PolSAR Image Classification Based on Robust Low-Rank Feature Extraction and Markov Random Field

Haixia Bi, Jing Yao, Zhiqiang Wei, Danfeng Hong, *Member, IEEE,* and Jocelyn Chanussot, *Fellow, IEEE*

Abstract—Polarimetric synthetic aperture radar (PolSAR) image classification has been investigated vigorously in various remote sensing applications. However, it is still a challenging task nowadays. One significant barrier lies in the speckle effect embedded in the PolSAR imaging process, which greatly degrades the quality of the images and further complicates the classification. To this end, we present a novel PolSAR image classification method, which removes speckle noise via low-rank (LR) feature extraction and enforces smoothness priors via Markov random field (MRF). Specifically, we employ the mixture of Gaussian-based robust LR matrix factorization to simultaneously extract discriminative features and remove complex noises. Then, a classification map is obtained by applying convolutional neural network with data augmentation on the extracted features, where local consistency is implicitly involved, and the insufficient label issue is alleviated. Finally, we refine the classification map by MRF to enforce contextual smoothness. We conduct experiments on two benchmark PolSAR datasets. Experimental results indicate that the proposed method achieves promising classification performance and preferable spatial consistency.

Index Terms—PolSAR image classification, low-rank matrix factorization, mixture of Gaussian, Convolutional neural network, Markov random field.

I. INTRODUCTION

POLARIMETRIC synthetic aperture radar (PolSAR) is an advantageous earth observation technique nowadays. It captures abundant and high-resolution earth surface information by transmitting and receiving radar signals in different polarimetric ways. Earth terrains can be retrieved via effective PolSAR image classification techniques, which exposes huge theoretical and application significance in civil and military fields. PolSAR image classification has been delved into extensively in recent years.

However, despite the rapid development, there are still some challenges confronting the PolSAR image classification task, wherein the presence of speckle noise is one of the most significant ones. Researchers have put forward lots of methods to handle the adverse impact of various complex noises [1]–[11]. One conventional way is to design handcrafted features

H. Bi is with Faculty of Engineering, University of Bristol, Bristol BS8 1UB, United Kingdom. (e-mail: haixia.bi@bristol.ac.uk)

J. Yao is with the School of Mathematics and Statistics, Xi'an Jiaotong University, 710049 Xi'an, China. (e-mail: jasonyao@stu.xjtu.edu.cn)

Z. Wei is with the Xi'an Electronics and Engineering Institute, 710100 Xi'an, China. (e-mail: zqwei@fudan.edu.cn)

D. Hong is with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Wessling, Germany. (e-mail: danfeng.hong@dlr.de)

J. Chanussot is with the Aerospace Information Research Institute, Chinese Academy of Sciences, 100094 Beijing, China. (e-mail: jocelyn@hi.is)

using sliding window techniques. Wavelet technique [1] was leveraged to extract polarization wavelet features in the space of polarization states. Literature [2] proposed to use Gabor filtering and wavelet analysis to obtain contextual features. A more heuristic way is to extract features based on learning-based methods, e.g., convolutional neural network (CNN) [3], [12], [13]. They have been employed to automatically extract data-driven polarimetric features, and were further combined with semi-supervised learning and active learning in literature [4], [5]. A complex-valued CNN was applied in PolSAR image classification in [10]. An alternative strategy is to encourage local consistency on pixel labels. Typical methods include the over-segmentation technique [14] used in preprocessing stage and Markov random field (MRF)-based optimization employed as a post-processing process [8], [9].

Although the previous approaches achieved good results on mitigating the performance degradation caused by speckle noise, they do not explicitly consider removing noise while extracting features. To solve this problem from the source, we propose to simultaneously perform noise removal and feature extraction. Specifically, we leverage the mixture of Gaussian (MoG)-based robust low-rank matrix factorization (RLRMF) to eliminate noise disturbances and draw compact and robust polarimetric features, where the universal approximation power of MoG model is exploited to characterize the PolSAR noise with complex and heterogenous distributions. With the extracted RLRMF features, a classification map is then obtained by applying data augmentation incorporated CNN classifier on them, where local consistency is implicitly enforced by the convolution operations. Finally, based on the output probabilities of the CNN classifier, the classification map is refined by solving an MRF optimization problem to further enforce spatial smoothness.

