SAR4LCZ-Net: A Complex-valued Convolutional Neural Network for Local Climate Zones Classification Using Gaofen-3 Quad-pol SAR Data

Rui Zhang, Yuanyuan Wang, *Member, IEEE*, Jingliang Hu, Wei Yang, *Member, IEEE*, Jie Chen, *Senior Member, IEEE*, Xiao Xiang Zhu, *Fellow, IEEE*

***Abstract*—The recent local climate zones (LCZ) classification scheme provides spatially fine granular descriptions of inner-urban morphology. It is universally applicable to cities worldwide and capable of supporting various urban studies. Although optical and dual-pol SAR data continue to push the frontiers of this task, the potential of quad-pol SAR data for LCZ classification is not yet explored. In this paper we propose a novel complex-valued convolutional neural network (CNN), *SAR4LCZ-Net*, to tackle this challenge. SAR4LCZ-Net improves the state of the art by exploiting two facts of this specific task: the semantic hierarchical structure of the LCZ classification scheme, and the complex-valued nature of quad-pol SAR data. To validate the performance of our algorithm, we generate a Chinese Gaofen-3 quad-pol SAR data set for LCZ which covers 31 cities around the world. Results show that the proposed SAR4LCZ-Net improves 2.4% on overall accuracy and 4.5% on average accuracy compared to the real-valued CNN with same structure. Gaofen-3 quad-pol SAR data also showed its advantage over the dual-pol Sentinel-1 data. It enhanced 5.0% on overall accuracy and 7.2% on average accuracy in LCZ classification, under a fair comparison with a model trained by Sentinel-1 of the same area.**

***Index Terms*—Quad-pol SAR, complex-valued convolutional neural networks, local climate zones, urban land cover.**

I. INTRODUCTION

LOCAL climate zones (LCZ) provides an universally applicable classification system for the urban environment. The 17 LCZ classes are based on climate-relevant surface properties on the local-scale, mainly related to the 3D structure, spatial density of the surface structure, and covering material of the surface. LCZ depicts 10 built classes and 7 natural classes [1] as shown in Fig.1, which offer the possibility to compare different areas of different cities with trenchant distinctions representing morphological structures of urban local neighborhood [2]. Apart from an increasing influence on various climate research such as the cooling effect of green infrastructure and micro-climatic effects on town peripheries [1], [3], LCZ can also be used to describe the internal structure of urban areas, providing information for infrastructure planning and population assessment [4]–[6]. For instance, compact built areas in LCZ classes 1 to 3 are normally located in downtown areas with massive population concentration, while open built area appearing as classes 4 to 6 are usually distributed in suburb areas with less population density. Therefore, thanks to the described morphological structure, LCZ maps can act as a valuable source for a wide variety of studies in urban areas [7]. However, reliable LCZ classification maps are not available on a global scale, making it hard to compare and combine their urbanization works. Because remote sensing data have been routinely used for land cover research, they have great potential for LCZ classification. Nevertheless, global LCZ mapping is still challenging due to the limited number of high quality remote sensing datasets related to LCZ maps in large scale. Due to the rich characteristics of natural processes and environments, it is rare for a single acquisition method to provide a complete understanding of certain phenomenon. It is necessary to explore multi-source remote sensing data to generate LCZ mapping.

A. Related work

Most studies of LCZ classification mainly focus on using optical remote sensing data [8], [9]. Among them, the World Urban Database and Portal (WUDAPT) is a well-known community-based global LCZ mapping using openly available Landsat data and software [10]. Within WUDAPT, almost 100 cities located across the globe have already been mapped in moderate resolution, providing sufficient details for a number of applications [11]. Recently researchers are noticing the potential of synthetic aperture radar (SAR) data for LCZ classification. The Sentinel-1 mission with a dual-pol SAR has been studied for LCZ classification in [12]. It has been proved that the combination of Sentinel-1 dual-pol SAR data and LandSat-8 data can improve the performance

This work is supported by the National Natural Science Foundation of China (NSFC) under Grant No.61861136008 and Grant No.61701012, TUM-Beihang doctoral exchange program and the China Scholarship Council under CSC NO.202006020127, as well as by the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program with the grant number ERC-2016-StG-714087 (Acronym: So2Sat, project website: www.so2sat.eu), and the Helmholtz Association under the framework of the Helmholtz Artificial Intelligence Cooperation Unit–Local Unit "Munich Unit @Aeronautics, Space and Transport (MASTr)," and Helmholtz Excellent Professorship "Data Science in Earth Observation–Big Data Fusion for Urban Research." (Corresponding authors: Jie Chen and Xiao Xiang Zhu).

Rui Zhang, Wei Yang, and Jie Chen are with the School of Electronic and Information Engineering, Beihang University, Beijing 100191, China.(e-mail: zrui201@buaa.edu.cn; chenjie@buaa.edu.cn; yangweigigi@sina.com). Jie Chen and Wei Yang are also with Collaborative Innovation Center of Geospatial Technology, Wuhan, China.

Yuanyuan Wang, Jingliang Hu, and Xiao Xiang Zhu are with Data Science in Earth Observation, Technical University of Munich (TUM), Arcisstr.21, 80333 Munich, Germany. Yuanyuan Wang and Xiao Xiang Zhu are also with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Wessling, Germany.(e-mail: Yuanyuan.Wang@dlr.de; Jingliang.Hu@dlr.de; Xiaoxiang.Zhu@dlr.de).

of LCZ classification. Large-scale LCZ classification maps [7] are furtherly attempted to produce for the first time using dual-pol SAR data. A newly valuable benchmark dataset named "So2Sat LCZ42" [13], which consists of LCZ labels of about half a million Sentinel-1 dual-pol SAR and Sentinel-2 optical image patches in 42 urban agglomerations across the globe was shared to the community for machine learning LCZ model development. Normally dual-pol modes simply do away with one of the transmitted polarizations. Comparing to dual-pol SAR, quad-pol SAR collect the full polarimetric scattering matrix and complete polarimetric scattering information. Complete polarimetric information obviously permits better image classification and analysis in land cover land use. However, there is no study to our best knowledge utilizing quad-pol SAR data for LCZ classification. Therefore, this work tries to comprehensively investigate quad-pol SAR data for the LCZ classification task.

From the perspective of methodology, several convolutional neural networks (CNNs) have been applied in LCZ classification study due to their feature grasping ability. One CNN model achieved effective LCZ classification based on So2Sat LCZ42 dataset [14]. A modified Densenet framework with SAR and multispectral images as inputs was proposed for LCZ classification [15]. However, the hierarchical structure of categories in LCZ classification is seldom considered corresponding to CNN. Shallower layers of CNN extract feature as salient corners and edges [20], which can distinguish coarse classes in LCZ such as built classes and natural classes. While deeper layers extract representations that can differ urban local neighborhood of different morphologies. CNN outputs multiple predictions ordered from coarse to fine along the concatenated convolutional layers corresponding to the hierarchical structure of the target classes, which can be regarded as a form of prior knowledge on the output. A CNN with hierarchical layered structure is designed in [21] to learn internal representations that abstract from the input data. Hierarchical deep CNN was introduced by embedding deep CNN into a two-level category hierarchy in [22], achieves state-of-the-art results on optical benchmark datasets.

Indeed, amounts of land cover land use classifications have already adopted quad-pol SAR image as data source and CNNs as methodology. Terrain classification framework [16] used a novel quad-pol SAR based on deep CNN. In order to classify all pixels in a patch of urban area simultaneously, a fixed-feature-size CNN was used in [17]. A deep self-paced CNN [18] was employed for river area classification, which retains mutual information between channels and texture information from the raw quad-pol SAR data. For building and vegetation classification [19], a quad-pol SAR image classification method based on residual network (ResNet) and deep autoencoder was proposed. Therefore, the application potential of quad-pol SAR image in LCZ classification is anticipated.

In aforementioned CNNs for image classification, input features of those CNNs are mainly real-valued amplitude. As quad-pol SAR dataset is in complex-valued domain, several researches have been focusing on complex-valued CNN for quad-pol SAR image classification. In [23], amplitude and phase of complex-valued quad-pol SAR data are extracted as the input for the first time to maintain the integrity of original information while avoiding complex-valued operations, then a multi-task CNN architecture is proposed to match the improved input form and achieve better classification results. Differently, a complex-valued CNN uses real part and imaginary part of quad-pol SAR images as input features [24] for land cover classification. All elements of CNN including input and output layer, convolutional layer, activation function, and pooling layer are extended to the complex domain.

There exist different strategies to work with complex-valued PolSAR data. The elements of the covariance matrix and the coherence matrix with the most abundant polarization information were selected in [25]. Feature combinations [26] were selected rather than individual features for quad-pol SAR image classification. The polarization features from the complex coherency matrix and the spatial features of the Pauli image were extracted for quad-pol SAR image classification in [27]. In addition to the input features, activation functions are also important for complex-valued CNN. The phase information can be easily influenced by the activation layer. Some complex-valued CNN adopt CReLU and ZReLU as activation functions in the network [28], and some use Cardioid [29] to scale the input magnitude but retain the input phase. For output layer, previous complex-valued CNNs output real units and imaginary units directly then use the least-squares loss function [30].

B. Main Contributions

In this work, a complex-valued CNN named SAR4LCZ-Net was developed. SAR4LCZ-Net is designed to fully utilize the information in quad-pol SAR data for LCZ classification. Main contributions of this work are as follows:

(1) Shallower and deeper representations of CNN layers were associated to hierarchical semantic meaning of LCZ classes and formulated them into loss functions.

(2) Spatial features and polarimetric features were combined to feed into the network, activated by CReLU complex-valued activation function in convolutional layers and complex-valued softmax in output layer.

(3) Gaofen-3 quad-pol SAR images covering 31 cities around the world were collected to validate the feasibility of proposed network.

The rest of the paper is organized as follows. Section II demonstrates the principle of selecting a study area and describes the Gaofen-3 quad-pol SAR data and its data preparation. Section III elaborates on the proposed CNN architecture, SAR4LCZ-Net, and the strategy of hierarchical structure of LCZ classification. Section IV introduces the baseline CNNs to be compared, and the experimental setup. The classification accuracy of LCZ is evaluated and LCZ results for several test scenes are visualized. Lastly, Section V concludes this work.

