

OpenSARUrban: A Sentinel-1 SAR Image Dataset for Urban Interpretation

Juanping Zhao, Zenghui Zhang, *Member, IEEE*, Wei Yao,
Mihai Datcu, *Fellow, IEEE*, Huilin Xiong, *Member, IEEE*, and Wenxian Yu

Abstract—The Sentinel-1 mission provides a freely accessible opportunity for urban image interpretation based on synthetic aperture radar (SAR) data with specific resolution, which is of paramount importance for Earth observation. In parallel, with the rapid development of advanced technologies, especially deep learning, we urgently need a large-scale SAR dataset supporting urban image interpretation. This paper presents *OpenSARUrban*: a Sentinel-1 dataset dedicated to the content-related interpretation of urban SAR images, including a well-defined hierarchical annotation scheme, data collection, well-established procedures for dataset compilation and organization as well as properties, visualizations, and applications of this dataset. Particularly, our *OpenSARUrban* collection provides 33,358 image patches of urban SAR scenes, covering 21 major cities of China, including 10 different target area categories, 4 kinds of data formats, 2 kinds of polarization modes, and owning 5 essential properties: large-scale coverage, diversity, specificity, reliability, and sustainability. These properties guarantee the achievement of several goals for *OpenSARUrban*. The first one is to support urban target characterization. The second one is to help develop well-applicable and advanced algorithms for Sentinel-1 urban target classification. The third one is to explore content-based image retrieval for this kind of data. In addition, dataset visualization is implemented from the perspective of manifolds to give an intuitive understanding. Besides a detailed description and visualization of the dataset, we present results of some benchmarking algorithms, demonstrating that this dataset is practical and challenging. Notably, developing algorithms to enhance the classification performance on the whole dataset and considering the data imbalance are especially demanding.

Index Terms—Sentinel-1 dataset, synthetic aperture radar, OpenSARUrban, urban interpretation.

I. INTRODUCTION

SENTINEL-1 is an imaging radar mission providing continuous all-weather, day-and-night imagery of relatively low-resolution C-band data [1]–[4]. Potentially imaging all global landmasses, Sentinel-1 allows for comprehensive urban

target interpretation [5], [6]. In particular, the interferometric wide (IW) swath mode is the primary operational mode over land. The data are publicly accessible and provide sufficient resources for land cover applications, such as urban deformation mapping [5], [7], or forest and agriculture monitoring [8]–[10].

Considering the explosion of Sentinel-1 satellite data, the lack of urban data interpretation tools [11]–[17] as well as the rapid development of new deep learning techniques [18]–[24], the user community urgently needs a large-scale Sentinel-1 image dataset to develop more sophisticated and robust algorithms for the interpretation of urban synthetic aperture radar (SAR) images. The challenges are how the ever-increasing data can be indexed, organized into a dataset, and utilized for specific applications. These issues cause a crucial problem yet to be solved.

There are already large-scale datasets, having been compiled in the optical remote sensing field to satisfy different requirements. The existing literatures include the UC Merced land use dataset (UC-Merced for short) [25], the local climate zone (LCZ) dataset [26], the aerial image dataset (AID) [27], AID++ [28], the dataset for object detection in aerial images (DOTA) [29], and the EuroSAT dataset [30]. Because of the clear visual appearance of optical images, any dataset compilation is relatively easy to perform. On the contrary, in the SAR community, a dataset compilation faces more severe challenges. On the one hand, the non-intuitive visual image appearance—caused by the active imaging of SAR—poses the biggest obstacle in SAR image annotation. On the other hand, the SAR data themselves are rather expensive to acquire, which also is an important factor impeding any SAR dataset compilation. Despite these difficulties, researchers have developed several datasets in this field. For instance, the Western North America Interferometric SAR (WinSAR) Consortium (<https://winsar.unavco.org/>) acquires SAR imagery aiming to promote the development and the use of InSAR technology. Further, the moving and stationary target recognition (MSTAR) dataset [31], covering different aspect angles, depression angles, and target configurations, is composed of ten types of military vehicle targets. The dataset is extensively adopted to develop automatic target recognition (ATR) algorithms for SAR images [32]–[34]. In addition, the authors of [35] designed a large-scale TerraSAR-X dataset based on very high resolution (HR) imagery, aiming to promote information mining from HR and X-band SAR images. In contrast, the authors of [36] developed a HR and X-band land cover dataset for classification benchmarking of temporal changes. Later, the OpenSARShip [37] image

Manuscript received XXX, XXX; revised XXX, XXX. This work was supported in part by National Natural Science Foundation of China (Grant No. U1830103) and in part by China Scholarship Council (CSC). (*Corresponding author: Zenghui Zhang.*)

J. Zhao is with the Department of Electric Information and Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China. She is currently a joint Ph.D. student in the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Wessling 82234, Germany (e-mail: juanpingzhao@sjtu.edu.cn).

Z. Zhang, H. Xiong, and W. Yu are with the Department of Electric Information and Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: zenghui.zhang@sjtu.edu.cn, hlxiong@sjtu.edu.cn, wxyu@sjtu.edu.cn).

W. Yao and M. Datcu are with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Wessling 82234, Germany (e-mail: wei.yao@dlr.de, mihai.datcu@dlr.de).

collection, containing 11,346 SAR ship chips, was designed to promote Sentinel-1 ship interpretation. More recently, the SEN1-2 dataset [38] is to foster deep learning research in SAR-optical data fusion. However, none of these SAR datasets can focus on the interpretation of urban Sentinel-1 images with their relatively low resolution. This dataset benefits researchers also due to the advantages of large coverage, fewer layovers, and easiness due to their freely accessible priority.

With the goal of filling this gap and to advance interpretation research with urban SAR images, in this study, we present a benchmarking SAR dataset called *OpenSARUrban*, which has been collected from 19 Sentinel-1 images, mainly covering areas of 21 individual metropolises of China. In the very beginning, a coarse-to-fine annotation scheme was proposed, which was initially implemented according to the urban operation functionalities and then hierarchically divided into more detailed categories (see Fig. 2). The dataset, comprising 33,358 SAR image patches (i.e., image chips) with a size of 100×100 pixels each, supports 10 different functional urban types. The design of the *OpenSARUrban* dataset follows the idea of annotation transition from the optical domain to the SAR domain. Owing to five essential properties, namely large-scale, diversity, specificity, reliability, and sustainability, *OpenSARUrban* achieves several goals. The first one is to support urban target characterization analysis. The second one is to foster applicable and advanced classification algorithms for Sentinel-1 urban targets. The third one is to explore content-based image retrieval [39]–[42] of this kind of data. The visualization of this dataset is performed from the perspective of manifolds via a combination of fast compression distances (FCDs) and t-distributed stochastic neighbor embedding (t-SNE), which offers an intuitive way to understand the structure within the given dataset. In the case of image classification of this dataset, some representative benchmarking algorithms are provided.

The three main contributions of this paper can be summarized as follows. Firstly, a hierarchical coarse-to-fine annotation scheme for urban target interpretation is proposed, which takes the urban requirements into account. Secondly, via organizing and exploiting a rapidly growing set of Sentinel-1 SAR images, we compiled the *OpenSARUrban* dataset, which is particularly applicable to urban target interpretation. Thirdly, five essential properties were achieved and some benchmarking experimental analysis was made, which contributes to the practicality and the quality of this dataset.

The remainder of this paper is organized as follows. Section II presents detailed procedures for compiling the *OpenSARUrban* dataset. The layout and properties of *OpenSARUrban* are illustrated in Section III and Section IV, respectively. Section V visualizes the manifolds within this dataset. Section VI provides some preliminary applications on urban target classification of this dataset as benchmarking algorithms. Finally, conclusions are drawn and future work is illustrated in Section VII.

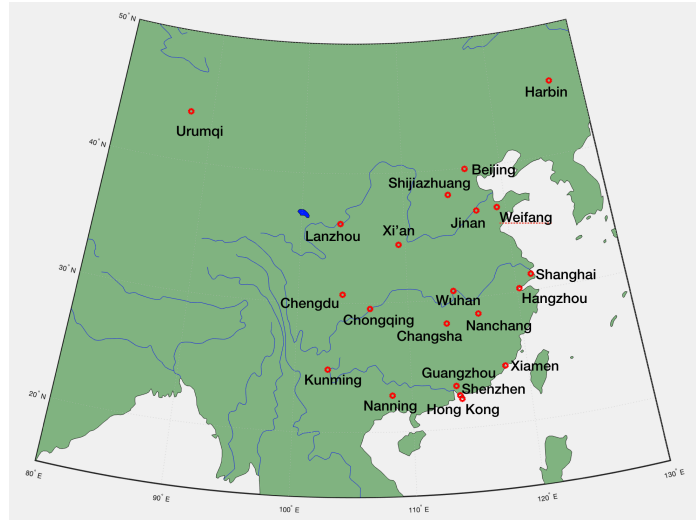


Fig. 1. Data source distributions of the *OpenSARUrban* dataset.

II. CONCEPTION AND COMPILATION OF THE OPENSARURBAN DATASET

In this section, we present procedures for conceiving the *OpenSARUrban* dataset; we also explain how these procedures guarantee the properties of large-scale image data, their diversity, specificity, reliability, and sustainability. The dataset compilation can be explained from three aspects: data collection and preprocessing, a well-defined annotation scheme, and the step-by-step compilation procedures.

A. Data Collection and Preprocessing

Before compiling the dataset, it is necessary to collect some typical original images from the Sentinel-1 data access hub, and the corresponding preprocessing has to be done. In this study, we focus on urban targets from major cities distributed across China.

During a data collection phase, a large amount of initial Sentinel-1 SAR images were selected and downloaded from a SAR image archive, containing typical regions of interest (RoIs). In our case and in this work, the selected Sentinel-1 images cover areas of 21 major Chinese cities from 17 administrative provinces. Most of them are located around provincial cities. Table I shows details of the dataset source particularities. The geographical distribution of this dataset is shown in Fig. 1. Notably, the red circles in this figure denote the different cities. The green area, the blue colors, and the gray-colored region are land areas, rivers, and ocean areas, respectively.