This paper presents a CNN-based PolSAR image classification method incorporating RLRMF and MRF to suppress speckle noises, further promoting the classification performance. The main contributions of this letter are two-fold:

1) The speckle noise is suppressed in three different dimensions, i.e., mainly reduced via RLRMF-based noise removal and feature extraction, suppressed by CNN-based classification, and MRF prior-based post-processing. To our knowledge, this is the first work that considers suppressing complex noises from diverse views and extracting discriminative features for classification.

2) RLRMF features with data augmentation can provide robust and discriminative information for CNN classi-
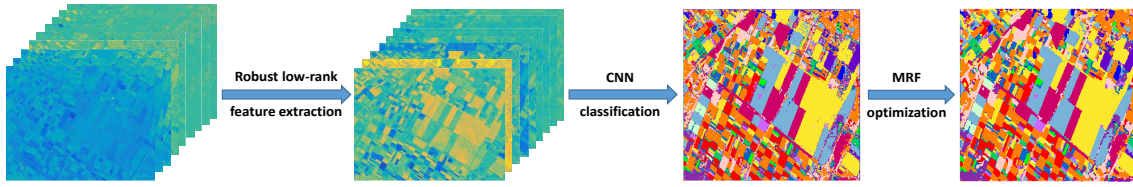
Figure 1. Illustrating the PolSAR image classification method via robust low-rank feature extraction and Markov random field.

fier, which effectively alleviates the insufficient label problem. Beyond the traditional MRF model that only enforces label smoothness priors, our MRF not only encourages the contextual consistency, but also enforces the alignment of label boundaries with image edges.

II. PROPOSED METHOD

A. Problem Formulation

The input polarimetric data for each pixel is a 3×3 complex conjugate symmetric coherency matrix $\mathbf{T}$. For convenience, $\mathbf{T}$ can be vectorized to a 9×1 real vector as

$$[\mathbf{T}_{11}, \mathbf{T}_{22}, \mathbf{T}_{33}, \Re(\mathbf{T}_{12}), \Im(\mathbf{T}_{12}),$$
$$\Re(\mathbf{T}_{13}), \Im(\mathbf{T}_{13}), \Re(\mathbf{T}_{23}), \Im(\mathbf{T}_{23})]^T, \quad (1)$$

where $\Re(\cdot)$ and $\Im(\cdot)$ extract the real and imaginary components of the cross-polarization items. Superscript ^{T} denotes the transpose operation. Each vector dimension in Eq. (1) is normalized to [0,1] before feeding into the proposed method.

Then for a given PolSAR image, the raw PolSAR data set can be represented by $\mathbf{X} = \{\mathbf{x}_i \in \mathbb{R}^d\}_{i=1}^N$, $N = H \times W$. Here, H and W are height and width of the image, d is the dimension of the raw feature, which takes the value of 9 as defined in Eq. (1). We define the labeled training samples as $(\mathbf{x}_i, \mathbf{y}_i)_{i=1}^n$, where $n \ll N$, $\mathbf{y}_i \in \{1, \cdots, C\}$, C is the total number of classes. Then the proposed method is designed to assign class label $\mathbf{y}_i$ to each pixel i. We further denote $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^N$ in the following sections. In addition, we denote $\mathbf{F}_{ij}$ as the pixel vector at location (i, j), and $\mathbf{S}_{ij} = \cup_{\{(u,v):|u-i|\leq\lfloor\frac{s}{2}\rfloor,|v-j|\leq\lfloor\frac{s}{2}\rfloor\}}\mathbf{F}_{uv}$ as the neighborhood of $\mathbf{F}_{ij}$ in the spatial dimension by using a sliding window box with size $s \times s$. Here, s is the window size, and $\lfloor\frac{s}{2}\rfloor$ is the rounded down integer of number $\frac{s}{2}$.