II. STUDY AREA AND DATASET

In this section, we show Gaofen-3 mission and the corresponding LCZ patches, the global scale coverage, as well as the setup of training and testing dataset.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3137911, IEEE Transactions on Geoscience and Remote Sensing

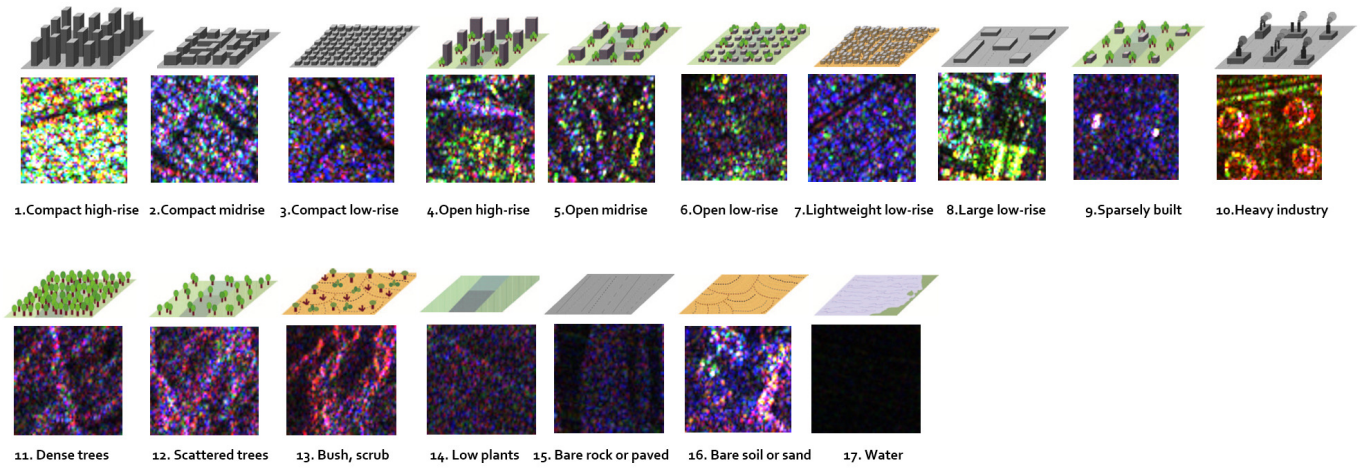


Fig. 1. Illustration of LCZs [1] and corresponding SAR images. Bottom row is Gaofen-3 quad-pol mode Pauli RGB image.

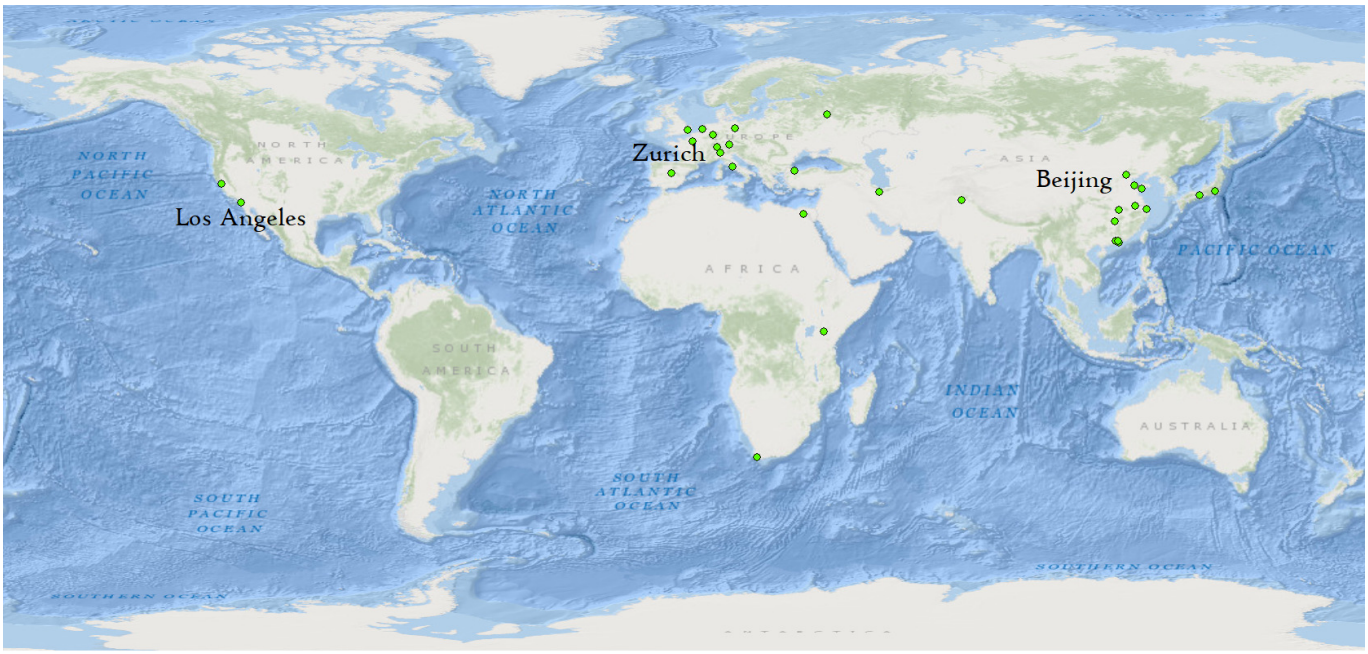


Fig. 2. The location of 31 cities covered by Gaofen-3 quad-pol mode in global scale (those cities are included in So2Sat LCZ42 list), the complete list of the 31 cities is in the Appendix. The background world map comes from ArcGIS ArcMap.

A. Gaofen-3 mission

Gaofen-3 is a spaceborne multi-polarization imaging radar mission in C-band [31]. The satellite is in a sun-synchronous dusk-dawn orbit at a 755 km altitude. The repeat cycle of orbit is about 29 days. In order to meet the technical requirements from various types of users, Gaofen-3 have twelve observing modes. The resolution varies from 1 m to 500 m, and the swath varies from 10 km to 650 km. The SAR payload can operate in single-pol (HH or VV), dual-pol (HH+HV or VH+VV), and quad-pol (HH+HV+VH+VV) modes. The quad-pol mode works in time-division operation architecture. Gaofen-3's quad-pol mode has 8m nominal resolution both in range and azimuth directions, which is the mode we use in this work. Its parameters are included in Table I. As Sentinel-1's dual-pol mode is adopted as the comparison setting, Table I also list the parameters of Sentinel-1's dual-pol mode.

B. Study area

In order to study the potential of quad-pol SAR data in LCZ classification, we collected Gaofen-3 quad-pol SAR data for cities in the So2Sat LCZ42 dataset [13]. So2Sat LCZ42 contains labeled data of 42 cities cross the world. This geographical distribution ensures that regions of interest include transcultural, transnational, and cross-environmental areas. While selecting these cities, population was another criteria under consideration. Among all cities, the population of each city is at least one million and is expected to grow in the future according to UN statistics. In total, 31 cities were selected in our Gaofen-3 study area. The other 11 cities were not covered by Gaofen-3 quad-pol mode. Fig.2 shows the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3137911, IEEE Transactions on Geoscience and Remote Sensing

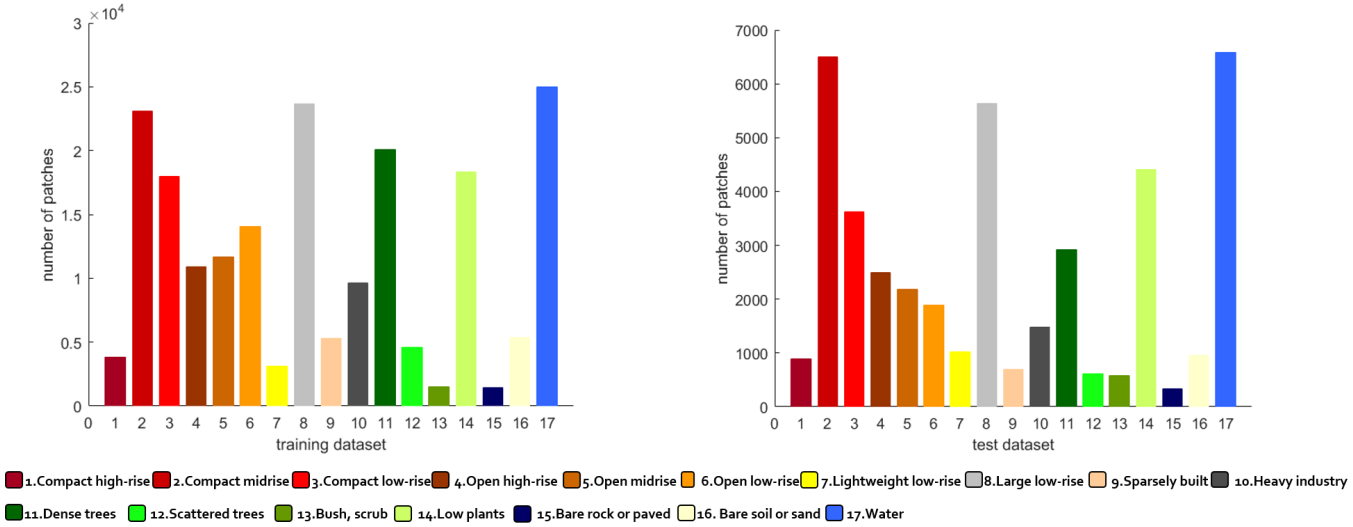


Fig. 3. For each city in the 31-cities list, we split the labels of each LCZ class into the west and east parts of the city, comprising the training and testing dataset. Then we get 200k patches of training dataset and 40k patches of testing dataset, which are geographically separated.

TABLE I
RADIOMETRIC ACCURACY AND SPECS BETWEEN GAOFEN-3 AND SENTINEL-1

SAR satellite	Absolute radiometric accuracy(dB)	Relative radiometric accuracy(dB)	Radiometric resolution(dB)	Spatial resolution in range direction(m)	Spatial resolution in azimuth direction(m)
Sentinel-1	1.0	0.1	3.0	3.0	22.0
Gaofen-3	1.5	1.0	3.5	8.0	8.0

geographic location of the 31 cities. The complete list of the 31 cities is in the Appendix. Considering Gaofen-3's average swath width of 30km in the quad-pol mode, we need 6 non-overlapping images to cover the whole urban and suburb area for most cities. In total, 179 quad-pol Gaofen-3 SAR images were collected for this study.

C. Training and testing dataset production

We employed the LCZ label in the So2Sat LCZ42 dataset to produce training and testing LCZ patches. The So2Sat LCZ42 labels determine the distribution of LCZ patches. So2Sat LCZ42 labels have been balanced by label post-processing procedures. Because the percentage of nonurban areas is naturally larger and results in many more samples for nonurban classes. To balance the number of samples among all of the LCZ classes, for each city, the samples of each nonurban classes are reduced [13]. By the way, the impact of LCZ imbalance has been analyzed based on Sentinel-2 dataset in So2Sat LCZ42 which uses the same ground truth as this paper. Experimental results by using class weights did not show obvious benefits [32]. This indicates that the imbalance problem in the So2Sat LCZ42 dataset has been addressed to some extent during the data preparation process. For one city, we split the labels of each LCZ class into the west and east parts, comprising the training and testing dataset respectively. Therefore, the training and testing dataset were geographically separated from each other.

In order to crop out Gaofen-3 patches by the corresponding So2Sat LCZ42 labels, we adopted orthorectification using RPC (Rational Polynomial Coefficient) in ENVI (The Environment for Visualizing Images) software to geocode each Gaofen-3 quad-pol SAR Single Look Complex (SLC) image. SRTM (Shuttle Radar Topography Mission) DEM was used in the geocoding. The data was then re-sampled to a 5-m GSD WGS84/UTM coordinate system by nearest-neighbor interpolation. In total, 200,000 training samples and 40,000 testing samples were generated. The patch size is 64×64 pixels. The distribution of LCZ classes in training and testing dataset are visualized in Fig.3. The postprocessed So2Sat LCZ42 labels controlled the number of patches of most the classes to be on the same order of magnitude. It is true that there still exists class imbalance after the label postprocessing. However, the imbalanced number of samples of different LCZs accords with land cover distribution. For instance, compact high-rise, lightweight low-rise and bare rock are comparatively less than other LCZ classes on the global scale.