In this study, we focus on Level-1 ground range detected (GRD) data with IW swath products, typically regarded as the default acquisition mode over land. These original images were downloaded from the official *Sentinels Scientific Data Hub* (<https://scihub.copernicus.eu/dhus/#/home>), which is freely accessible to researchers.

The original images and their radiometrically calibrated versions are included in this dataset. It is notable that the pixel values of radiometrically calibrated data can be directly related

TABLE I
MAJOR CHINESE CITIES INVESTIGATED IN THIS STUDY

Imagery scene (covered provincial areas)	City name	Image acquisition date (Year/Month/Day)
Shanghai	Shanghai	2017/02/03
Beijing	Beijing	2017/02/01
Zhejiang	Hangzhou	2017/03/28
Guangdong & Hong Kong	Shenzhen, Guangzhou, Hong Kong	2017/02/28
Guangxi	Nanning	2017/03/10
Yunnan	Kunming	2017/02/05
Sichuan	Chengdu	2017/04/13
Chongqing	Chongqing	2017/05/02
Hubei	Wuhan	2017/01/06
Hunan	Changsha	2017/05/11
Jiangxi	Nanchang	2017/03/02
Shandong	Jinan, Weifang	2017/05/15
Hebei	Shijiazhuang	2016/11/14
Shaanxi	Xi'an	2017/03/05
Gansu	Lanzhou	2017/05/15
Xinjiang Uygur Autonomous Region	Urumqi	2017/02/06
Heilongjiang	Harbin	2016/09/14

to the radar backscattering of the Earth's surface. In contrast to the qualitative usage of the original SAR data, their calibrated version is essential for quantitative applications. We used the SNAP 3.0 software to perform radiometric SAR image calibration. For both uncalibrated data and their corresponding calibrated version, we used GeoTIFF format images.

Annotating SAR images directly by expert inspection is a very laborious and time-consuming task. Apart from this, the relatively low image resolution causes great difficulty in target type determination. In order to overcome these challenges, the optical images from Google Earth Engine provide an optimal solution, which fills to a great extent the gap between the human visual system (HVS) [43] and the radar's active imaging mode. Firstly, the optical images can be easily recognized by human observation, which directly leads to qualified annotations. Secondly, geographical annotations can be generated by 91 Weitu [44], which can be imported into SNAP 3.0 and then matched geographically with the corresponding SAR images. Thirdly, several petabytes of optical remote sensing images are also provided by 91 Weitu, which is achieved by plugging in the Google Earth Engine.

B. Discernible Categories

The annotation scheme proposed in this paper is actually a coarse-to-fine hierarchical scheme, confirmed by two factors. The first one is related to the different visual patterns in Sentinel-1 SAR images, which are confirmed by several SAR experts. Even though there are subtle differences, most categories from different images are distinguishable. The other factor is the different use of urban areas, i.e., business areas, residential areas, industrial areas, and others. Fig. 2 shows our annotation scheme. The first level is differentiated by the overall functionality. In what follows, the second level gives more detailed semantic sub-categories, which have to be annotated individually. The annotation scheme is defined based on the data collection from major Chinese cities. However, this will not limit the generation of the dataset. Firstly, our

dataset is mainly composed of urban buildings, which vary a lot between China and other countries, like France we have already investigated, but where we found that it doesn't work. Considering this, we cannot give a uniform annotation scheme from a worldwide perspective. Secondly, the buildings in Chinese cities vary a little, but not too much. So the annotation scheme is relatively easy to define. Thirdly, the dataset is large enough to support our research.

For instance, skyscrapers are a representative building type in business areas, which often consist of extremely bright pixels in SAR images. In contrast, residential areas can be described by four sub-categories, i.e., general residential areas, high-rise buildings, dense and low-rise residential areas, and villas. General residential areas in today's China contain as the most common buildings those with not more than six floors. On the contrary, high-rise areas consist of residences with tens of floors and the distance between the buildings is usually very large. Dense and low-rise areas are commonly very crowded and their buildings are very low. Generally, villas are located in suburbs and they are often surrounded by trees, vegetation, and lakes, while a storage area is a typical category in industrial areas, usually located at the city periphery. Playing a paramount role in any municipal transportation system are its hubs, including airports, railways, and highways that are essential for the city operation. It is worthwhile to point out that highways are also an important indicator of overpasses in urban areas. In addition, another indispensable category needed in urban areas is vegetation. Fig. 3 illustrates typical examples of optical samples and their corresponding SAR samples, covering all our 10 categories. The colored masks on each example are actually the main surface cover of the current category. It can be seen that the labeled areas have different shapes between the optical images and their corresponding SAR images. The reasons can be attributed to three points. Firstly, the imaging mechanisms of SAR and optical instruments are different. Specifically, SAR is an active imaging technique which transmits and receives the reflected signals, whereas optical sensors work in a passive

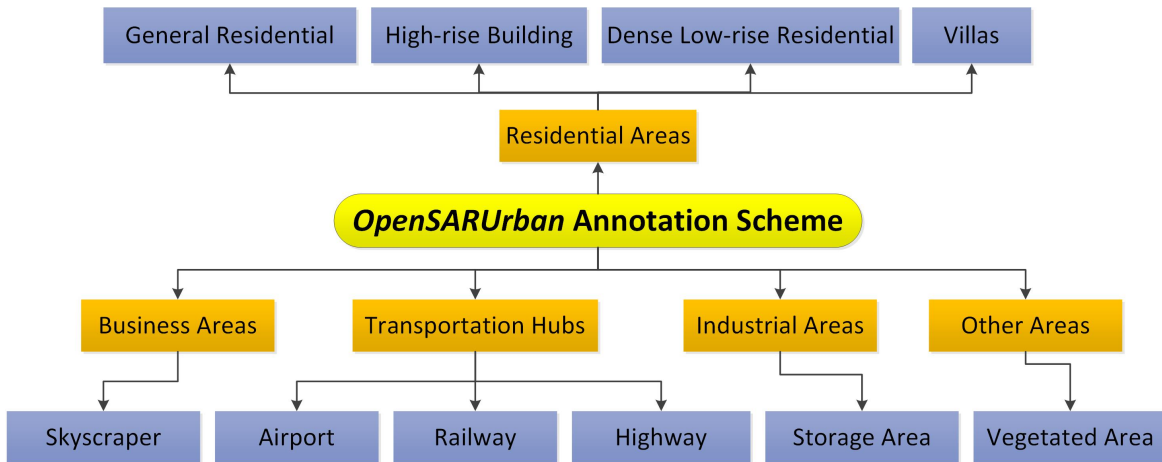


Fig. 2. Our two-level hierarchical annotation scheme.

mode, which receives the reflected signals of the natural illumination. Therefore, the geometrical distortions in SAR images are very common, and are not in accordance with our human visual experience. In contrast, optical images are more easily recognized by human observers. Secondly, a SAR sensor may observe the Earth’s surface from different directions, thus accentuating the phenomenon. Thirdly, the depression angle can also call for a rotation or a flipping of the images when comparing them with optical images. Even though there are significant differences between optical images and their corresponding SAR images, the geographical information plays an important role during image co-registration. These differences can make the interpretation of SAR images rather difficult.

C. Dataset Compilation

Fig. 4 displays the overall workflow of the *OpenSARUrban* dataset compilation. The annotation of this dataset was implemented by a transition from optical space to SAR space. In particular, targets were firstly annotated in our optical images with the assistance of the 91 Weitu software package [44]. The optical annotations were then saved in a “.shp” file together with their geographical information, which serves as a bridge linking the optical images with their corresponding (i.e., overlapping) SAR image counterparts. The time gap between optical images and SAR images was eliminated by aligning the image acquisition times recorded in the meta-data files. The target coordinates within the SAR images can be obtained by using the Sentinel-1 application platform (SNAP 3.0) software [45]. Finally, with the availability of the generated SAR annotations, the original SAR images were tiled into image patches separately for each land cover category. It is worth noting that the validity of the *OpenSARUrban* data is accurately monitored in every step and is provided in several data formats to satisfy different user requirements. The core procedures can be explained by nine steps as follows:

Step 1: Information Querying

In the meta-data file of Sentinel-1 packages, the geographical location, i.e., the longitude and latitude, of the image’s four corners can be easily queried. In addition, the image acquisi-

tion date/time can also be obtained from this file, formatted as “Date-Month-Year” plus the coordinated universal time (UTC) of the acquisition.

Step 2: Geo-Referencing of the Corresponding Optical Remote Sensing Images

With the information about location and image acquisition time being available, one can search the RoIs in optical remote sensing images, which are mostly in line with the HVS, making any semantic annotation much easier. Specifically, the image location is defined by the boundaries of the current imaging area and the exact image acquisition time can be traced. Here, the 91 Weitu software provides an optimal way to make this step.

Step 3: Creating a New Directory

Before annotating a specific category, we routinely create a new directory, naming it with the category name, such as “high-rise building”. In the following steps, the human-readable annotations belonging to this category can be automatically saved in this directory on condition that the directory has been successfully activated.

Step 4: Annotating Optical Remote Sensing Images

With a relatively low resolution of about 20 m for Sentinel-1 images, directly annotating SAR images is very challenging for image analysts. In order to overcome this problem, the optical images, being directly understandable by human observation, provide an optimal way to accurately annotate a given image. The optical images from Google Earth Engine have a resolution of 0.13 m both along azimuth and range directions. The annotations are marked with a series of polygons using the “Polygon Annotation” tool provided by the 91 Weitu software. The geographical locations of the annotated polygon vertices are also stored in the annotation file. Thus, with the geographical information as a bridge, annotations in optical images can be easily matched with co-aligned SAR images.

Step 5: Saving the Optical Annotations

This step is to save the optical annotations in a “.shp” file, which contains the geographical information of each annotation. The “.shp” annotation file is to be saved in the directory created in Step 3.