Taking the Flevoland area data set as example, Fig. 1 illustrates the pipeline of our proposed method. The proposed RLRMF feature extraction is first implemented to remove complex noise while extracting features (Section II-B). We then train a CNN classifier to obtain a classification map, and the classification map is finally refined by solving an MRF prior-based optimization problem. (Section II-C).

B. Feature Extraction Using RLRMF

Since neighboring pixels in PolSAR images are prone to have high affinities, i.e., PolSAR data has low-rank characteristics, RLRMF is thus an effective tool to capture the global structures of the data and remove noise [15]. For each spatial neighborhood matrix $\mathbf{S} \in \mathbb{R}^{d \times s^2}$ (we omit subscript ij for simplicity), the RLRMF problem can be formulated as,

$$\min_{\mathbf{U}, \mathbf{V}} \|\mathbf{S} - \mathbf{U}\mathbf{V}^T\|, \quad (2)$$

where $\mathbf{U} \in \mathbb{R}^{d \times r}$ and $\mathbf{V} \in \mathbb{R}^{s^2 \times r}$ are matrices with low-rank property ($r < \min(s^2, d)$). $\|\cdot\|$ represents the noise measure of PolSAR data. After obtaining an optimal solution $(\mathbf{U}^*, \mathbf{V}^*)$, $\mathbf{S}$ can be recovered by the product $\mathbf{U}^*\mathbf{V}^{*T}$.

Due to the complexity and heterogeneity of polarimetric noise, MoG assumption is used to model the unknown noise distributions. We denote the noise term as $\mathbf{E}$ with entries ϵ_{ij}, each of which is a sample from MoG distribution $p(\epsilon) \sim \sum_{k=1}^K \pi_k \mathcal{N}(\epsilon; 0, \sigma_k^2)$, where $\mathcal{N}(\epsilon; 0, \sigma^2)$ indicates the Gaussian distribution with mean 0 and variance σ^2. Here, $\pi_k > 0$, and $\sum_{k=1}^K \pi_k = 1$, where K is the number of mixed components. Under maximum likelihood estimation framework, the log-likelihood of noise term $\mathbf{E}$ is written as,

$$P(\mathbf{E}) = \sum_{i=1}^d \sum_{j=1}^{s^2} \log \sum_{k=1}^K \pi_k \mathcal{N}(\epsilon_{ij}|0, \sigma_k^2). \quad (3)$$

Since each element $\mathbf{s}_{ij}$ in $\mathbf{S}$ satisfies $\epsilon_{ij} = \mathbf{s}_{ij} - \mathbf{u}_i^T\mathbf{v}_j$, where $\mathbf{u_i}$ and $\mathbf{v_j}$ are the ith and jth column vectors of $\mathbf{U}$ and $\mathbf{V}$ respectively, let $\Pi = \{\pi_1, \pi_2, \cdots, \pi_K\}$, $\Sigma = \{\sigma_1, \sigma_2, \cdots, \sigma_K\}$, and $\Theta = \{\mathbf{U}, \mathbf{V}, \Pi, \Sigma\}$, our objective is to maximize the log-likelihood function as,

$$\max_{\Theta} \mathcal{L}(\Theta) = \sum_{i=1}^d \sum_{j=1}^{s^2} \log \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{s}_{ij} - \mathbf{u}_i^T\mathbf{v}_j|0, \sigma_k^2). \quad (4)$$

The Expectation-Maximization (EM) algorithm is employed to solve the model defined in Eq. 4. That is, the parameters are estimated by iteratively conducting *E-step*, i.e., calculating posterior probabilities of all Gaussian components, and *M-step*, i.e., re-estimating the parameters by maximizing the Q-function which is the upper-bound of the original likelihood.