Fig.1 shows the cropped patches from Gaofen-3 Pauli RGB image of different LCZ classes. According to the scattering mechanism of Pauli decomposition, RGB channel represents double bounce scattering, volume scattering and surface scattering, respectively. As double bounce scattering often reflects from dihedral angle, compact high-rise Gaofen-3 patches contain more red pixels than compact midrise and compact low-rise ones. For lightweight low-rise and sparsely built,

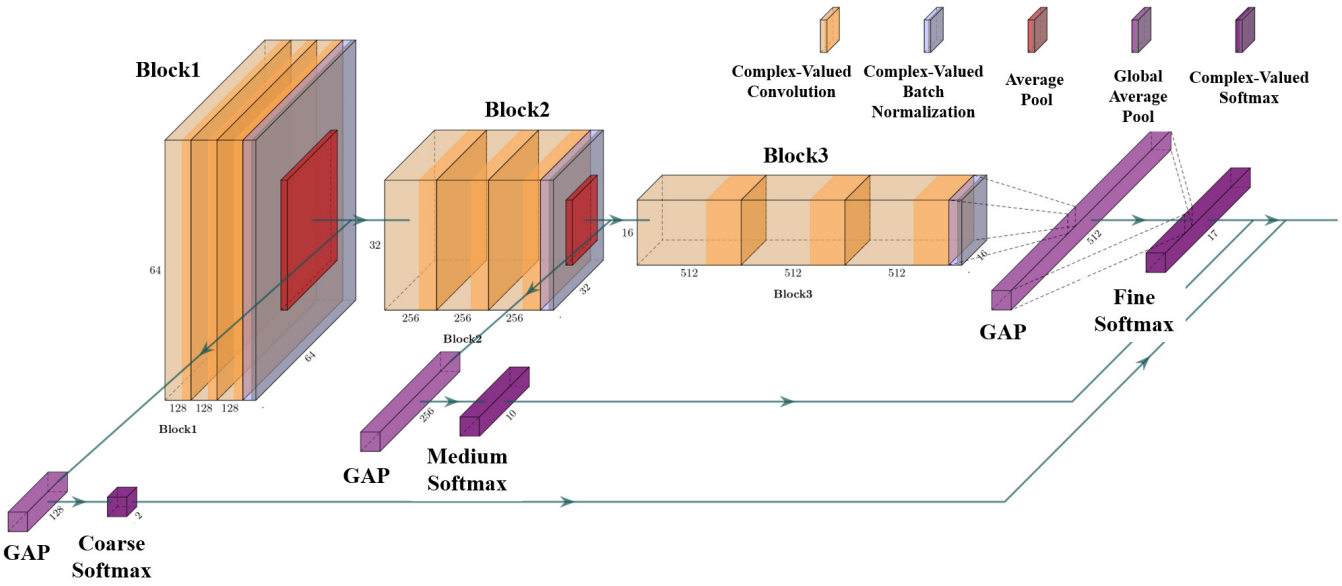


Fig. 4. Architecture of SAR4LCZ-Net: the loss function of SAR4LCZ-Net is a weighted summation of all output prediction losses, from coarse prediction to medium prediction and to fine prediction. The loss function takes all levels' loss into account to make sure the structure prior can play a role of internal guide to the whole model and make it easier to flow the gradients back to the shallow layers.

they are almost flat area and reflect as surface scattering in blue color. Oil collector of heavy industry are obvious red as double bounce scattering. It can be expected that the additional polarimetric information in the Ganfen-3 quad-pol data brings advantage in distinguishing certain land uses over single polarization or dual polarization data.

III. SAR4LCZ-NET DESCRIPTION

SAR4LCZ-Net is a complex-valued CNN that cooperates the hierarchical semantic meanings of the LCZ classification scheme. Shallower layers in the CNN capture textural features such as lines and corners whereas deeper layers extract features with more semantic meanings. Correspondingly, as shown in Table II, the 17 LCZ classes can be first grouped into two coarse classes: built area and nature area. In the second level of classes, the definition relies on areas where semantic meaning becomes more decisive, such as compact built and open built. Finally, the 17 LCZ classes are extracted explicitly.

A. SAR4LCZ-Net architecture

Fig.4 is the SAR4LCZ-Net architecture. As this work mainly focuses on validating the potential of complex-valued CNN using quad-pol SAR images for LCZ classification, three sequential convolutional blocks instead of sophisticated structure are embedded in SAR4LCZ-Net. It is also reported in [32] that, in LCZ classification, CNN with a simple architecture can be very competitive to the ones with complicated structures.

The size of input patches is 64×64. SAR4LCZ-Net contains three convolutional blocks (Block1, Block2, Block3). Each of the 3 blocks consists of three consecutive complex-valued convolutional layers and one complex-valued batch normalization layer, where all convolutional kernel sizes are 3×3. Block1 has convolutional layers with 128 kernels, features are downsampled by average pool layer with a pooling size of 2×2 and stride of 1. The output feature maps size is 32×32 from Block1. Each convolutional layer of Block2 have 256 kernels. After batch normalization and average pooling, feature maps have size of 16×16. The number of kernels in each convolutional layer in Block3 is 512.

SAR4LCZ-Net has three hierarchical predictions. Coarse prediction comes from Block1, after connecting one global average layer and 2-class softmax. Medium prediction flows out from Block2, after connecting one global average layer and 10-class softmax. Fine prediction is from Block3, also downsampled by one global average layer and activated by a 17-class softmax. The global loss function is a weighted average of all levels' loss.

The core components of SAR4LCZ-Net include complex-valued convolutional layer, complex-valued batch normalization layer, average pool layer, global average pool layer and complex-valued softmax output layer. We will describe them in the following subsections.

B. Complex-valued components of the network

1) Complex-valued convolutional layer. A complex kernel $W = A + iB$ is convolved by a complex vector $h = x + iy$ where A and B are real matrices and x and y are real vectors. As the convolution operator is distributive, convolution of the vector h with the filter W is obtained:

$$W * h = (A * x - B * y) + i(B * x + A * y) \qquad (1)$$

Considering a typical real-valued 2D convolutional layer that has N feature maps such that N is divisible by 2. To represent these as complex numbers, the first $N/2$ feature maps are allocated to represent the real components and the remaining $N/2$ to represent the imaginary ones. Then the

TABLE II
HIERARCHICAL STRUCTURE IN LCZ CLASSIFICATION

Coarse Prediction	Medium Prediction	Fine Prediction
Built area	Compact built	Compact high-rise
		Compact mid-rise
		Compact low-rise
	Open built	Open high-rise
		Open mid-rise
		Open low-rise
	Lightweight low-rise	Lightweight low-rise
	Large low-rise	Large low-rise
	Sparsely built	Sparsely built
	Heavy industry	Heavy industry
Nature area	Vegetation	Dense trees
		Scattered trees
		Bush, scrub
		Low plants
	Bare rock or paved	Bare rock or paved
	Bare soil or sand	Bare soil or sand
	Water	Water

weight tensor size of complex-valued convolution layer can be reduced to half of that in real-valued convolution layer.

2) Complex-valued activation function. Three complex-valued activation functions having good performance in other reported experiments are investigated in this work. Complex ReLU ($CReLU$) [28] is a complex activation that applies separate ReLUs on both of the real and the imaginary part of a neuron:

$$CReLU(z) = ReLU(Re(z)) + iReLU(Im(z)) \quad (2)$$

$ZReLU$ [28] is similar to $CReLU$ and is defined as follows:

$$ZReLU(z) = \begin{cases} z, & if\ \theta_z \in [0, \pi/2] \\ 0, & otherwise \end{cases} \quad (3)$$

$Cardioid$ is proposed in [33], which is sensitive to the input phase rather than the input magnitude. The output magnitude is attenuated based on the input phase, while the output phase remains equal to the input phase. The complex cardioid is defined as:

$$cardioid(z) = \frac{1}{2}(1 + cos(\phi z))z \quad (4)$$

3) Complex-valued batch normalization. Standard mean variance normalization does not ensure equal variance in both the real and imaginary components. In order to guarantee that the resulting distribution is circular, we scale the data by the square root of their variances along each of the real and imaginary components [28]. This can be done by multiplying the zero-centered data by the inverse square root of the covariance matrix V.

$$\tilde{x} = (V)^{-\frac{1}{2}}(x - E[x]) \quad (5)$$

where the covariance matrix V is

$$\begin{aligned} V &= \begin{pmatrix} V_{rr} & V_{ri} \\ V_{ir} & V_{ii} \end{pmatrix} \\ &= \begin{pmatrix} Cov(Re(x), Re(x)) & Cov(Re(x), Im(x)) \\ Cov(Im(x), Re(x)) & Cov(Im(x), Im(x)) \end{pmatrix} \end{aligned} \quad (6)$$

4) Complex-valued softmax output layer. In real-valued CNN, the output layer is usually a softmax classifier predicting the probability distribution over different classes. Then, the entire network is learned by minimizing the cross-entropy loss function. Softmax [34] is hardly applied to complex-valued CNN directly, previous complex-valued CNN uses softmax on the amplitude of complex-valued input and adopts the least-squares loss function. In this work, softmax classifier is adopted in real units and imagery units separately. The output is the mean value of the real and imagery softmax classifier.

$$\begin{aligned} F(n) = \frac{1}{2}[&SoftMAX(Re(F(n-1))) \\ &+SoftMAX(Im(F(n-1)))] \end{aligned} \quad (7)$$

The SoftMAX is defined as:

$$p_c(x) = \frac{exp(a_c(x))}{\sum_{i=1}^{C} exp(a_i(x))} \quad (8)$$

where a_i denotes the activation in the i th feature channel, C is the number of LCZ categories which are 2 in the coarse prediction, 10 in the medium prediction and 17 in the fine prediction. $p_c(x)$ is the approximated maximum function, which represents the probability for each category. The cross-entropy calculates the loss function:

$$L = -\sum_{i=1}^{C} p_l(x) log(p_c(x)) \quad (9)$$

where $p_l(x)$ is the ground truth label and $p_c(x)$ is the predicted label.

IV. EXPERIMENTS AND DISCUSSIONS

A. Input features

Quad-pol SLC image pixel is expressed by a complex scattering matrix:

$$S = \begin{Bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{Bmatrix} \quad (10)$$

where H, V denotes the orthogonal horizontal and vertical polarization bases. Spatial features in SAR images are not obvious when real and imaginary parts are represented separately. The amplitudes of four elements in scattering matrix can better describe spatial textures: $|S_{HH}|$,$|S_{HV}|$,$|S_{VH}|$ and $|S_{VV}|$. In order to fit the complex-valued input of SAR4LCZ-Net, spatial features are rebuilt as:

$$Sc = \begin{Bmatrix} Sc_1 \\ Sc_2 \end{Bmatrix} = \begin{Bmatrix} |S_{HH}| + j * |S_{HV}| \\ |S_{VH}| + j * |S_{VV}| \end{Bmatrix} \quad (11)$$

Fig.5 shows the complex-valued spatial features, it can be seen that phase value of Sc_1 and Sc_2 show clear spatial pattern.