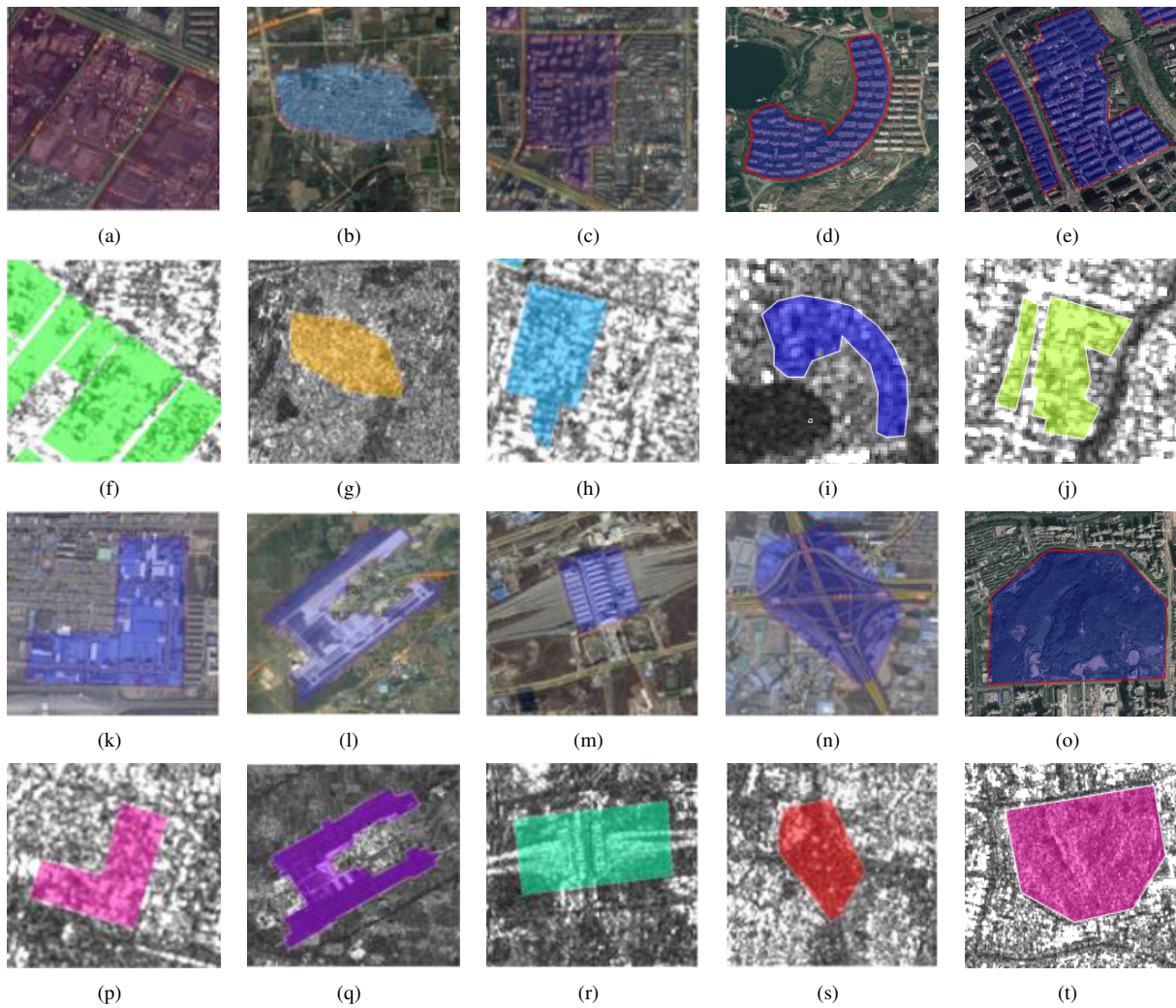


Fig. 3. Optical examples and their corresponding SAR examples for each category. The first row and the second row show optical and SAR examples of skyscraper, dense and low-rise residential buildings, high-rise buildings, villas, and general residential areas, respectively; The third row and the fourth row show optical and SAR examples of storage areas, airports, railways, highways, and vegetation, respectively. The colored masks are the main land cover locations of the given category.

Step 6: Checking the Optical Annotations

When all the categories in the optical images have been annotated and the corresponding “.shp” files have been created, these annotations should be checked in the optical domain with the assistance of the 91 Weitu software. During this process, if the targets are correctly annotated, these files are saved into a directory named “Optical Annotation”. Otherwise, the optical annotations should be corrected before saving.

Step 7: Image Co-Registration

With the geographical information being available, the optical annotations can be easily matched with the corresponding SAR images with the help of SNAP 3.0 desktop. In particular, the Sentinel-1 images should be opened first by the SNAP 3.0 software. In this study, the selected image is usually a VH polarized image, because of its better visual quality when comparing it with VV polarization. The better visual appearance of VH polarization can be demonstrated in this step. Then, the optical annotations, formatted with “.shp” files, are imported into SNAP. The annotations will be checked and matched with the corresponding SAR image by using

this software, where the Sentinel-1 image and the optical annotations are co-aligned automatically.

Step 8: Exporting SAR Annotations

With the help of the SNAP 3.0 toolbox, the annotation information of SAR images is stored in a “.txt” file in the “SAR Annotation” directory. Actually, the SAR annotation contains the coordinates of the categorical locations in the Sentinel-1 images. More specifically, the coordinates are the locations of the annotated polygon vertices.

Step 9: Generating SAR Image Patches with Different Formats

In this step, SAR image patches of each category are generated with MATLAB. Using the polygon annotations, we tile them into image patches with a size of 100×100 pixels. Here, the overlap of the annotated polygons and the tiled image patches is set to be larger than a given pixel margin. The overlap computation mechanism is illustrated in Fig. 5. In this figure, the purple square and the yellow polygon represent a tiled image patch and an annotated SAR polygon, respectively.

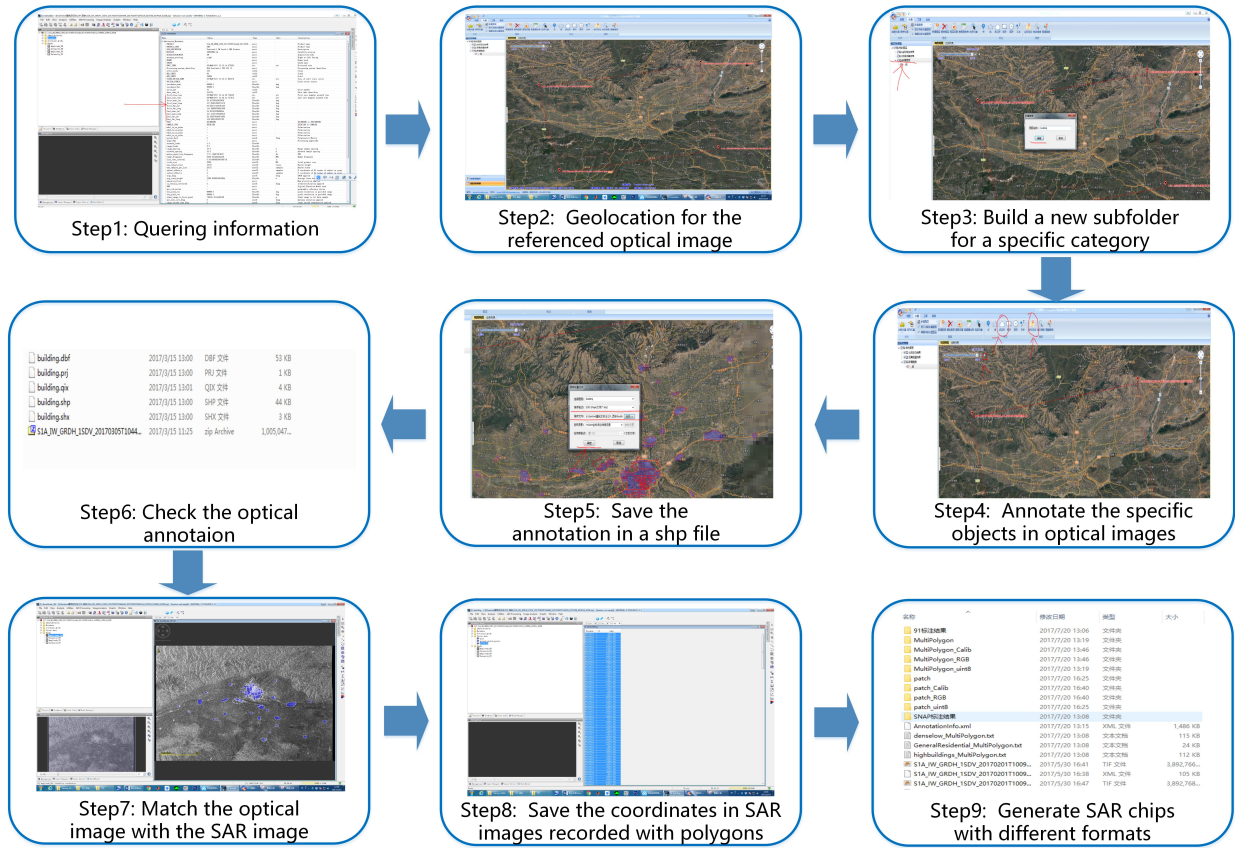


Fig. 4. The workflow of compiling the *OpenSARUrban* dataset.

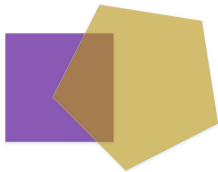


Fig. 5. Our overlap computation mechanism. The purple square and the yellow polygon represent a tiled image patch and an annotated SAR polygon, respectively. If their overlap is larger than a pre-set margin, the corresponding image patch is saved into the dataset.