E-Step: We assume a latent variable z_{ijk} in the model with $z_{ijk} \in \{0, 1\}$ and $\sum_{k=1}^K = 1$, indicating the assignment of the noise ϵ_{ij} to a kth component of the mixture. Then, the posterior possibility of mixture k for generating the noise of $\mathbf{s}_{ij}$ can be calculated by,

$$E(z_{ijk}) = \gamma_{ijk} = \frac{\pi_k \mathcal{N}(\mathbf{s}_{ij}|\mathbf{u}_i^T\mathbf{v}_j, \sigma_k^2)}{\sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{s}_{ij}|\mathbf{u}_i^T\mathbf{v}_j, \sigma_k^2)}. \quad (5)$$

M-Step: M step maximizes the Q-function constructed from the posterior possibility in E-step,

$$E_{\mathbf{z}}p(\mathbf{S}, \mathbf{Z}|\Theta) = \sum_{i,j,k} \gamma_{ijk}(\log \frac{\pi_k}{\sqrt{2\pi}\sigma_k} - \frac{(\mathbf{s}_{ij} - \mathbf{u}_i^T\mathbf{v}_j)^2}{2\pi\sigma_k^2}). \quad (6)$$

Then we solve this optimization problem by alternatively updating the MoG parameters $\mathbf{U}, \mathbf{V}$ and the factorized matrices Π, Σ as follows:

(1) Update Π, Σ: Closed-form solutions for the MoG parameters can be given by,

$$\sigma_k^2 = \frac{1}{\sum_{i,j} \gamma_{ijk}} \sum_{i=1}^{d} \sum_{j=1}^{s^2} \gamma_{ijk} (\mathbf{S}_{ij} - \mathbf{u}_i^T \mathbf{v}_j)^2. \tag{7}$$

(2) Update $\mathbf{U}, \mathbf{V}$: By reformulating the objective function in Eq. 6 w.r.t $\mathbf{U}$ and $\mathbf{V}$, we have below optimization problem,

$$\min_{\mathbf{U}, \mathbf{V}} \left\| \mathbf{W} \odot (\mathbf{S} - \mathbf{U}\mathbf{V}^T) \right\|_F^2, \tag{8}$$

where the element w_{ij} of $\mathbf{W}$ is $\sum_{k=1}^{K} \frac{\gamma_{ijk}}{2\pi\sigma_k}$. Problem (8) can be solved by [16].

With the solution $(\mathbf{U}^*, \mathbf{V}^*)$ obtained using the EM algorithm, the spatial neighborhood $\mathbf{S}$ can be recovered by $\hat{\mathbf{S}} = \mathbf{U}^*\mathbf{V}^{*T}$. Then, the extracted RLRMF feature $\hat{\mathbf{F}}$ is used as input in the consequent classification process instead of the original feature $\mathbf{F}$.

C. CNN-MRF Classification

Our proposed CNN-MRF PolSAR classification method is composed of two steps: CNN classification and MRF prior-based optimization.

1) CNN Classification: CNN is leveraged to predict class label based on the RLRMF features. We adopt a CNN architecture [17] which consists of two groups of convolutional layer and pooling layer, followed by two fully connected layers and an output softmax layer. Batch normalization is performed after the two convolutional layers and the first fully connected layer. We exploit data augmentation technique following [5] to enlarge the training set, alleviating the reliance of CNN on a large amount of labels. Let Φ denote the combination of the filters, weight matrices and biases of the CNN, the loss function of the CNN is defined as

$$E_d = -\sum_{i=1}^{n} \sum_{j=1}^{C} \mathbf{1}\{\mathbf{y}_i = j\} \log P(\mathbf{y}_i = j | F_\Phi, \hat{\mathbf{f}}_i), \tag{9}$$

where $P(\mathbf{y}_i = j | F_\Phi, \hat{\mathbf{f}}_i)$ denotes the probability of $\hat{\mathbf{f}}_i$ having label j. We first train a CNN by minimizing the loss function in Eq. (9). Then the learned classifier F_Φ is applied on the whole data set to obtain class probabilities for the MRF optimization.