In monostatic case, S_{HV} and S_{VH} are often assumed identical under the reciprocity theorem. Thus, the matrix can

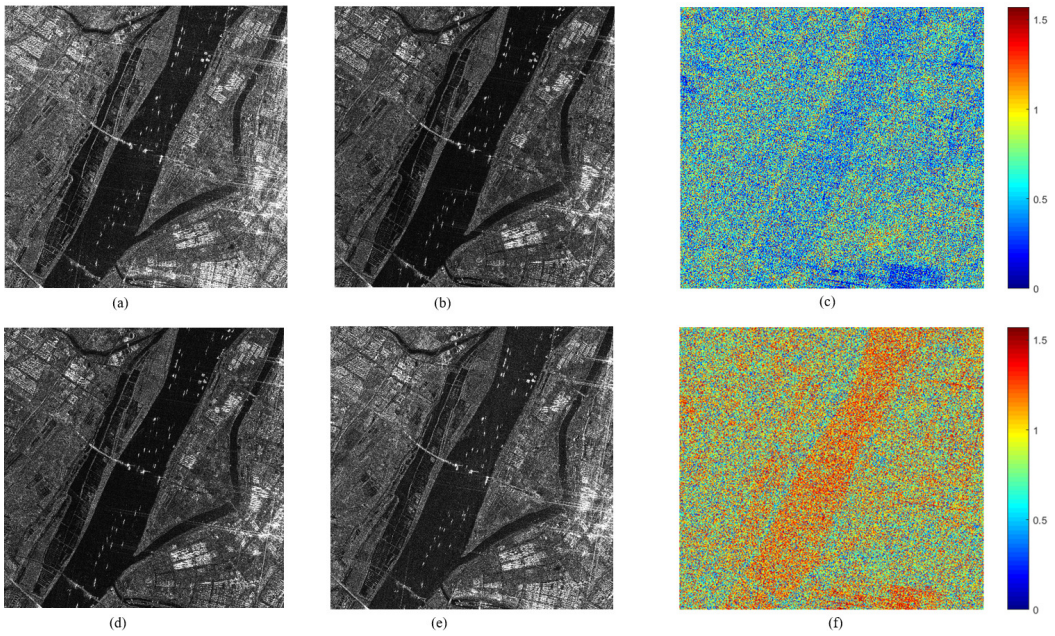


Fig. 5. Equation (11) Combines spatial features into complex-valued input for SAR4LCZ-Net: $Sc1$ consists of $|S_{HH}|$ and $|S_{HV}|$, $Sc2$ consists of $|S_{VH}|$ and $|S_{VV}|$. Spatial pattern in phase image show that the combination is feasible for SAR4LCZ-Net. (a) $|S_{HH}|$,(b) $|S_{HV}|$,(c) phase of $Sc1$ (in radians), (d) $|S_{VH}|$,(e) $|S_{VV}|$,(f) phase of $Sc2$ (in radians).

be reduced to a 3-D scattering vector K. Using the Pauli decomposition, K can be expressed as:

$$K = \frac{1}{\sqrt{2}}\{S_{HH} + S_{HV}, S_{HH} - S_{HV}, 2S_{HV}\}^T \tag{12}$$

The coherency matrix of K is obtained as:

$$C = KK^H = \begin{Bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{Bmatrix} \tag{13}$$

The superscript T denotes the transpose and H denotes the conjugate transpose. The coherency matrix C is a Hermitian matrix whose diagonal elements are real number. The off-diagonal complex elements are the polarimetric features. In order to reduce the speckle on quad-pol SAR images, the coherency matrix is filtered by a sequential approach [35].

Combination of features extracted from scattering matrix and coherency matrix in quad-pol SAR data have been verified effective for terrain classification [27], [36]. We also assemble polarimetric features and spatial features as input features for SAR4LCZ-Net. Here, polarimetric features are up-diagonal complex-valued elements from the polarimetric coherency matrix as shown in the last paragraph. The scattering matrix is not filtered in order to retain the original spatial textures.

In short, the input features consist of the denoised polarimetric features C_{12}, C_{13}, C_{23}, and the original spatial features Sc_1, Sc_2.

B. Experimental Settings and Implementation Details

In order to validate the effectiveness of the proposed SAR4LCZ-Net, we set a series of comparative experiments. The baseline setup is a Random Forest experiment. We also would like to affirm if complex-valued CNN can have better performance with less learning parameters in LCZ classification. Therefore, the second setting is a real-valued CNN with the same structure. The number of learning parameters of the real-valued CNN are twice of those in complex-valued CNN.

In addition, we also set some ablation experiments to verify the feasibility of SAR4LCZ-Net. They are listed as follows.

1) Complex-valued softmax output layer: In order to know the effectiveness of designed complex softmax output, we set two identical complex-valued CNNs, except their output layers. One uses output without activation by softmax layer and applies MSE as loss function, and the other uses designed complex softmax output as shown in Formula (7) and makes use of cross-entropy as loss function.

2) Complex-valued activation function: We set three complex-valued activation functions for convolutional layer: CReLU, Cardioid and ZReLU. This setting is to confirm if CReLU is the most suitable activation function for LCZ classification. We also analyze the heatmaps from the intermediate convolutional layer to explain the experiments, which was not present in other applications using complex-valued CNN and the above mentioned activation functions.

3) Combination of spatial and polarimetric features: Most complex-valued CNNs for quad-pol SAR image classification only use polarimetric features. In this work, we combine spatial features with the polarimetric features as a comparative experiment to demonstrate the effect of our proposed input feature.

4) Hierarchical structure in LCZ classification: To prove that the designed LCZ hierarchical structure in SAR4LCZ-Net is valuable, we set a same complex-valued CNN without the hierarchical structure as an ablation experiment.

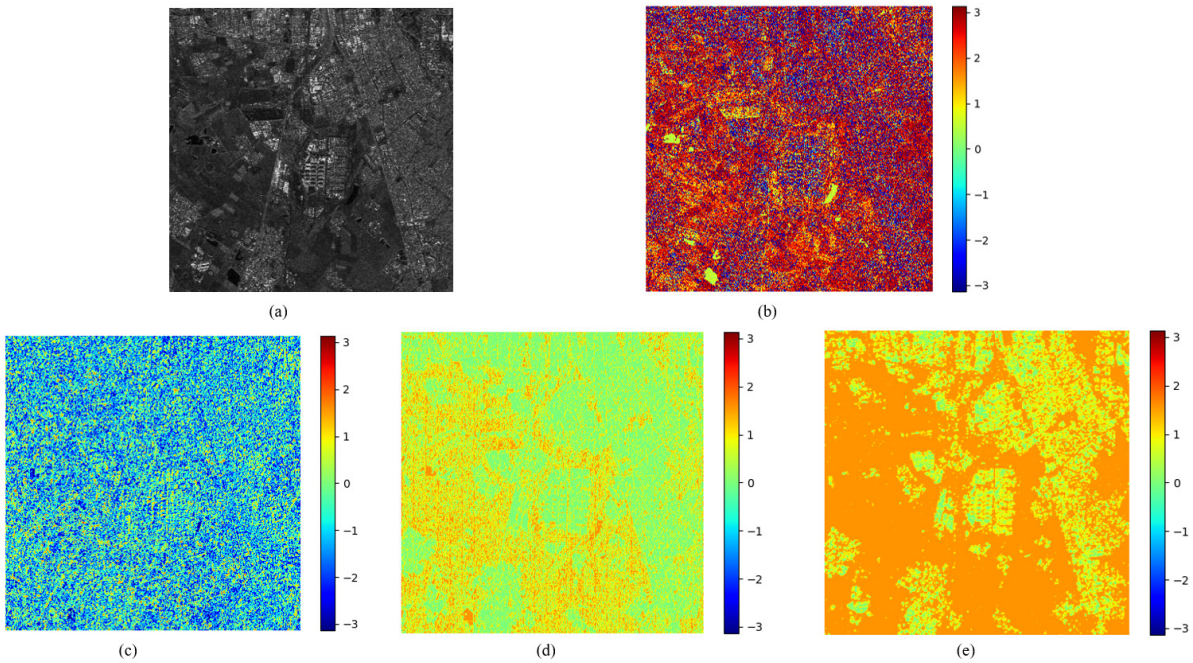


Fig. 6. The element C_{12} of the polarimetric coherency matrix as input element: (a) amplitude image, (b) original phase. After 5 complex-valued convolution layers activated by different functions: (c) cardioid, (d) ZReLU, (e) CReLU. CReLU has the most significant phase pattern comparing to those from other functions. The contour of built area is clearly outlined in CReLU phase pattern. Although all three activiation functions distorted the phases as expected, our main point is not to preserve the phases, but to extract high level features that are suitable for the classification task.

As for the implementation details, all models were trained from scratch following the same training settings in order to have fair comparisons. Keras was employed as the experimental software platform. All networks employed the Adam optimizer. We used a batch size of 128 patches. The initial learning rate was 10^{-4} and was decreased by half after every fifth epoch, and was kept as 10^{-5} after total 20 epochs. To control the training time and avoid overfitting, early stopping was activated based on the validation loss with a patience of 15 epochs. Model weights with the highest validation accuracy were saved. The distribution of the loss weights indicates the importance of each block. This can play a role as an internal guide to the whole model, which made the gradients easier to flow back to the shallow layers. [0.8, 0.1, 0.1] was assigned as the initialization loss weights, lower layers were frozen after 10 epochs at loss weights of [0, 0.9, 0.1], and medium layers were frozen after 60 epochs at loss weights of [0, 0, 1]. Metrics used for performance assessment include overall accuracy (OA), average accuracy (AA), and Kappa coefficient.

C. Results and Discussions

Table III summarizes the performances of SAR4LCZ-Net and some comparative experiments. Firstly, the OA of real-valued CNN (RV-CNN) (64.8%) is much higher than that of Random Forest (48.2%), which validates the excellent feature grasping ability of CNN. Besides, results show that complex-valued CNN (CV-CNN) can get 1.2% improvement in OA compared to RV-CNN with the same architecture and double the number of learning parameters. This can support that choosing CV-CNN as the backbone of SAR4LCZ-Net is beneficial. In the following paragraphs, we will analyze designed features in SAR4LCZ-Net.

1) Impact of complex-valued softmax output layer: In CV-CNNs with CReLU as activation functions, the improvements of applying complex-valued softmax output layer are evident. OA and AA both rise 4.0%, which is remarkable in LCZ classification. This proved the feasibility of proposed complex-valued softmax output layer.

Compared to the MSE as loss function adopting output without activation by softmax layer, cross-entropy using designed complex-valued softmax output are better loss function for 17-category classification. Cross-entropy typically optimizes a surrogate loss function instead, which acts as a proxy but has advantages. The negative log-likelihood of the correct class is typically used as a surrogate for the 0-1 loss. The negative log-likelihood allows the model to estimate the conditional probability of the classes, then the model can pick the classes that yield the least classification error in expectation. A surrogate loss function actually results in being able to learn more. When the expected 0-1 loss is zero, one can improve the robustness of the classifier by further pushing the classes apart from each other, obtaining a more confident and reliable classifier, thus extracting more information from the training data than would have been possible by simply minimizing the average 0-1 loss on the training set.