III. THE LAYOUT OF THE OPENSARURBAN DATASET

The *OpenSARUrban* dataset is organized in different folders for different image categories and formats. For a given Sentinel-1 SAR image, the content of the generated dataset is illustrated in Fig. 6. When it comes to the formats of image patches, each of them consists of four different formats, i.e., the original data, the visualized data in gray-scale representation, the visualized data in pseudo-color, and the radiometrically calibrated data. Image patches with different formats are provided in different subfolders. Different formats shall achieve the goal to satisfy different user requirements. At present, ten different target categories are stored separately. Each image patch is named by the combination of its category name, pixel coordinates, polygon index, and the polarization mode, thus supporting the patch retrieval.

The original data is stored in 32-bit format, which is in line with the original data format of Sentinel-1. Considering the GRD format of a Sentinel-1 image, each image patch is stored in a two-channel matrix, where each one contains the amplitude values of the pixels with VH and VV polarizations, respectively. Based on the original image patches, image enhancement is applied to visualize the data in gray-scale as unsigned 8-bit integers (UINT8) for VH and VV polarizations. The radiometrically calibrated data, obtained by using the SNAP 3.0 software, are stored in a matrix, which contains the normalized radar cross section (NRCS) σ_0 [46] data for VH and VV polarizations. Fig. 7 shows an example for each category of VH and VV polarizations with UINT8 format. The visualization of pseudo-color image is implemented by using the MATLAB function ‘imagesc’. The values of the image elements are indexed into the current ColorMap (jet), determining the pixel color. A color bar is provided to illustrate the radiometric boundary of each pseudo-color image patch.

The detailed information about each image patch, its SAR signatures, and the generated messages provided by their meta-data are listed in an XML-file named “Annotation.xml”. The annotation information, including the annotation times, the locations, etc., is also provided in this file. Via the pixel coordinates contained in the name of each image patch, the corresponding information can be easily retrieved in the XML-file.

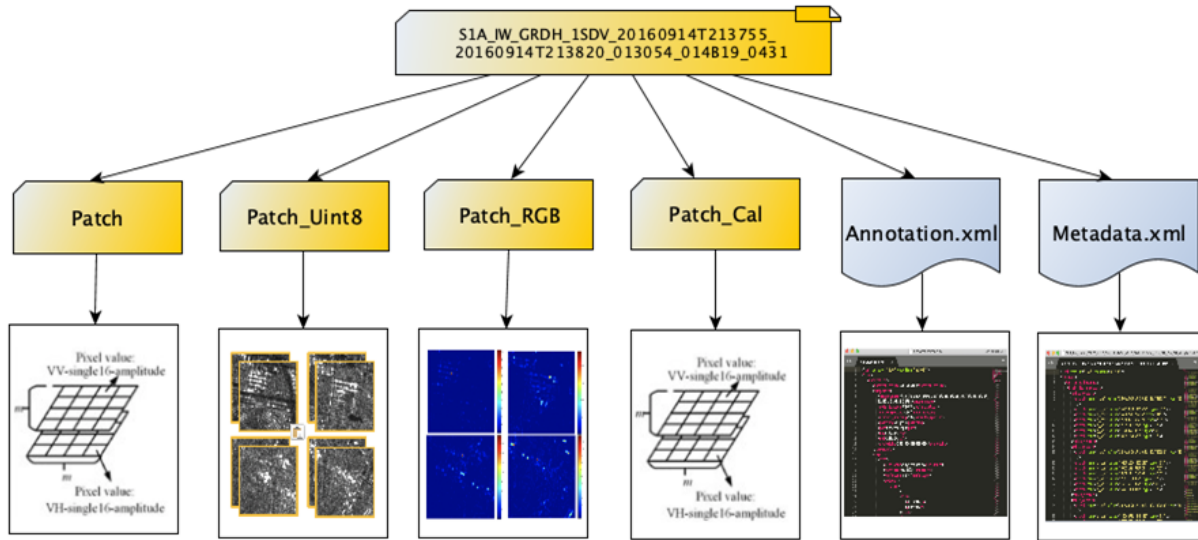


Fig. 6. Data organization of the *OpenSARUrban* dataset.

IV. THE PROPERTIES OF OPENSARURBAN

A. Large-Scale

The original data cover 21 Chinese major cities, related to 17 political provinces. Consequently, the *OpenSARUrban* dataset provides as much as 33,358 image patches, where each image patch includes 4 different kinds of image formats and 2 different kinds of polarization modes.

B. Diversity

- Data Format Diversity

For each image patch, there are four different formats available, i.e., the original 32-bit image, the enhanced gray-scale image, the radiometrically calibrated image, and the pseudo-colored radiometrically calibrated image. These four different formats are generated to satisfy different requirements.

- Geographical Diversity

The original extent of *OpenSARUrban* comprises data from 21 Chinese major cities in 17 administrative provinces. The image patches from each city are separately stored in the corresponding folders. Fig. 8 explains the image patch distributions across different cities. In this figure, bars with different colors represent different categories and image patches are grouped according to the cities they come from. In this figure, Guangzhou actually represents image patches jointly collected from the city of Guangzhou, Shenzhen, and Hong Kong, because of the wide-swath capability of Sentinel-1.

- Categorical Diversity

Related categories in this study and the validity of the annotation scheme are illustrated in Section II-B. All these typical categories comprise the *OpenSARUrban* dataset's categorical diversity. The dataset distributions among different categories are illustrated in Fig. 9. The categories of transportation hubs are very limited in this dataset, calling for more research on data imbalance.

- Polarization Diversity

For each image patch in this dataset, VH- and VV-polarized data are included. Different polarizations, conveying different scattering signatures and visual effects, have different potentials in describing different kinds of urban types.

C. Specificity

This study provides a large-scale C-band urban categorical dataset with a resolution of 20 m. This dataset essentially aims to provide a theorem study on SAR signatures and the characteristics of different urban categories, as well as paving the way for applications of this kind of data. Specifically, the *OpenSARUrban* dataset is designed to: 1) study characteristics and potentials for different urban areas by using Sentinel-1 SAR images; 2) develop sophisticated urban target interpretation algorithms for this kind of data; 3) support content-based image retrieval.

To the best of our knowledge, representative C-band urban target data for SAR images with a resolution of 20 m are not yet available for users to study their characteristics. Therefore, we created the *OpenSARUrban* dataset, expecting to fill this gap.

D. Reliability

Initially, the *OpenSARUrban* dataset was annotated by analyzing optical remote sensing images, which can provide annotators with visually clearly understandable information; furthermore, the optical images can give annotators some additional understandings about the Earth's surface. The image co-registration between optical images and their corresponding SAR images is supported by geographical information, which is very helpful when verifying the annotation coordinates, and the image acquisition dates, which eliminates the time gap between optical images and SAR images effectively. Moreover, some confused classifications can be corrected by the additional information. Thus, the annotation quality of this dataset can be improved.

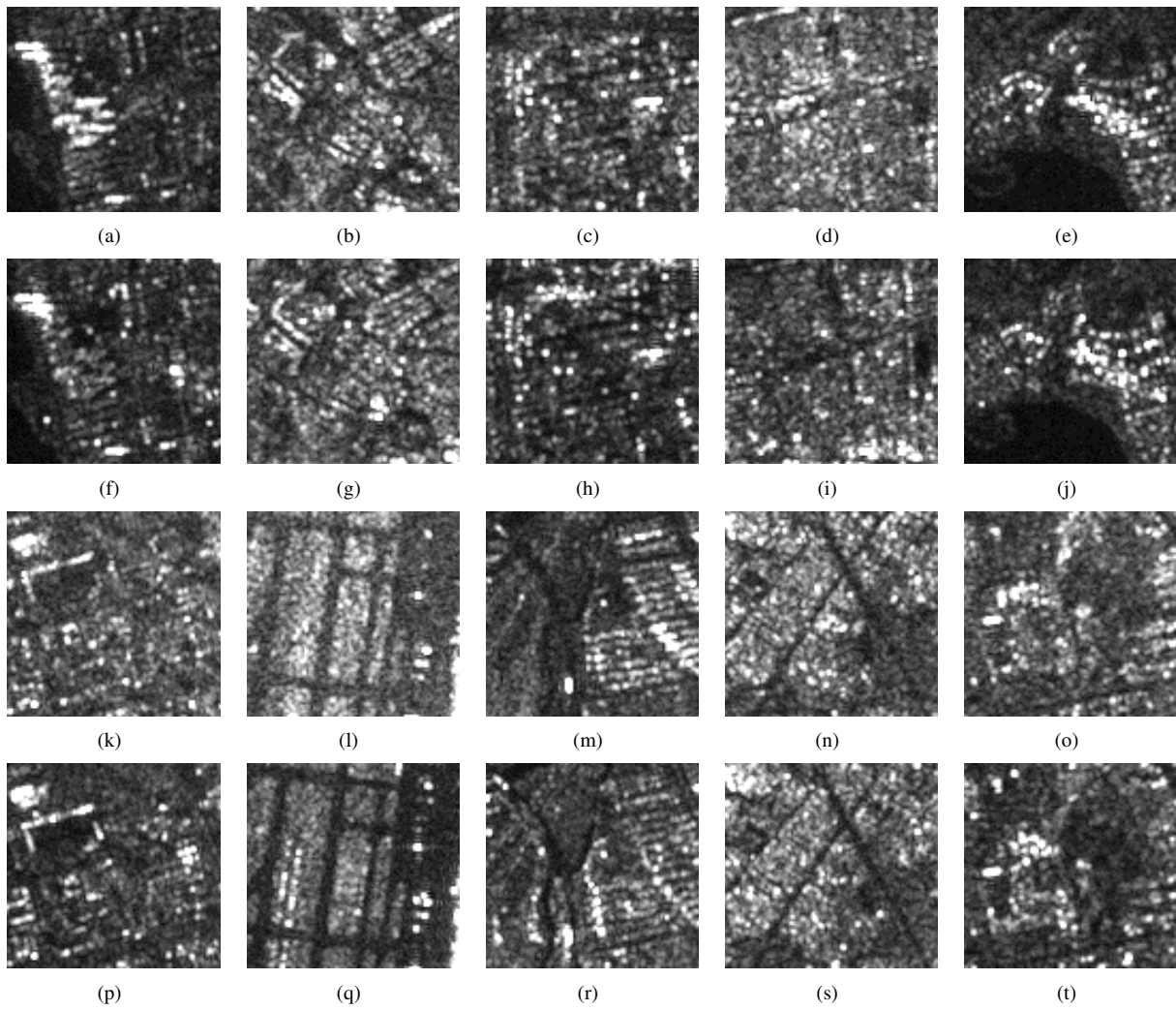


Fig. 7. Image patches for each category with UINT8 format. The first row shows VH-polarized examples of skyscrapers, general residential areas, high-rise building blocks, dense and low-rise residential areas, and villas; the second row shows the corresponding VV-polarized examples; the third row displays VH-polarized data of airports, railways, highways, industrial storage areas, and vegetated areas; the fourth row exhibits the corresponding VV-polarized patches.