2) MRF Prior-based Optimization: This step is used to predict the final labels $\mathbf{Y}$ under the constraint of MRF priors where the smoothness of estimated label map and the alignment of label boundaries with image edges are enforced. This task can be expressed as the following optimization problem,

$$\mathbf{Y} = \arg\min_{\mathbf{y} \in \{1,...,C\}} \left(-\sum_{i=1}^{N} \log P(\mathbf{y}_i) + \alpha_s \sum_{i=1}^{N} \sum_{j \in \mathcal{N}(i)} \mathcal{S}_{ij} \right). \tag{10}$$

The first part of the optimization objective is the CNN output probability, and the second part is the label smoothness term with $\mathcal{S}_{ij} = |\mathbf{y}_i - \mathbf{y}_j| \exp(-\frac{\|\mathbf{z}_i - \mathbf{z}_j\|_2^2}{2\sigma})$, α_s as the label smoothness factor, and $\mathcal{N}(i)$ as the neighboring pixel set of pixel i. $\mathbf{z}_i$ should take features significantly changing values across the image edges. Due to the proper delineation of homogeneous areas and edges, Pauli matrix components are taken as $\mathbf{z}_i$, where δ indicates the mean squared distance between features

of two adjacent pixels. Solving Eq. 10 is a combinatorial optimization problem. Belief propagation [18] optimization algorithm is employed here due to its quick convergence.

III. EXPERIMENTS

In this section, we validate our proposed method on two real PolSAR data sets. We first introduce the experimental data sets and parameter settings in Section III-A. Next, the experimental results are reported in Section III-B, where comparison is performed between the proposed CNN-RLRMF-MRF method (denoted as Ours) and one traditional method SVM [19], and three state-of-the-art competitors, including CNN [3], CNN-MRF [20] and complex-valued CNN (CV-CNN) [10]. We conduct ablation study in Section III-C, to demonstrate the validity of the key components of the proposed method.

A. Experimental Data and Experimental Settings

Figure 2 and 3 display the experimental images employed for evaluation. The first data set is an L-band image collected by AIRSAR over Flevoland area in 1991 with size 1020×1024. The second data set is an E-SAR L-band image provided by German Aerospace Center over Oberpfaffenhofen area in Germany with size 1300×1200. It is noteworthy that, for both data sets, we utilize much less labeled data than compared methods, which are summarized in Table I.

In our experiments, we empirically set patch size s as 7. We performed experiments to assess the impact of different rank values $\{2, 3, 4, 5\}$ on the classification accuracy over the Flevoland area dataset. Experimental results show that $r = 2$ achieves the best performance. Therefore, rank r is set to 2 throughout the experiments. An automatic K estimation method is employed to determine the number of Gaussian noise components. The EM iterations stop when the iteration reaches 100 or the difference between two consecutive $\mathbf{U}_s$ is smaller than 0.01. For the CNN architecture, the first convolutional layer contains 20 filters with size 3×3, while the second convolutional layer involves 50 filters with size 2×2. The kernel size of the two max-pooling layer is 2×2. The first fully connected layer contains 500 units, and the unit number of the second one takes the value of the class number. We empirically set the learning rate τ, weight decay and momentum parameter to 0.001, 0.0005 and 0.9, respectively. The batch size is set as 50, and early stopping criteria is utilized to prevent overfitting by stopping training when the validation accuracy continuously decreases. The smoothness factor α_s is set as 5.

Table I
USED LABELS ON TWO DATA SETS OF DIFFERENT METHODS

Method	Flevoland	Oberpfaffenhofen
SVM [19]	13535 (10%)	137429 (10%)
CNN [3]	13535 (10%)	137429 (10%)
CNN-MRF [20]	13535 (10%)	137429 (10%)
CV-CNN [10]	11900 (8.8%)	13743 (1%)
Ours	2707 (2%)	6872 (0.5%)