2) Validation on complex-valued activation function: Results from different kinds of complex-valued activation function show that CReLU achieves the best performance. To illustrate the effectiveness of complex-valued activation function, we visualize the phase pattern after complex convolution using those activation functions. Fig.6 shows the phase pattern of

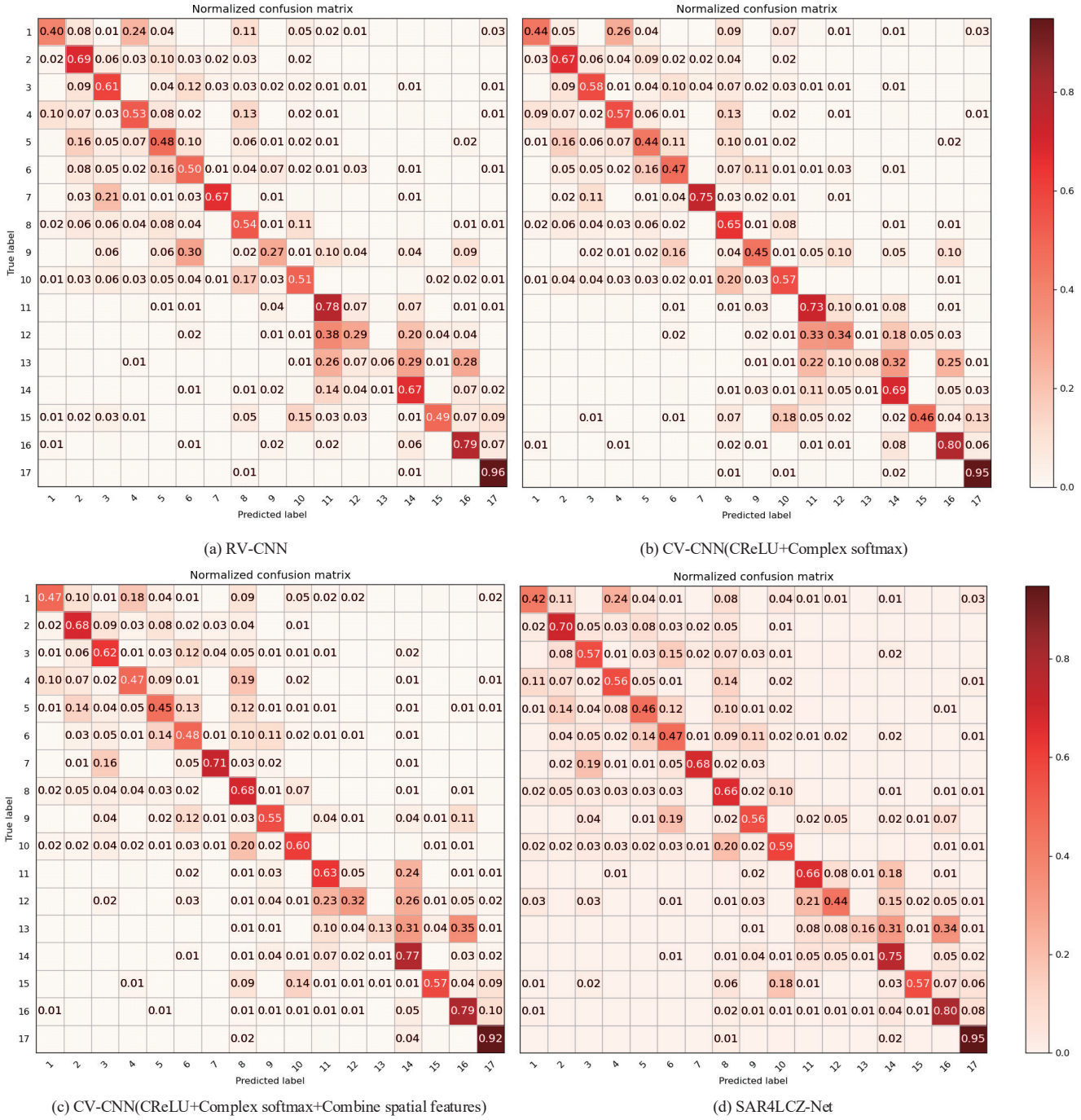


Fig. 7. Confusion Matrixes of SAR4LCZ-Net and the comparison networks based on the test dataset of Gaofen-3 quad-pol SAR patches.

C_{12} from the polarimetric coherency matrix in Formula (13) after 5 convolution layers. It is obvious that CReLU can grasp the phase pattern more effectively, which produces the most significant phase pattern comparing to those of other functions. The contour of built area is clearly outlined in the phase pattern of CReLU. Although all three activiation functions distorted the phases as expected, our main point is not to preserve the phases, but to extract high level features that are suitable for the classification task. Therefore, CReLU was selected as the activation function of SAR4LCZ-Net, and employed in the experiments hereafter.

3) Effectiveness on the combination of spatial and polarimetric features: When combining the designed spatial features, the AA was improved by 1.3% which is decent for 17 classes. This comparison validates the effectiveness of the proposed combination of spatial features and polarimetric features for LCZ classification.

4) Feasibility of hierarchical structure for LCZ classification: In the comparison between SAR4LCZ-Net and CV-CNN without the hierarchical structure, the OA and AA of SAR4LCZ-Net can both improve by 1%. The results demon-

TABLE III
COMPARATIVE EXPERIMENTS

Model	Polarimetric features	Parameters	Complex Relu	Complex softmax	Spatial features	Use Hierar-chical structure	OA	AA	Kappa
RF	Despeckling						48.2%	35.6%	0.422
ResNet-50	Despeckling	1.40M					52.0%	37.9%	0.466
RV-CNN	Despeckling	7.68M					64.8%	54.3%	0.611
CV-CNN	Despeckling	3.85M	CReLU				62.1%	52.6%	0.582
CV-CNN	Despeckling	3.85M	CReLU	✓			66.0%	56.6%	0.626
CV-CNN	Despeckling	3.85M	Cardioid	✓			65.0%	53.6%	0.612
CV-CNN	Despeckling	3.85M	ZReLU	✓			59.6%	49.7%	0.554
CV-CNN	Without Despeckling	3.85M	CReLU	✓			65.8%	56.2%	0.636
CV-CNN	Despeckling	3.85M	CReLU	✓	✓		66.5%	57.9%	0.630
SAR4LCZ-Net	Despeckling	3.85M	CReLU	✓	✓	✓	**67.2%**	**58.8%**	**0.638**

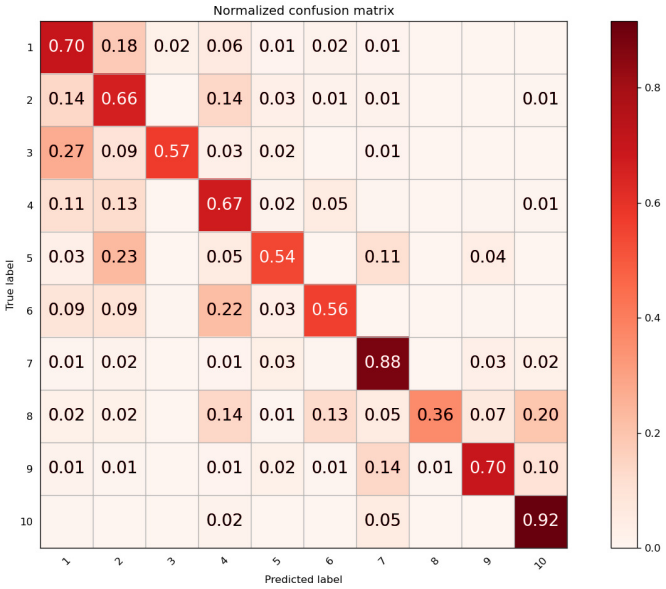


Fig. 8. Medium prediction results of SAR4LCZ-Net using Gaofen-3 quad-pol SAR test dataset: The aggregated classes such as vegetation, compact bulit and open built, can achieve relatively higher accuracy.

strate the feasibility of applying hierarchical structure for LCZ classification. The confusion matrixes of LCZ classification from SAR4LCZ-Net and the comparison networks are shown in Fig.7. Class 2, 7, 8 ,11, 14, 16 and 17 can achieve accuracy higher than 65%. The confusion matrix of the medium prediction from SAR4LCZ-Net is shown in Fig.8. Its OA and AA can reach 74.5% and 65.6% respectively. The classification accuracies of vegetation and open built can achieve 88% and 70%.

Fig.9 show the intermediate feature maps from SAR4LCZ-Net and RV-CNN in different stages. It can be seen that the SAR image covered urban areas in the lower center part and natural areas in the upper part. The hierarchical feature grasping ability is demonstrated through the feature maps flowing from Block1 to Block3, lower convolutional layers obtained the contour features and higher convolutional layers extracted abstract semantic features. The effectiveness of hierarchical structured network design for LCZ classification is verified. Besides, the feature attention on urban areas from SAR4LCZ-Net is more constraint, especially in the middle network after Block2 and deeper network after Block3. This can certify that SAR4LCZ-Net has better feature grasping ability compared to RV-CNN.

5) Comparison to Sentinel-1 dual-pol data: We also would like to compare the LCZ classification results between Gaofen-3 quad-pol SAR data and Sentinel-1 dual-pol SAR data. A strict quantitative comparison is not possible, because the coverages of our Gaofen-3 quad-pol SAR image is different from that of the So2Sat LCZ42 dataset. 11 cities were not covered. In addition, the spatial resolution, and many other parameters of the sensors are different. However, this is the best possible data preparation one could do to bring the comparison as fair as possible. The large amount of Gaofen-3 data in this study is sufficient to bring a qualitative comparison to Sentinel-1 data. Table IV shows the performances of Gaofen-3 and Sentinel-1. The LCZ classification accuracy from Sentinel-1 has an OA of 61.0% and an AA of 51.0%. Compared with Sentinel-1's results, Gaofen-3's OA and AA improve 6.2% and 7.8%, respectively. Fig.10 shows the LCZ map predictions in Beijing, Los Angeles and Zurich. It can be seen that compared with Sentinel-1, the results from Gaofen-3 are more spatially continuous and homogenous. In order to have a more detailed comparison, we zoom in to small areas of LCZ map predictions, and compare with the corresponding optical images from Google Earth. This shown in Fig.11. The small area in the Beijing LCZ map shows a wetland park area, where the waterbody, dense tree and the compact buildings around the park are predicted relatively more accurately in the Gaofen-3 LCZ map than those in the Sentinel-1 LCZ map. The small area in the Los Angeles LCZ map is a residential area.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3137911, IEEE Transactions on Geoscience and Remote Sensing