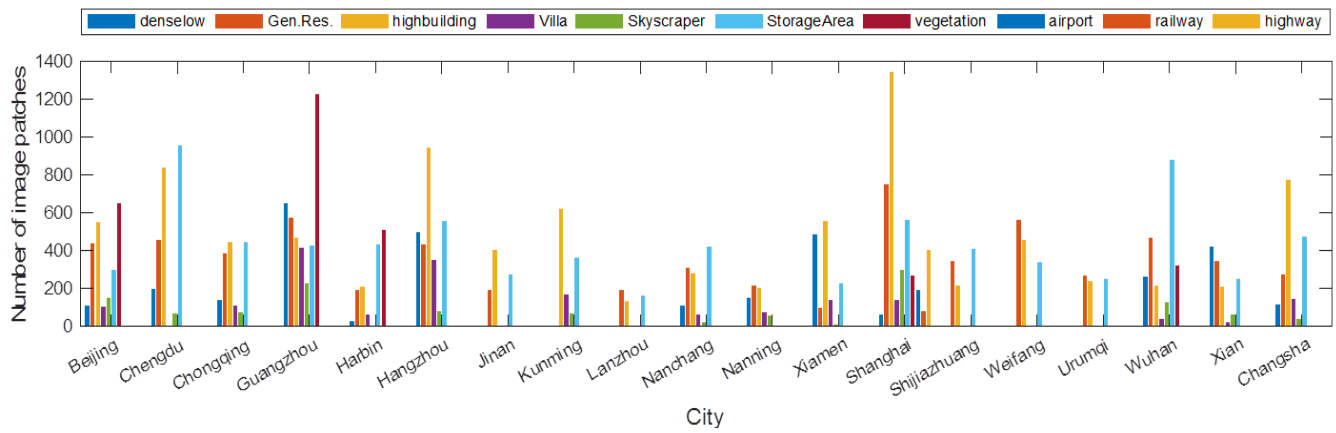


Fig. 8. Dataset distributions among different cities. Differently colored bars represent the number of image patches from different categories. Image patches are grouped according to the city distribution.

E. Sustainability

The optical annotations recorded as “.shp” files can be used for further co-registration with newly available SAR data. The

geographical information contained in the optical annotation files can be regarded as an effective bridge linking new SAR

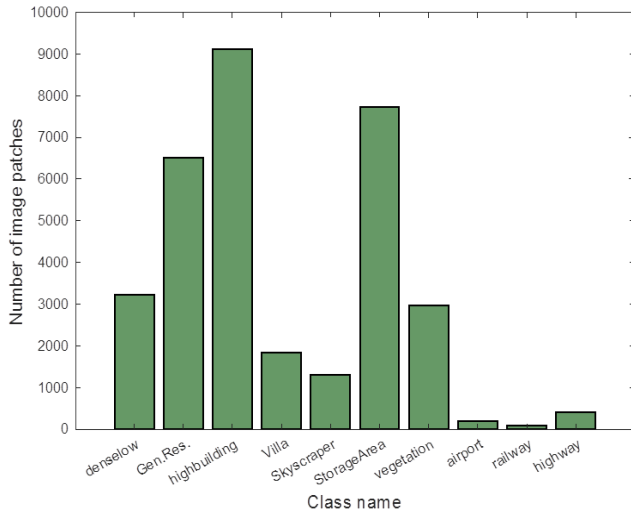


Fig. 9. Dataset distributions among different categories.

images located in the same area and acquired at the same date. Thus, the manual labeling effort can be greatly reduced, more Sentinel-1 SAR datasets can be easily generated, and more urban categories can be added to the dataset. In these cases, the process makes the *OpenSARUrban* dataset to be enriched in a relatively simple way.

V. VISUALIZATION OF THE OPENSARURBAN DATASET

To illustrate the data manifolds within *OpenSARUrban*, the algorithms of FCD [47], showing its advantages when applying a remote sensing dataset, and t-SNE [48] are combined to visualize this dataset. The idea behind this is to convert similarities between data points to joint probabilities and attempting to minimize the Kullback-Leibler (KL) divergence [49] between the joint probabilities of the low-dimensional manifolds and the high-dimensional data. This method is a parameter-free and thus unbiased data analysis technique, performing well on preserving the complete local structure and some global structures of the data points. The visualization and interpretation are based on a Vega style interactive tool. Users are able to zoom in, zoom out, and pick out datasets from *OpenSARUrban*. The visualization procedures are depicted in Fig. 10, including raw data extraction, dictionary extraction, pair-wise distance computation, dimension reduction, and visualization. Further details can be found in [50].

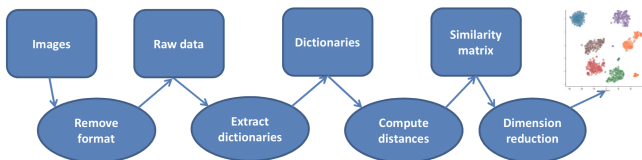


Fig. 10. Flowchart of the visualization procedure.

Considering the computational load of this method, in an experiment, we picked out 300 samples randomly from each

category, including the 2 polarization modes. The visualization results are analyzed in the manifold space, which provides an intuitive way to understand the dataset. Fig. 11(a) and Fig. 11(b) exhibit the category visualization results for VH and VV polarization, respectively. Different colors in these figures represent different urban categories. From the visualization results, one can observe that:

- The VH-polarized images are generally more clearly distinguishable than the VV-polarized data.
- The transportation hubs, including airports, railways, and highways, have their distinct manifold spaces, both for VH and for VV polarization. This is due to their specific image patterns and the large functional distances between them.
- The business areas, represented by skyscrapers in this dataset, dominate their own manifold space for both VH and VV polarization modes. The performance seems better for VH polarization than for VV polarization.
- The functional areas of residential regions, including general residential areas, high-rise building areas, dense and low-rise areas, and villas, are mostly assembled in one cluster, following the manifolds of the VH-polarized dataset. As for the VV-polarized results, they are sometimes mixed up with other urban categories.
- In this dataset, comprising the manifold spaces of the VH- and VV-polarized datasets, the vegetation areas can be well clustered. However, this cluster is close to industrial areas, represented by the storage category, from the perspective of manifold visualization.
- Both figures pose great challenges in distinguishing the 10 categories.

VI. EVALUATION OF OUR URBAN CATEGORIZATION AND DISCUSSIONS OF OPENSARURBAN

It is generally acknowledged that urban categorization for SAR images is very challenging. The difficulty is further increased when encountering such a relatively low resolution. As explained before, these challenges always exist and are rather severe for *OpenSARUrban* interpretation. We assume that the *OpenSARUrban* dataset classification is demonstrably reliable as shown by some prevailing deep learning methods and some traditional SAR image classification techniques, e.g., a combination of representative traditional feature descriptors of SAR images and a linear support vector machine (SVM) classifier [51].

A. Benchmarking Algorithms for Urban Categorization

In order to demonstrate the distinguishability of target area categories and to provide some representative benchmarking algorithms for this dataset, we carried out urban target categorization on our *OpenSARUrban* dataset. In order to achieve this task, we comprehensively analyzed the prevailing deep learning algorithms developed in recent years and several representative hand-crafted feature descriptors for SAR images.

The reason for using deep learning algorithms in this study were the astonishing achievements they have achieved. In particular, densely connected convolutional networks

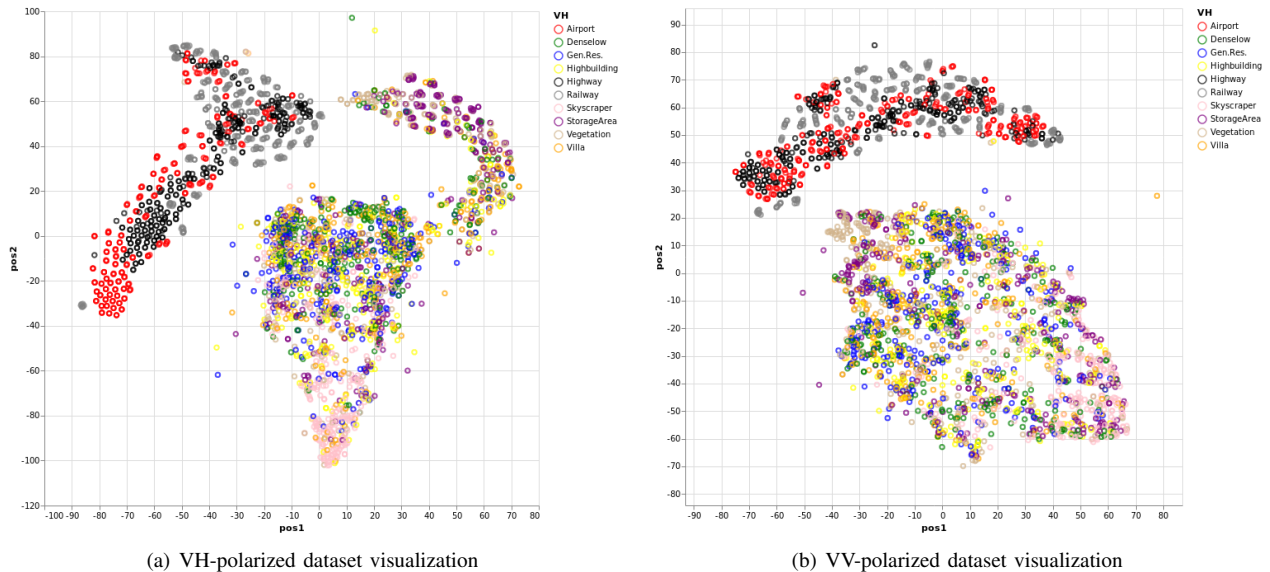


Fig. 11. Urban target visualization of the *OpenSARUrban* dataset.