Table II
CAs (%) AND OAs (%) OF FLEVOLAND AREA DATA SET WITH DIFFERENT METHODS

| Method	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	OA
SVM	83.82	26.93	41.25	90.56	89.55	16.67	92.66	60.54	14.26	54.88	86.82	87.17	37.80	48.88	80.82
CNN	97.06	99.79	92.54	95.01	93.92	88.12	97.80	95.75	99.95	70.93	97.72	99.72	92.34	96.78	96.58
CNN-MRF	99.34	**99.89**	98.71	98.68	97.08	**96.81**	99.80	**96.77**	**99.95**	82.09	98.84	**99.98**	94.72	97.76	98.70
CV-CNN	**99.80**	98.30	98.90	96.10	**99.60**	93.20	**99.90**	90.70	99.00	**98.20**	97.00	99.80	96.60	**99.20**	99.00
Ours	99.69	96.35	**99.43**	98.93	**99.60**	95.69	99.67	93.51	98.47	92.74	**99.64**	99.47	**97.74**	96.71	**99.07**
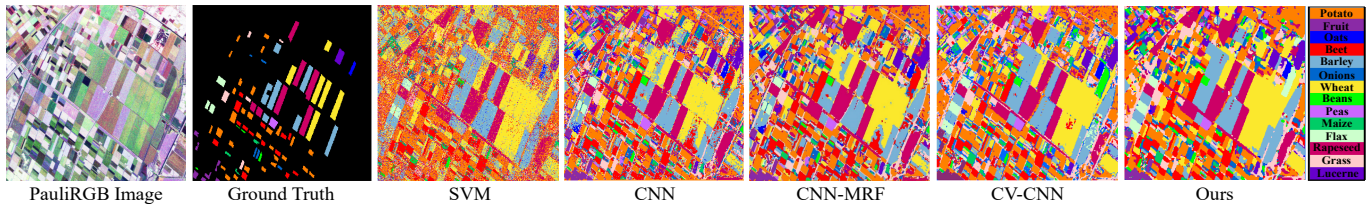
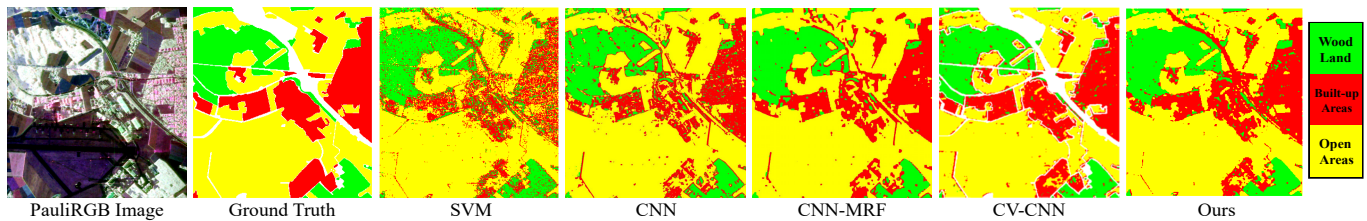Figure 2. Experimental data and results of Flevoland area data set.

Figure 3. Experimental data and results of Oberpfaffenhofen area data set.

Table III
CAs (%) AND OAs (%) OF OBERPFAFFENHOFEN AREA DATA WITH
DIFFERENT METHODS

Method	C1	C2	C3	OA
SVM	60.15	86.59	94.87	84.67
CNN	76.09	86.79	97.42	90.07
CNN-MRF	82.85	90.42	97.65	92.84
CV-CNN	**91.30**	92.20	94.60	93.40
Ours	86.09	**95.68**	**97.95**	**94.58**

B. Results and Comparisons

1) Flevoland Area Data set: The classification results of the five compared methods on the Flevoland area data set are illustrated in Fig. 2. Table II presents the classification accuracy (*CA*) and overall CA (*OA*) values. Observing the figure and table, we highlight below main observations:

(a) The class label map of our proposed method presents a desirable visual effect with preferable contextual consistency and clear label boundaries while well preserving the image details, which is due to noise removal of RLRMF feature representation, feature spatial consistency encouragement of CNN, and enforcement of label smoothness of MRF.

(b) Our method achieves the best OA with only 2% labels used, compared with SVM, CNN and CNN-MRF using 10% labels, and CV-CNN using 8.8% labels. This is attributed to the stronger feature representation and classification performance of our method, which greatly alleviates the reliance of deep learning (DL) method on large labeled sample set.