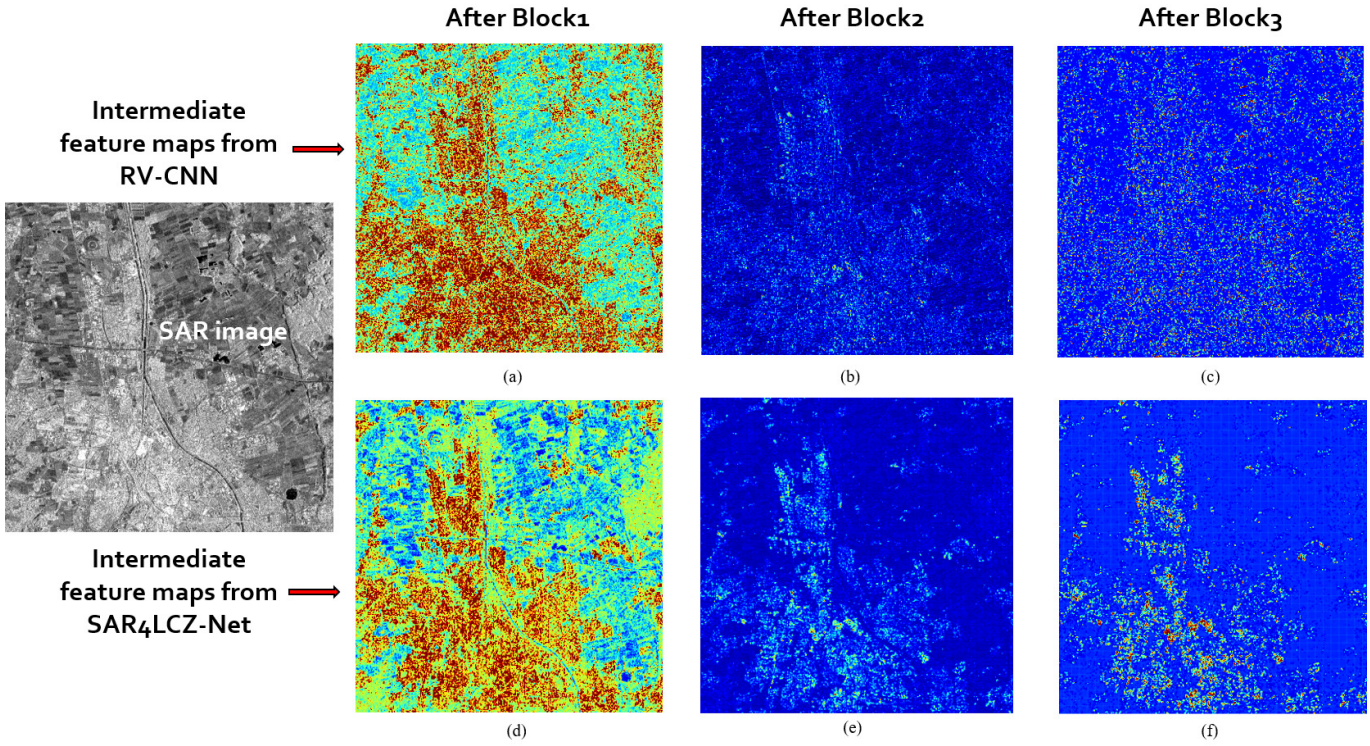


Fig. 9. Intermediate feature maps: Upper row are intermediate feature maps from RV-CNN, lower row are intermediate feature maps from SAR4LCZ-Net. Stages from left to right: Block1, Block2 and Block3.

TABLE IV
COMPARATIVE ACCURACY BETWEEN GAOFEN-3 AND SENTINEL-1

SAR satellite	OA	AA	Kappa
Sentinel-1	62.2%	51.6%	0.584
Gaofen-3	**67.2%**	**58.8%**	**0.638**

The compact low-rise areas are predicted more precisely by the Gaofen-3 LCZ map. The small area in Zurich LCZ map shows a residential area near forest. The dense trees in forest and open low-rise areas predicted by the Gaofen-3 LCZ map are closer to the ground truth.

Apart from Gaofen-3's generally better performance than Sentinel-1 primarily due to its higher spatial resolution, we also believe that Gaofen-3 quadpol data have unique advantage in distinguishing certain LCZ classes. Examples of the Gaofen-3 and Sentinel-1 image scenes of 5 LCZ classes are compared and shown in Fig.12. For compact midrise, the built area is mainly red in the Gaofen-3 Pauli image which consists of double bounce scattering. The flatten area in bare soil or sand appears blue in the Gaofen-3 Pauli image, which means they are surface scattering. Correspondingly, the classification accuracy of those 5 LCZ classes improves significantly by using quad-pol SAR data. Bare soil or sand achieves 52% higher accuracy in Gaofen-3 dataset compared with that in Sentinel-1 dataset. Open midrise, compact midrise, lightweight low-rise and scattered trees increase by 21%, 20%, 20% and 20%. This confirms the effectiveness of quad-pol SAR data for LCZ classification.

Of course such improvement may also be partially due to other factors such as higher spatial resolution and radiometric resolution. But it is more likely that polarimetric information helps the certain LCZ classes, as the image resolution or radiometric resolution improvements should be the same for all LCZ classes. The spatial resolution in range direction of Gaofen-3 quad-pol mode SLC data is 8m, which is lower than 3m of the Sentinel-1 dual-pol mode SLC data. In azimuth direction, Gaofen-3 quad-pol mode'8m is higher than Sentinel-1 dual-pol mode'22m. With respect to radiometric accuracy and specs, the parameters are listed in Table I and the Sentinel-1 radiometric performance is better. Therefore, the contributions of spatial resolution and radiometric performance are relatively balanced. A more thorough analysis shall be conducted to quantify the improvements caused by higher spatial resolution and the polarimetric bands.

6) General discussions: We also tried deeper CNN in this work such as ResNet-50 which is frequently used in computer vision. But LCZ classification results show that a properly designed simple CNN leads better performance as illustrated in Table III. A possible reason is that the patch sizes in classification tasks in remote sensing especially SAR images are comparatively smaller than those in computer vision research. For instance, the patch size in this work is 64×64 while the patch size of face recognition is normally 256×256. Therefore, shallower CNN is sufficient to grasp features from small patches. A deeper network is much easier to be overfitted and has worse generalizability.

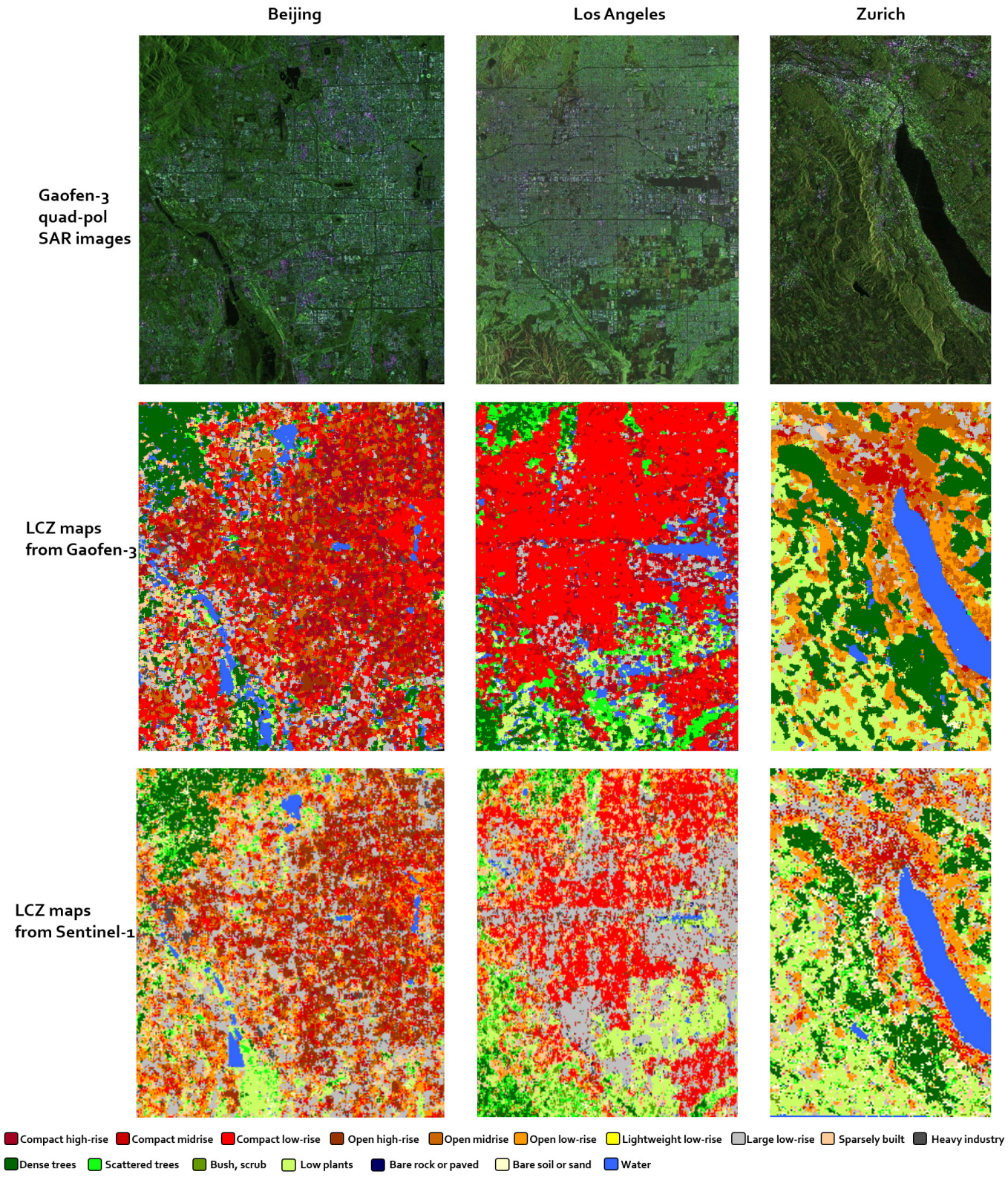


Fig. 10. LCZ map of Cities: Top row are Gaofen-3 quad-pol SAR Pauli RGB images, middle row are Gaofen-3's LCZ classification maps, and bottom row are the corresponding LCZ classification maps from Sentinel-1. Cities from left to right: Beijing, Los Angeles and Zurich.