(DenseNet) [52], the deep residual network with 50 residual blocks (ResNet50) [53], SqueezeNet [54], very deep convolutional networks with 19 layers (VGG19) [55], and AlexNet [56] were evaluated for the classification of this dataset.

For the evaluation of traditional methods, six prevailing SAR image descriptors, including local binary patterns (LBPs) [57], LogGabor features [58], Gabor features [59], Weber local descriptors (WLDs) [60], histograms of oriented gradients (HoGs) [61], and principle component analysis (PCA) [62], were selected to evaluate the usefulness of *OpenSARUrban*. The number of scales and orientations for Gabor features and LogGabor features were set to 6 and 4, respectively. Then, a PCA feature descriptor reduces each image patch to a 30-dimensional vector. For simplicity and giving a fair comparison, a linear SVM was chosen as a classifier in each experiment.

B. Implementation Details and Evaluation Metrics

During the process of implementing urban target classification with this dataset, it was split into a training part and a testing part. Among them, 60% of each category was chosen randomly for training, and the remaining data were used for testing and evaluation. In order to overcome the training data imbalance among the different categories, different Gaussian noise was added to augment the training data. In addition, the *OpenSARUrban* data were evaluated comparatively with different polarization modes [63].

The network training for deep learning techniques was started with an initial learning rate of 0.001 and the learning rate changes were set in accordance with the “poly” descending policy described in [64]. The parameters were iteratively updated until convergence was reached.

Without loss of generality, the overall accuracy (OA) [65], [66] and a confusion matrix [27], [28] were applied to evaluate the performance of the available benchmarking algorithms.

C. Overall Performance of *OpenSARUrban* for Urban Target Categorization

The OA of the *OpenSARUrban* dataset by each benchmarking algorithm is shown in Fig. 12. In this figure, the purple bar and the yellow bar indicate the performances of the VH and VV polarization modes of this dataset, respectively. One can observe from this figure that:

- Generally speaking, deep learning methods surpass traditional methods by a large margin, except for LBP.
- The best performance for both VH polarization and VV polarization, in terms of OA, can be given by VGG19, being 89.49% and 89.53%, respectively.
- Among the traditional methods, LBP, with 71.06% and 70.82% of OA for the VH and VV polarization modes, is far more advantageous than the other methods and even surpasses some deep learning techniques.

In real situations, the categorization depends on the quality of the samples, the feature selections, etc.. However, in general, the VH-polarized man-made structures perform better than VV-polarized ones. We would like to explain this phenomenon from three different perspectives. Firstly, this can be understood from the perspective of radar backscattering theory. The categories related to this study are mostly man-made structures, except for the category of vegetation. It has been demonstrated that VH-polarized data contributes the most to the classification of man-made structures [67]. The second viewpoint is to explain the phenomenon from the perspective of visual appearance. A comparison of visual appearance between VH-polarized images and VV-polarized images is shown in Fig. 7. By human observations, it is visually clear that the VH-polarized images show more intuitive details than the VV-polarized ones. Thirdly, we attempt to give some explanations from the perspective of manifold visualization. The even more separable visual intuitive details of VH-polarized images can also be demonstrated by the manifold visualization results, shown in Fig. 11(a) and Fig. 11(b).

The easier separation between different categories of VH polarization shows us that the VH-polarized images are more distinctive than the corresponding VV-polarized ones.

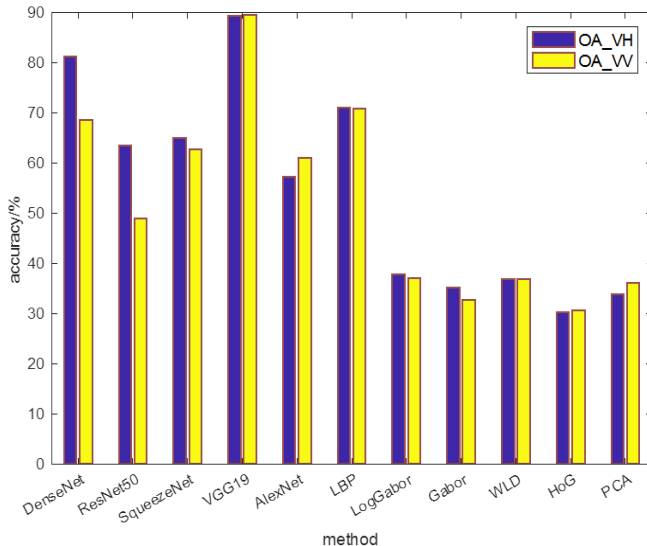


Fig. 12. Overall classification accuracy of the whole dataset.

In order to give a comprehensive explanation of this dataset, we provide the confusion matrices of the best investigated deep learning method (i.e., VGG19) and the best investigated traditional method (i.e., LBP) for the overall classification of this dataset. The confusion matrices produced by VGG19, including both VH and VV polarizations, are depicted in Fig. 13(a) and Fig. 13(b), respectively. Observers can easily conclude from these two confusion matrices that the urban targets in this dataset can be well distinguished by using VGG19, even though their visualization contains some confusions (for comparison, see Fig. 11(a) and Fig. 11(b)). However, even with the VGG19 algorithm, there are still some categories, e.g., general residential areas (*Gen.Res.* for short), storage areas, and dense and low-rise residential areas (simplified as *denselow*) that cannot achieve a highly satisfactory classification accuracy, indicating potentials for developing more advanced classification algorithms on this kind of data.

Figure 14(a) and Fig. 14(b) represent the confusion matrices obtained by the combination of an LBP feature descriptor and a linear SVM classifier. These figures are based on the following points: 1) Most categories can be well categorized except for three typical building types: general residential areas, industrial storage areas, and high-rise buildings, which, comparatively, are better distinguishable by using the VGG19 method; 2) These three most challenging building types are very prone to be confused with each other; 3) An LBP descriptor performs almost the same for both VH and VV polarizations. The inadequacies of LBPs, for one thing, demonstrate the limitations of hand-crafted feature descriptors. For another thing, the transportation hubs, i.e., airports, railways, and highways, are easy to be distinguished because of the great differences in image patterns with urban buildings.

To further explore the effectiveness of LBP features on the

three dominating challenging building types, i.e., *Gen.Res.*, storage areas, and high-rise buildings, we picked them out from the whole dataset and carried out a classification run purely on these three categories. The classification accuracy for each category is shown with their confusion matrices depicted in Fig. 15(a) and Fig. 15(b). These two figures are for VH polarization and VV polarization, respectively. It is concluded that an LBP feature descriptor has very limited capability for recognizing general residential areas, industrial storage areas, and high-rise buildings, which greatly reduces the overall performance of LBPs in differentiating this dataset.

D. Analysis of Specific Urban Functionalities

1) *Residential Area Evaluation*: The classification accuracy of residential areas, including general residential areas, high-rise building areas, dense and low-rise residential areas, and villas, is shown for each algorithm in Fig. 16. Among which, Fig. 16(a), Fig. 16(b), Fig. 16(c), and Fig. 16(d) denote the accuracy of general residential areas, high-rise building areas, dense and low-rise residential areas, and villas, respectively. The purple bar and the yellow bar in these figures illustrate the results of VH and VV polarization, respectively. Based on these figures, some conclusions can be drawn as follows:

- Traditional methods are seriously limited for recognizing general residential areas and high-rise buildings.
- AlexNet is solely effective for high-rise building areas, while it has very limited capabilities for recognizing other residential areas.
- LBPs show that they are advantageous for dense and low-rise areas and villa areas.
- For the recognition of villa areas, DenseNet, VGG19, and LBP stand at a comparably favorable position, showing almost the same level of classification accuracy for both VH and VV polarizations.

2) *Transportation Hub Evaluation*: Fig. 17 exhibits the classification accuracy of transportation hubs by applying the prevailing five deep learning methods and six traditional methods. Both VH and VV polarizations are compared and evaluated. In particular, the classification accuracies of airports, railways, and highways are shown in Fig. 17(a), Fig. 17(b), and Fig. 17(c), respectively. The results demonstrate that: 1) The deep learning algorithms, LBPs, and PCA features have relatively satisfactory capabilities for interpreting both VH and VV polarized transportation hubs; 2) The PCA feature representation is simple yet powerful for transportation hub description, achieving results as good as deep learning methods.

3) *Business Area Evaluation*: The classification accuracy of skyscrapers, including VH and VV polarizations, is shown in Fig. 18. Notably, skyscrapers are the most representative building types in business areas, which are chosen for annotation in the *OpenSARUrban* dataset. We can observe that: 1) The algorithms of VGG19 and LBP show great superiority over others, and for both of them, the VH polarization performs slightly better than VV; 2) The performance differences between VH and VV for DenseNet and ResNet50 are very striking, precisely being 43% and 41%, respectively.