2) Oberpfaffenhofen Area Data set: Figure 3 shows the classification results of Oberpfaffenhofen data set with five competing methods, and the numerical results are displayed in Table III. Since the original paper of CV-CNN [10] did not provide the classification result of the whole image but gave

the result overlaid with the ground truth labels instead, there are some void areas in the CV-CNN result. We can find from the results that, due to the combinatorial effect of RLRMF features, data augmented CNN and MRF optimization, CNN-RLRMF-MRF yields the highest OA value using less labels (0.5%), compared with SVM, CNN and CNN-MRF using 10% labels, and CV-CNN using 1% labels.

C. Ablation Study

To validate the effectiveness of the three components of the proposed method, i.e., RLRMF features, CNN and MRF optimization, we conduct following six groups of experiments – (1) RF-Raw: using raw features and Random forest (RF) classifier; (2) RF-Lee: using features with refined Lee noise reduction [21] and RF classifier; (3) RF-RLRMF: using RLRMF features and RF classifier; (4) CNN-Lee: using features with refined Lee noise reduction [21] and CNN classifier; (5) CNN-RLRMF: using RLRMF features and CNN; (6) Ours: using RLRMF features, CNN and MRF prior-based optimization. 2% of the ground truth labels are used in training.

Taking the Flevoland area data set for example, Fig. 4 presents the class label maps and Table IV shows the *CA* values of the six compared methods, where the best results are highlighted in bold. From the figure and table, we can conclude:

(1) The class label map of RF-Raw method exhibits clear speckle-like appearance, while this issue is obviously relieved in the result of RF-RLRMF method, which gracefully demonstrates the noise removal effect of RLRMF features. The RLRMF features improve the OA by 16.17 percentage points.

Table IV
CAs (%) AND OAs (%) OF FLEVOLAND AREA DATA SET FOR ABLATION STUDY

| Method	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	OA
RF-Raw	85.76	13.26	36.59	91.07	88.51	6.10	92.66	34.10	6.99	58.06	82.66	86.73	17.55	32.83	78.83
RF-Lee	98.23	77.14	0	98.76	97.38	0.28	99.24	91.50	31.90	20.39	95.54	97.90	66.56	73.54	91.55
RF-RLRMF	96.54	93.29	82.93	96.01	97.77	33.33	97.67	90.76	95.83	87.83	96.21	96.30	89.27	84.18	95.00
CNN-Lee	99.45	**99.82**	99.00	94.66	99.25	53.85	99.65	**97.60**	**98.81**	**97.91**	98.30	98.86	93.65	89.46	97.42
CNN-RLRMF	99.19	95.73	98.06	98.13	99.34	88.77	99.29	88.72	98.43	87.88	98.98	99.16	96.74	95.43	98.56
Ours	**99.69**	96.35	**99.43**	**98.93**	**99.60**	**95.69**	**99.67**	93.51	98.47	92.74	**99.64**	**99.47**	97.74	**96.71**	**99.07**
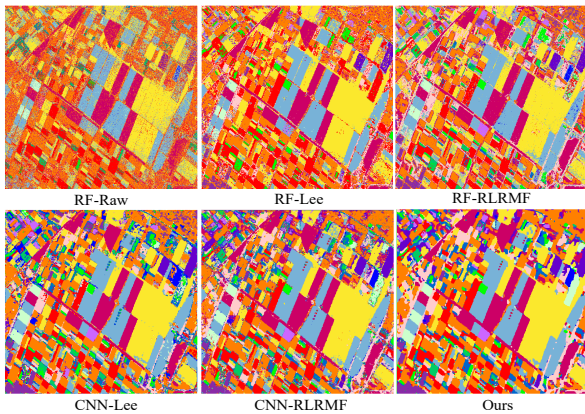
Figure 4. Experimental results for ablation study.

The RF-RLRMF and CNN-RLRMF outperform RF-Lee and CNN-Lee by 3.45 and 1.14 percentage points, respectively, which demonstrates the advantage of RLRMF features.

(2) CNN and MRF further promote the classification performance and spatial consistency, which bring gains of 3.56 and 0.51 percentage points in OA respectively. Furthermore, the label boundaries are better aligned with the image edges after MRF optimization.