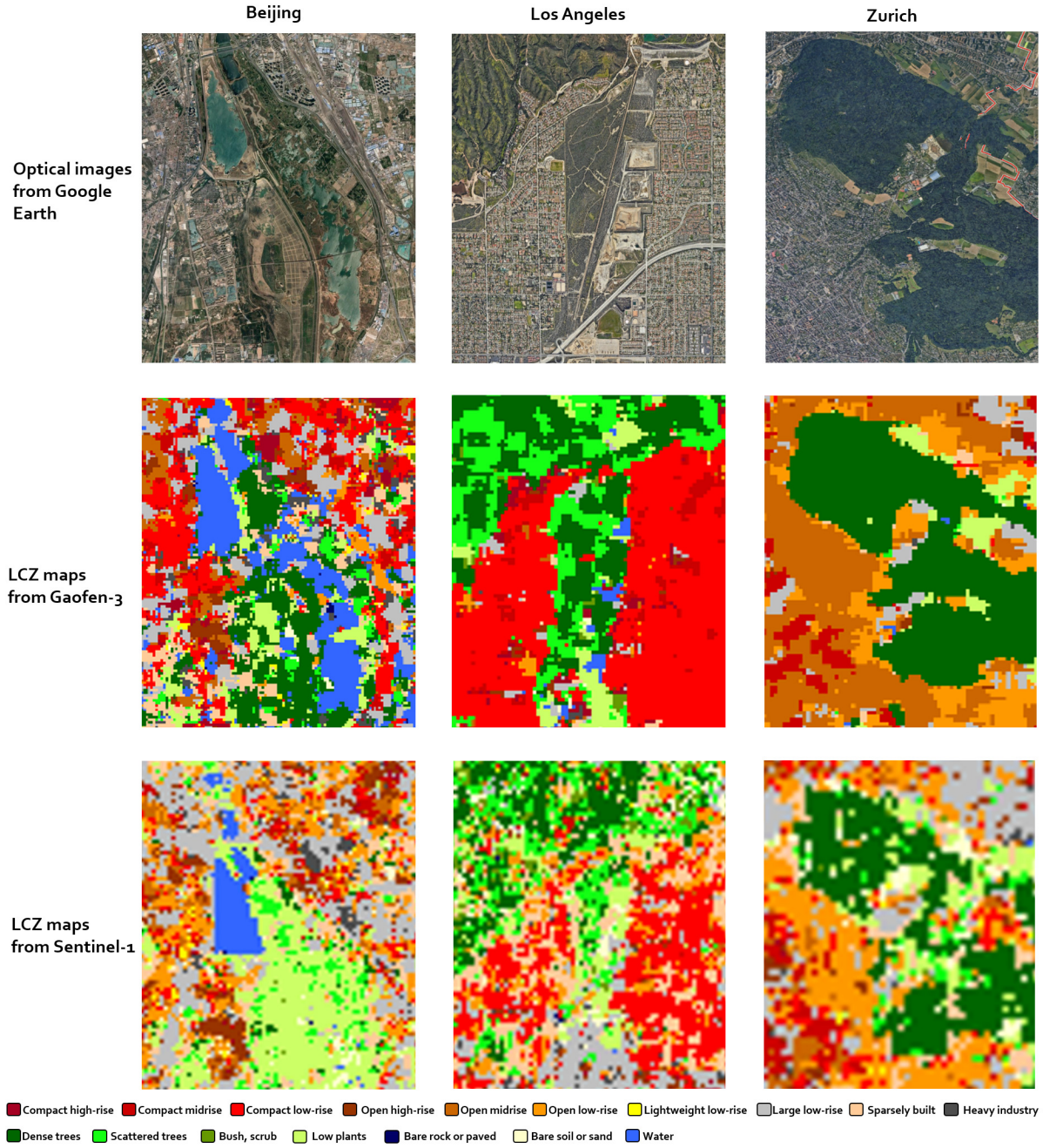


Fig. 11. LCZ map of small areas: Top row are optical images, middle row are Gaofen-3's LCZ classification maps, and bottom row are the corresponding LCZ classification maps from Sentinel-1. Cities from left to right: the wetland park area in Beijing, the residential areas in Los Angeles and Zurich.

The proposed complex-valued network SAR4LCZ-Net can be applied to single-channel SAR images for LCZ classification. As SAR4LCZ-Net are complex-valued structure with complex-valued softmax output layer and activation function, single-channel SAR images can be fed into the network so long as the input consists of real part and imaginary part. Besides, the structure associated with hierarchical semantic meaning of LCZ classes remains effective even though there is only one channel to flow the features. Therefore, the proposed SAR4LCZ-Net has ubiquity for LCZ classification when the adopted SAR datasets are in single channel.

V. CONCLUSION

We propose a complex-valued CNN, SAR4LCZ-Net, which is designed to take full advantage of the four polarization channels in quad-pol SAR images for LCZ classification. We introduced two novel improvements to the conventional LCZ classification network: hierarchical architecture, and complex-valued network realized by CReLU activation function and a complex-valued softmax layer. Comparing to a same network without hierarchical structure, SAR4LCZ-Net can achieve 1% improvements on overall accuracy and average accuracy using the big dataset of Gaofen-3 quad-pol SAR images that covers

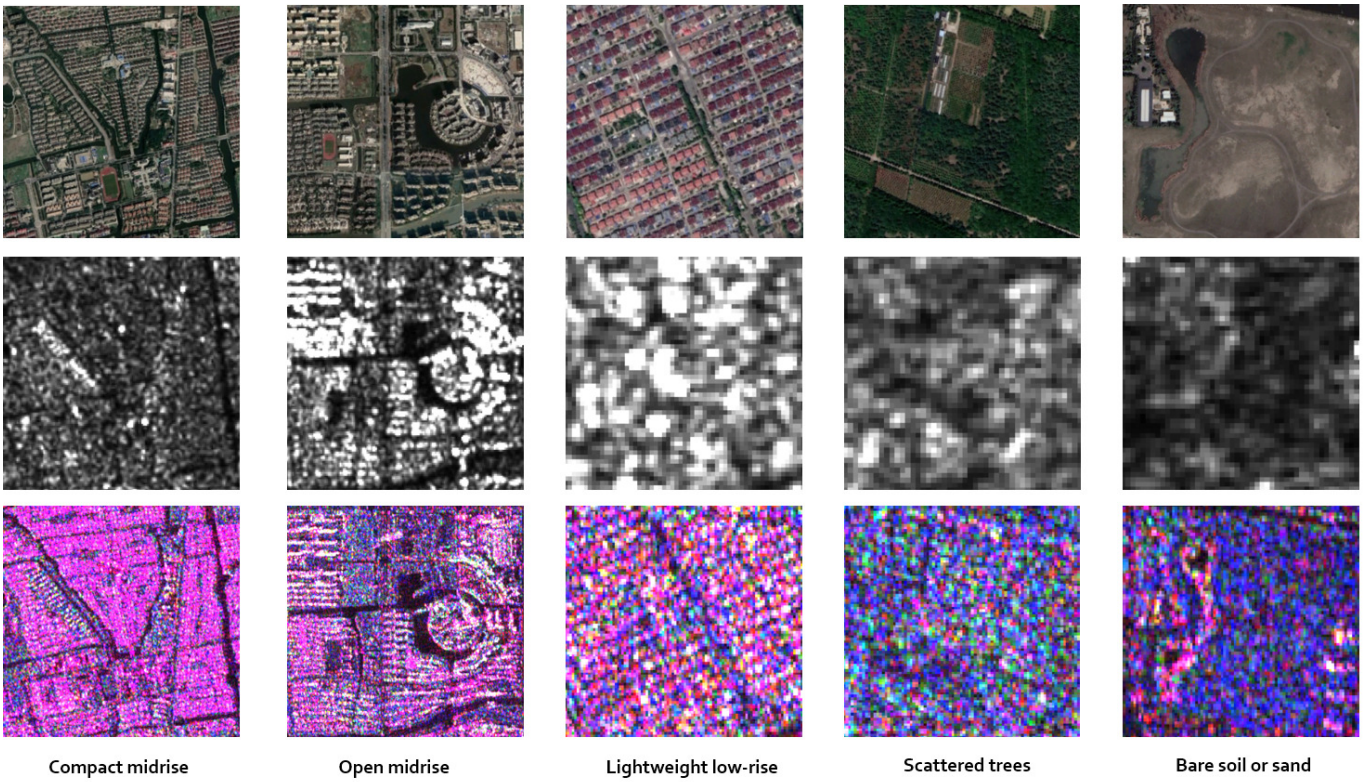


Fig. 12. Representative examples of the Sentinel-1 and Gaofen-3 image scenes of certain LCZ classes. In each example, the upper image is the high resolution aerial image from Google as a reference, the middle one is the intensity of the Sentinel-1 scene, and the lower image is the corresponding Gaofen-3 scene in Pauli RGB. This figure shows the typical urban morphology of LCZ classes which scatter more differently in quad-pol SAR images.

31 cities around the world.

Different kinds of complex-valued activation function were also analyzed in the study. CReLU was validated as the optimal complex-valued activation function for our task and adopted in SAR4LCZ-Net. A complex-valued softmax output layer is invented in order to use cross-entropy loss function for LCZ classification. This complex implementation also brings an improvement of overall accuracy by another 3.9%. In complex-valued CNN, CReLU is recommended to be the activation function in convolutional layer.

Comparing to most quad-pol SAR classification works utilizing polarimetric features from the coherency matrix, our approach combines spatial features from the scattering matrix with polarimetric features as the input, which improved 1.3% on average accuracy.

Besides, Gaofen-3 quad-pol SAR data also outperforms Sentinel-1 dual-pol SAR data in LCZ classification.

APPENDIX

The 31 cities selected in this study: Amsterdam, Beijing, Berlin, Cairo, Cape Town, Changsha, Cologne, Dongying, Guangzhou, Hong Kong, Islamabad, Istanbul, Kyoto, London, Los Angeles, Madrid, Milan, Moscow, Munich, Nairobi, Nanjing, Paris, Qingdao, Rome, San Francisco, Shanghai, Shenzhen, Tokyo, Tehran, Wuhan, Zurich.

REFERENCES

[1] I. D. Stewart and T. R. Oke, "Local climate zones for urban temperature studies," *Bulletin of the American Meteorological Society*, vol. 93, no. 12, pp. 1879–1900, 2012.

[2] B. Bechtel, P. J. Alexander, J. Böhner, J. Ching, O. Conrad, J. Feddema, G. Mills, L. See, and I. Stewart, "Mapping local climate zones for a worldwide database of the form and function of cities," *ISPRS International Journal of Geo-Information*, vol. 4, no. 1, pp. 199–219, 2015.

[3] J. A. Quanz, S. Ulrich, D. Fenner, A. Holtmann, and J. Eimermacher, "Micro-scale variability of air temperature within a local climate zone in berlin, germany, during summer," *Climate*, vol. 6, no. 1, p. 5, 2018.

[4] A. Wicki and E. Parlow, "Attribution of local climate zones using a multitemporal land use/land cover classification scheme," *Journal of Applied Remote Sensing*, vol. 11, no. 2, p. 026001, 2017.

[5] H. C. Ho, K. K.-L. Lau, R. Yu, D. Wang, J. Woo, T. C. Y. Kwok, and E. Ng, "Spatial variability of geriatric depression risk in a high-density city: A data-driven socio-environmental vulnerability mapping approach," *International journal of environmental research and public health*, vol. 14, no. 9, p. 994, 2017.

[6] J. Hu, Y. Wang, H. Taubenböck, and X. X. Zhu, "Land consumption in cities: A comparative study across the globe," *Cities*, vol. 113, p. 103163, 2021.

[7] J. Hu, P. Ghamisi, and X. X. Zhu, "Feature extraction and selection of sentinel-1 dual-pol data for global-scale local climate zone classification," *ISPRS International Journal of Geo-Information*, vol. 7, no. 9, p. 379, 2018.

[8] O. Danylo, L. See, B. Bechtel, D. Schepaschenko, and S. Fritz, "Contributing to wudapt: A local climate zone classification of two cities in ukraine," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 5, pp. 1841–1853, 2016.

[9] Y. Xu, C. Ren, M. Cai, N. Y. Y. Edward, and T. Wu, "Classification of local climate zones using aster and landsat data for high-density cities," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 7, pp. 3397–3405, 2017.

[10] J. Ching, G. Mills, B. Bechtel, L. See, J. Feddema, X. Wang, C. Ren, O. Brousse, A. Martilli, M. Neophytou, *et al.*, "Wudapt: An urban weather, climate, and environmental modeling infrastructure for the anthropocene," *Bulletin of the American Meteorological Society*, vol. 99, no. 9, pp. 1907–1924, 2018.

[11] B. Bechtel, P. J. Alexander, C. Beck, J. Böhner, O. Brousse, J. Ching, M. Demuzere, C. Fonte, T. Gál, J. Hidalgo, *et al.*, "Generating wudapt level 0 data–current status of production and evaluation," *Urban climate*, vol. 27, pp. 24–45, 2019.