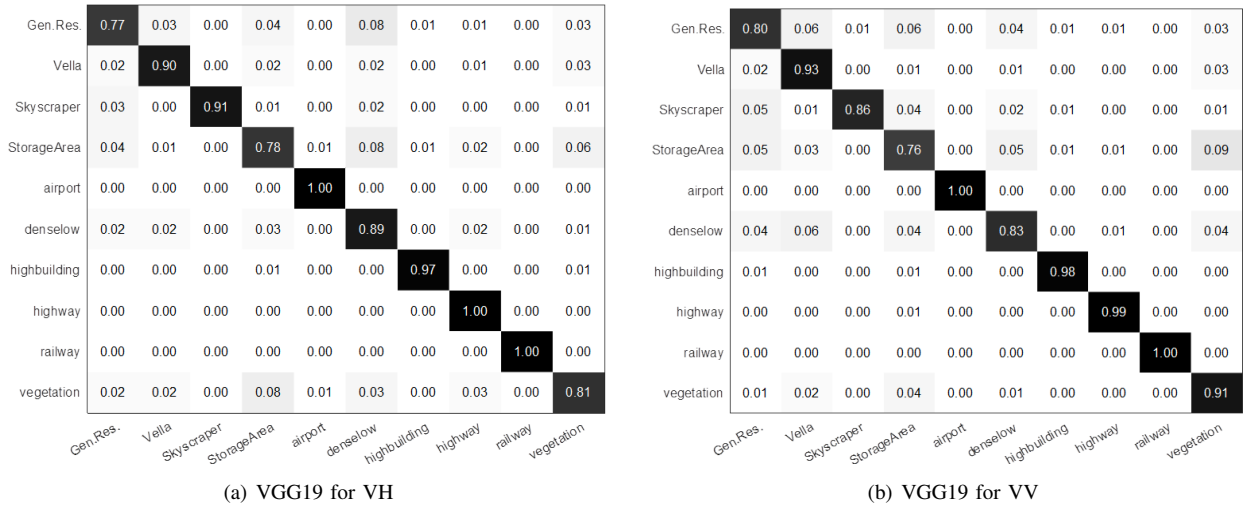


Fig. 13. The confusion matrix of VGG19 when classifying the whole dataset containing different polarizations. ‘Gen.Res.’ denotes general residential areas.

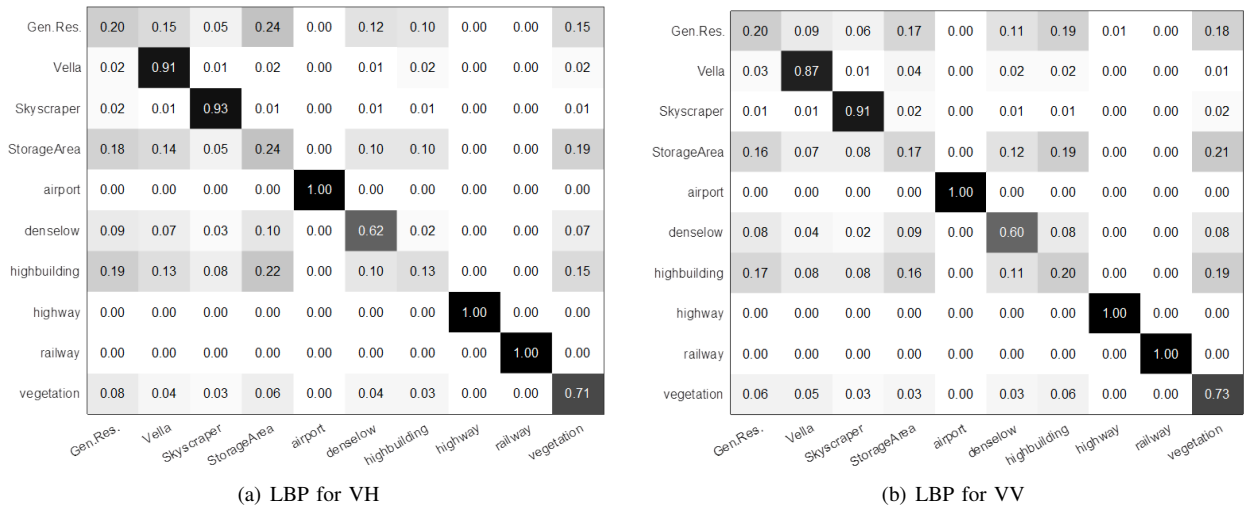


Fig. 14. The confusion matrix of LBP when classifying the whole dataset containing different polarizations. ‘Gen.Res.’ denotes the general residential areas.

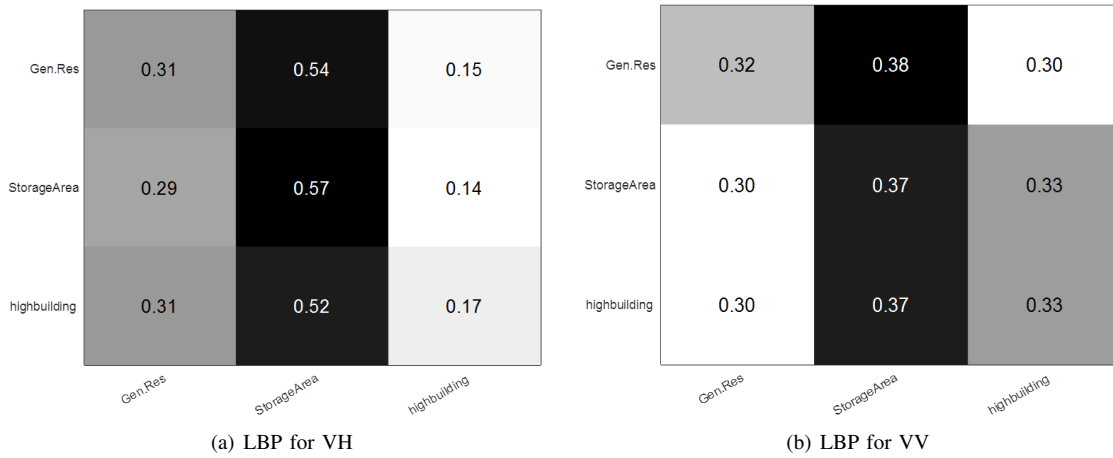


Fig. 15. Confusion matrix of the three typical building types by using LBP features. Both VH and VV polarization modes are evaluated.

4) *Industrial Area Evaluation*: Fig. 19 depicts the classification accuracy of each benchmarking algorithm for industrial storage areas, where both VH and VV polarizations are

included. For the identification of this category, the results listed in this figure tell us that: 1) The deep learning algorithms show their advantages over traditional methods, both for VH

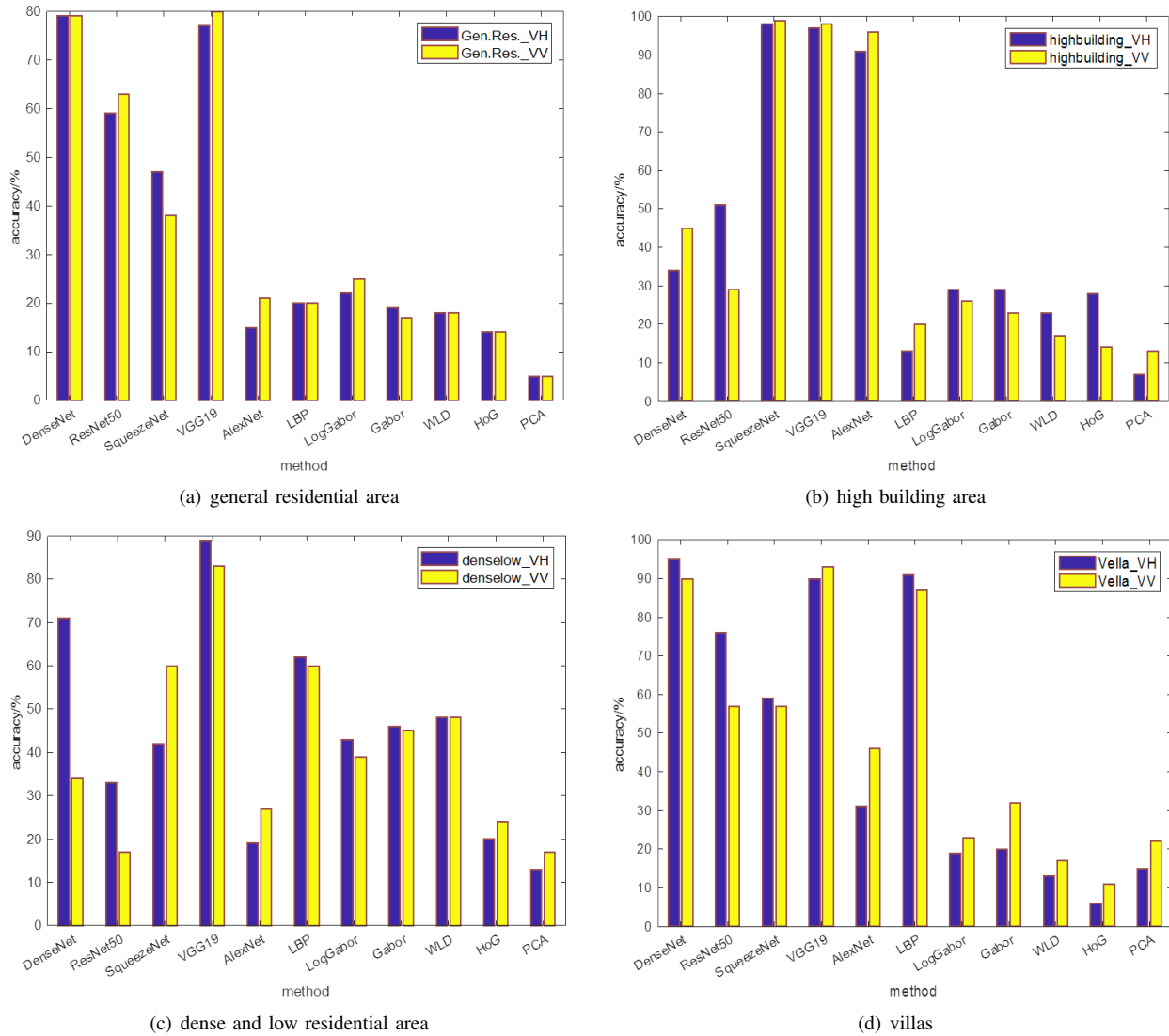


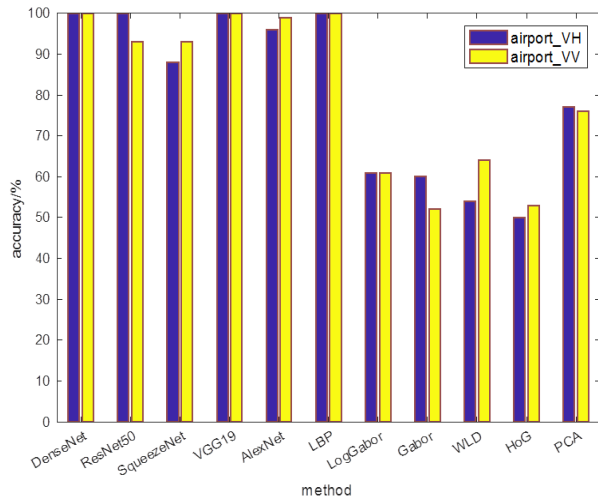
Fig. 16. Classification accuracy of residential areas, including general residential areas, high-rise building areas, dense and low-rise residential areas, and villas, by using eleven benchmarking algorithms. Both the VH and VV polarizations are compared and evaluated.

