the PhD degree in circuit and system from Xidian University, Xi'an China. He is currently a member of Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, and Joint International Research Laboratory of Intelligent Perception and Computation. He is the chair of IEEE Xidian university student branch. Xidian University, Xi'an, China. His current research interests include machine learning, deep learning and image processing.



Licheng Jiao (F'17) received the B.S.degree from Shanghai Jiaotong University, Shanghai, China, in 1982 and the M.S. and PhD degree from Xi'an Jiaotong University, Xi'an, China, in 1984 and 1990, respectively. Since 1992, he has been a Professor with the school of Electronic Engineering, Xidian University, Xi'an, where he is currently the Director of Key Laboratory of Intelligent Perception and Image Understanding of the Ministry of Education of China. His research interests include image processing, natural computation, machine learning, and intelligent information processing.



Dan Zhang received the B.S. degree in intelligent science and technology from Xidian University, Xi'an, China, in 2011. She was a Counselor with Xidian University, Xi'an, China, from 2011 to 2013. And she received the M.S. degree in circuits and systems from Xidian University, Xi'an, China, in 2016. She is currently working in the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University.

Her current research interests include image processing, machine learning.



Xu Tang (M'17) received the B.S., M.S., and Ph.D. degrees from Xidian University, Xi'an, China, in 2007, 2010, and 2017, respectively, where he is currently pursuing the Ph.D. degree in circuit and systems. He is currently a member with the Key Laboratory of Intelligent Perception and Image Understanding, Ministry of Education, Xidian University. His current research interests include remote sensing image processing, remote sensing image contentbased retrieval, and reranking.