IV. CONCLUSIONS

In this paper, we presented a PolSAR classification method involving RLRMF-based feature extraction and MRF-based label optimization. The advantages of our work lie in two points: (1) The proposed method simultaneously extracts robust features and removes noises via the MoG-based low-rank modeling. (2) The low-rank feature extraction, CNN, and MRF effectively improve the classification performance, mitigating the reliance of DL method on massive labels and promoting the spatial smoothness. In the future, we are interested in developing a DL-based multi-level feature fusion method [22] to suppress noise and investigating the inner relationship between the coherency matrix elements to further boost the PolSAR image classification performance.

REFERENCES

[1] C. He, S. Li, Z. Liao, and M. Liao, "Texture classification of polsar data based on sparse coding of wavelet polarization textons," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 8, pp. 4576–4590, 2013.

[2] A. Masjedi, M. J. V. Zoej, and Y. Maghsoudi, "Classification of polarimetric sar images based on modeling contextual information and using texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 932–943, 2015.

[3] 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.

[4] H. Bi, J. Sun, and Z. Xu, "A graph-based semisupervised deep learning model for polsar image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2116–2132, 2018.

[5] H. Bi, F. Xu, Z. Wei, Y. Xue, and Z. Xu, "An active deep learning approach for minimally supervised polsar image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9378–9395, 2019.

[6] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, 2019.

[7] J. Yao, D. Meng, Q. Zhao, W. Cao, and Z. Xu, "Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2991–3006, 2019.

[8] H. Bi, L. Xu, X. Cao, and Z. Xu, "Polsar image classification based on three-dimensional wavelet texture features and markov random field," in *Proc. IGARSS*, pp. 3921–3928, IEEE, 2017.

[9] H. Bi, J. Sun, and Z. Xu, "Unsupervised polsar image classification using discriminative clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3531–3544, 2017.

[10] 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.

[11] L. Gao, D. Hong, J. Yao, B. Zhang, P. Gamba, and J. Chanussot, "Spectral superresolution of multispectral imagery with joint sparse and low-rank learning," *IEEE Trans. Geosci. Remote Sens.*, 2020. DOI:10.1109/TGRS.2020.3000684.

[12] S.-W. Chen and C.-S. Tao, "Polsar image classification using polarimetric-feature-driven deep convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 627–631, 2018.

[13] D. Hong, L. Gao, J. Yao, B. Zhang, P. Antonio, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *arXiv preprint arXiv:2008.02457*, 2020.

[14] D. Hong, X. Wu, P. Ghamisi, J. Chanussot, N. Yokoya, and X. X. Zhu, "Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3791–3808, 2020.

[15] D. Meng and F. De La Torre, "Robust matrix factorization with unknown noise," in *Proc. ICCV*, pp. 1337–1344, 2013.

[16] D. Hong, N. Yokoya, N. Ge, J. Chanussot, and X. X. Zhu, "Learnable manifold alignment (LeMA): A semi-supervised cross-modality learning framework for land cover and land use classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 147, pp. 193–205, 2019.

[17] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, and J. Benediktsson, "Feature extraction for hyperspectral imagery: The evolution from shallow to deep (overview and toolbox)," *IEEE Geosci. Remote Sens. Mag.*, 2020. DOI: 10.1109/MGRS.2020.2979764.

[18] M. F. Tappen and W. T. Freeman, "Comparison of graph cuts with belief propagation for stereo, using identical mrf parameters," in *Proc. ICCV*, vol. 2, pp. 900–907, IEEE, 2003.

[19] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "Cospace: Common subspace learning from hyperspectral-multispectral correspondences," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4349–4359, 2019.

[20] X. Cao, Z. Xu, and D. Meng, "Spectral-spatial hyperspectral image classification via robust low-rank feature extraction and markov random field," *Remote Sens.*, vol. 11, no. 13, p. 1565, 2019.

[21] J.-S. Lee, M. R. Grunes, and G. De Grandi, "Polarimetric sar speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2363–2373, 1999.

[22] D. Hong, L. Gao, N. Yokoya, J. Yao, J. Chanussot, D. Qian, and B. Zhang, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Trans. Geosci. Remote Sens.*, 2020. DOI: 10.1109/TGRS.2020.3016820.