[12] B. Bechtel, L. See, G. Mills, and M. Foley, "Classification of local climate zones using sar and multispectral data in an arid environment," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 7, pp. 3097–3105, 2016.

[13] X. Zhu, R. Huang, L. H. Hughes, H. Li, and Y. Hua, "So2sat lcz42: A benchmark dataset for global local climate zones classification," *IEEE Geoscience and Remote Sensing Magazine*, vol. 8, no. 3, 2020.

[14] H. Jing, Y. Feng, W. Zhang, Y. Zhang, S. Wang, K. Fu, and K. Chen, "Effective classification of local climate zones based on multi-source remote sensing data," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 2666–2669, IEEE, 2019.

[15] P. Feng, Y. Lin, J. Guan, Y. Dong, G. He, Z. Xia, and H. Shi, "Embranchment cnn based local climate zone classification using sar and multispectral remote sensing data," in *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 6344–6347, IEEE, 2019.

[16] Y. Zhou, H. Wang, F. Xu, and Y.-Q. Jin, "Polarimetric sar image classification using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1935–1939, 2016.

[17] L. Wang, X. Xu, H. Dong, R. Gui, and F. Pu, "Multi-pixel simultaneous classification of polsar image using convolutional neural networks," *Sensors*, vol. 18, no. 3, p. 769, 2018.

[18] C. Jiao, X. Wang, S. Gou, W. Chen, D. Li, C. Chen, and X. Li, "Self-paced convolutional neural network for polsar images classification," *Remote Sensing*, vol. 11, no. 4, p. 424, 2019.

[19] R. Yang, Z. Hu, Y. Liu, and Z. Xu, "A novel polarimetric sar classification method integrating pixel-based and patch-based classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 3, pp. 431–435, 2019.

[20] Y. Seo and K.-s. Shin, "Hierarchical convolutional neural networks for fashion image classification," *Expert systems with applications*, vol. 116, pp. 328–339, 2019.

[21] X. Zhu and M. Bain, "B-cnn: Branch convolutional neural network for hierarchical classification," 2017.

[22] Z. Yan, H. Zhang, R. Piramuthu, V. Jagadeesh, D. DeCoste, W. Di, and Y. Yu, "Hd-cnn: hierarchical deep convolutional neural networks for large scale visual recognition," in *Proceedings of the IEEE international conference on computer vision*, pp. 2740–2748, 2015.

[23] L. Zhang, H. Dong, and B. Zou, "Efficiently utilizing complex-valued polsar image data via a multi-task deep learning framework," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 157, pp. 59–72, 2019.

[24] Z. Zhang, H. Wang, F. Xu, and Y.-Q. Jin, "Complex-valued convolutional neural network and its application in polarimetric sar image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 7177–7188, 2017.

[25] Y. Duan, N. Chen, and Y. Chen, "A novel polsar image classification method based on optimal polarimetric features and contextual information," *Canadian Journal of Remote Sensing*, vol. 45, no. 6, pp. 795–813, 2019.

[26] C. Yang, B. Hou, B. Ren, Y. Hu, and L. Jiao, "Cnn-based polarimetric decomposition feature selection for polsar image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8796–8812, 2019.

[27] F. Gao, T. Huang, J. Wang, J. Sun, A. Hussain, and E. Yang, "Dual-branch deep convolution neural network for polarimetric sar image classification," *Applied Sciences*, vol. 7, no. 5, p. 447, 2017.

[28] C. Trabelsi, O. Bilaniuk, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. Pal, "Deep complex networks," *ArXiv*, vol. abs/1705.09792, 2018.

[29] E. K. Cole, J. Y. Cheng, J. M. Pauly, and S. S. Vasanawala, "Analysis of deep complex-valued convolutional neural networks for mri reconstruction," *arXiv preprint arXiv:2004.01738*, 2020.

[30] P. Virtue, *Complex-valued deep learning with applications to magnetic resonance image synthesis*. eScholarship, University of California, 2019.

[31] J. Sun, W. Yu, and Y. Deng, "The sar payload design and performance for the gf-3 mission," *Sensors*, vol. 17, no. 10, p. 2419, 2017.

[32] C. Qiu, X. Tong, M. Schmitt, B. Bechtel, and X. Zhu, "Multilevel feature fusion-based cnn for local climate zone classification from sentinel-2 images: Benchmark results on the so2sat lcz42 dataset," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 2793–2806, 2020.

[33] P. Virtue, X. Y. Stella, and M. Lustig, "Better than real: Complex-valued neural nets for mri fingerprinting," in *2017 IEEE international conference on image processing (ICIP)*, pp. 3953–3957, IEEE, 2017.

[34] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-margin softmax loss for convolutional neural networks.," in *ICML*, vol. 2, p. 7, 2016.

[35] Y. Cui, Y. Yamaguchi, H. Kobayashi, and J. Yang, "Filtering of polarimetric synthetic aperture radar images: A sequential approach," in *2012 IEEE International Geoscience and Remote Sensing Symposium*, pp. 3138–3141, IEEE, 2012.

[36] J. Shi, H. Jin, and X. Li, "A novel multi-feature joint learning method for fast polarimetric sar terrain classification," *IEEE Access*, vol. 8, pp. 30491–30503, 2020.

Rui Zhang received the B.Sc. degree in underwater acoustic engineering from the Northwestern Polytechnical University, Xi'an, China, in 2016. Afterwards, he started pursuing the M.Sc. degree in electronic information engineering from Beihang University, Beiing, China, since 2016, and approved to be a doctor candidate since 2018.

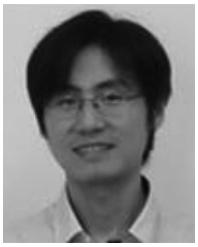
He was a visiting researcher in the Technical University of Munich (TUM), Munich, Germany, under the funding from TUM-Beihang doctoral exchange program and CSC program, since March 2020 to April 2021. His research interests include deep learning, polarimetric SAR image interpretation, SAR calibration.

Yuanyuan Wang (Member, IEEE) received the B.Eng. degree (Hons.) in electrical engineering from The Hong Kong Polytechnic University, Hong Kong, in 2008, and the M.Sc. and Dr. Ing. degrees from the Technical University of Munich (TUM), Munich, Germany, in 2010 and 2015, respectively. In June and July 2014, he was a Guest Scientist with the Institute of Visual Computing, ETH Zürich, Zürich, Switzerland.

He is currently a guest professor at the International AI Future Lab: Artificial Intelligence for Earth Observation, at the Department of Aerospace and Geodesy, Technical University of Munich, also with the Department of EO Data Science, Remote Sensing Technology Institute, German Aerospace Center, Weßling, Germany, where he leads the working group Big SAR Data. His research interests include optimal and robust parameters estimation in multibaseline InSAR techniques, multisensor fusion algorithms of synthetic aperture radar (SAR) and optical data, nonlinear optimization with complex numbers, machine learning in SAR, and high-performance computing for big data. Dr. Wang was one of the best reviewers of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2016.

Jingliang Hu received the B.Sc. degree in geographic information system from the Chengdu University of Technology, Chengdu, China, in 2011, the M.Sc. degree in cartography and geographic information system from Wuhan University, Wuhan, China, in 2014, and the M.Sc. degree in space science and technology from the Technical University of Munich (TUM), Munich, Germany, in 2014, and Ph.D. degree in signal processing in earth observation from the Technical University of Munich (TUM), Munich, Germany, in 2020.

Since 2020, he has been a Postdoctoral researcher with the Data Science in Earth Observation (SiPEO), Technical University of Munich (TUM), Munich, Germany. His research interests include topological data analysis (TDA), data fusion, machine learning, and deep learning.



Wei Yang received the M.S. and Ph.D. degrees in signal and information processing from Beihang University (previously known as Beijing University of Aeronautics and Astronautics (BUAA)), Beijing,China, in 2008 and 2011, respectively. In 2005, he studied the inner calibration signal analysis in SAR system. From 2006 to 2010, he focused on the system performance analysis and signal processing of high-resolution and wide swath mode in spaceborne SAR. From 2011 to 2013, he was a Postdoctoral Fellow with the School of Electronics and Information Engineering, BUAA. From 2013 to 2019, he was a Lecturer. Since July 2019 he has been an associate professor. He had been a Visiting Researcher in the Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, U.K. He has authored and coauthored more than 40 journal and conference publications. His current research interests include moving target detection, high-resolution spaceborne SAR image formation, and 3-D imaging.

Jie Chen received the B.S. and Ph.D. degrees in information and communication engineering from Beihang University (previously known as Beijing University of Aeronautics and Astronautics (BUAA)), Beijing, China, in 1996 and 2002,respectively. His Ph.D. study was focused on novel spaceborne synthetic aperture radar (SAR) system modeling and signal processing. From 2004, he was an Associate Professor with the School of Electronics and Information Engineering, BUAA.

He was a Visiting Researcher with the School of Mathematics and Statistics, University of Sheffield, Sheffield, U.K., from 2009 to 2010, working with Prof. S. Quegan on ionospheric effects on low-frequency space radars that measure forest biomass and ionospheric electron densities. He has been a Full Professor with the School of Electronics and Information Engineering, BUAA, since 2011. He has published more than 70 journal and conference papers, and he is the holder of 12 patents in the field of microwave remote sensing. His current research interests include multimodal remote sensing data fusion; topside ionosphere exploration with spaceborne HF/VHF-SAR system; high-resolution spaceborne SAR image formation; and SAR image quality enhancement. Dr. Chen was awarded New Century Excellent Talents in University by the Ministry of Education, China, in 2006, and the Excellent Young Teachers in University, by the Fok Ying Tong Education Foundation and Ministry of Education, China, in 2008.

Xiaoxiang Zhu (Fellow, IEEE) received the M.Sc., Dr.-Ing., and Habilitation degrees in signal processing from the Technical University of Munich (TUM), Munich, Germany, in 2008, 2011,and 2013, respectively.

She is Professor of signal processing in Earth observation with TUM and the Head of the Department EO Data Science, Remote Sensing Technology Institute, German Aerospace Center (DLR), Weßling, Germany. Since 2019, she has been a Co-Coordinator of Munich Data Science Research School. Since 2019, she also heads Helmholtz Artificial Intelligence Cooperation Unit (HAICU)—Research Field Aeronautics, Space and Transport. Since May 2020, she has been the Director of the International Future AI Lab AI4EO—Artificial Intelligence for Earth Observation: Reasoning, Uncertainties,Ethics and Beyond, Munich, Germany. She was a Guest Scientist or Visiting Professor with Italian National Research Council (CNR-IREA), Naples, Italy, Fudan University, Shanghai, China, the University of Tokyo,Tokyo, Japan, and University of California at Los Angeles, Los Angeles, CA,USA, in 2009, 2014, 2015, and 2016, respectively. Her main research interests are remote sensing and Earth observation, signal processing, machine learning and data science, with a special application focus on global urban mapping. Dr. Zhu is a member of young academy (Junge Akademie/Junges Kolleg) at the Berlin-Brandenburg Academy of Sciences and Humanities and the German National Academy of Sciences Leopoldina and the Bavarian Academy of Sciences and Humanities. She is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.