and VV polarizations; 2) DenseNet, ResNet50, and VGG19 can achieve almost the same classification accuracy for VH polarization mode, being roughly 78%; 3) DenseNet and ResNet50 have the most powerful capability to identify this category of VH polarization (with an accuracy of 86%) and VV polarization (with an accuracy of 81%), respectively.

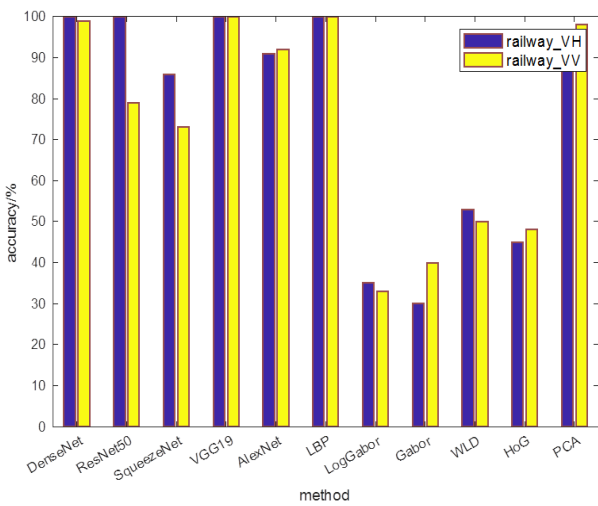
5) *Urban Vegetation Evaluation*: The classification accuracy for identifying urban vegetation areas, which account for a large amount of land covers in urban areas, by utilizing each benchmarking algorithm is illustrated in Fig. 20. This figure can be summarized based on the following points: 1) A satisfying performance can be achieved by using SqueezeNet, VGG19, AlexNet, and LBP for both VH and VV polarizations; 2) VGG19, which obtains the most convincing results for both polarizations of this category, is more suitable for VV polarization. The accuracy gap with results from VH-polarized data amounts to roughly 10%.

VII. CONCLUSIONS AND FUTURE WORK

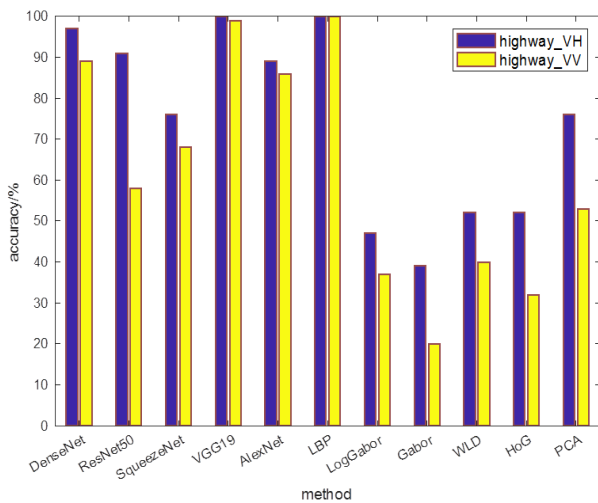
This paper describes a Sentinel-1 dataset for urban interpretation, called *OpenSARUrban*. The dataset is comprised of 10 different urban categories, including 4 kinds of formats and 2 kinds of polarization modes for each image patch, and covering 21 major cities of China. Specifically, the image formats of the original data, the visualized gray-scale data, the visualized data in pseudo-color, and the calibrated data are included. The polarization modes include VH and VV polarizations. With the five essential properties of large scale, diversity, specificity, reliability, and sustainability, the goals of this dataset can be achieved. The dataset structure is visualized from the perspective of data manifolds by using the FCD and the t-SNE. Finally, some benchmarking algorithms and experimental results are presented to demonstrate the practicality and the quality of this dataset. It is believed that developing methods enhancing the performance for the whole dataset is very challenging and the dataset is also expected to foster research on data imbalance. In the era of big data for the SAR



(a) airport



(b) railway



(c) highway

Fig. 17. Classification accuracy of transportation hubs, including airports, railways, and highways, by using eleven benchmarking algorithms. Both VH and VV polarizations are compared and evaluated.

community, *OpenSARUrban* is expecting to provide a dataset

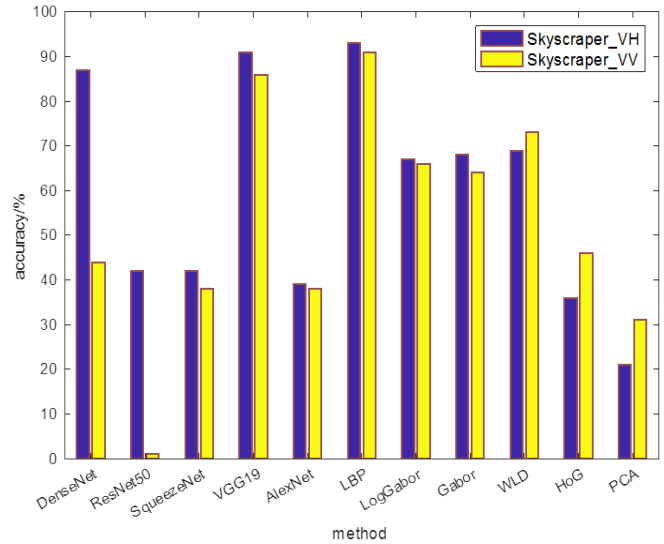


Fig. 18. Classification accuracy of business areas, represented by skyscrapers. Both VH and VV polarizations are compared and evaluated.

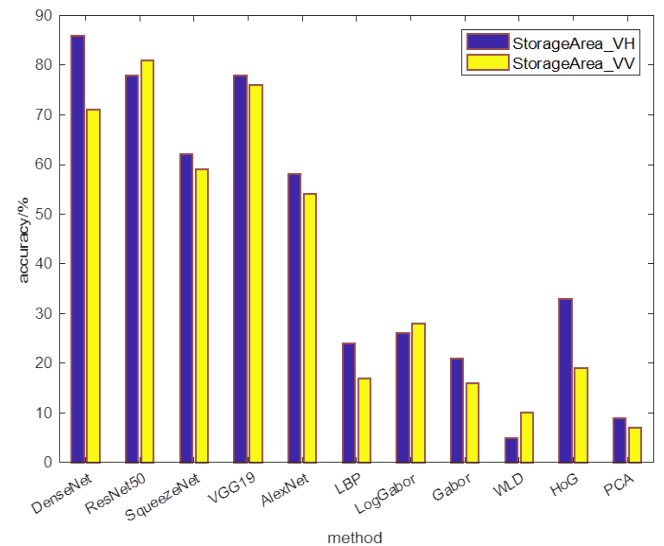


Fig. 19. Classification accuracy of industrial areas, represented by storage areas. Both VH and VV polarizations are compared and evaluated.

for developing much more advanced algorithms for Sentinel-1 urban interpretation and to foster the characterization of this kind of data. In future, this work will be extended throughout the world and also the time-series data.

The *OpenSARUrban* dataset can be found at <https://pan.baidu.com/s/1D2TzmUWePYHwTnHuHL7KdQ>.

ACKNOWLEDGMENT

The authors would like to thank the European Space Agency (ESA) for providing the Sentinel-1 data and the SNAP 3.0 software; we would like to thank Qianfanshijing Technology (Beijing) Co., Ltd. for providing the 91 Weitu software. The authors sincerely appreciate several colleagues in the Shanghai Key Laboratory of Intelligent Sensing and Recognition for

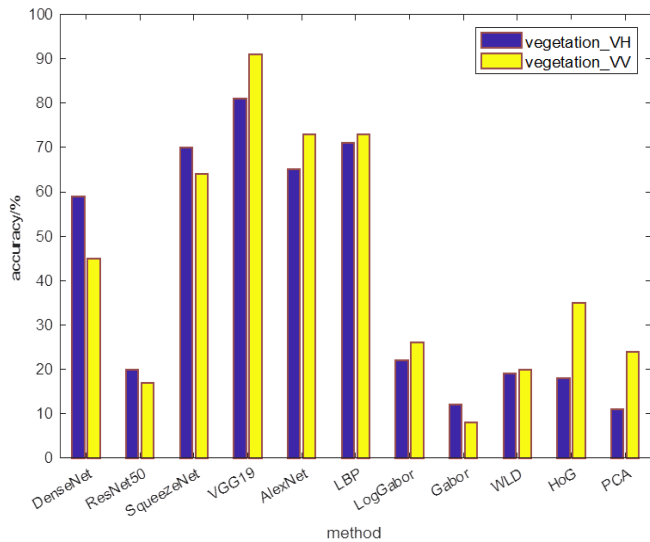


Fig. 20. Classification accuracy of urban vegetation. Both VH and VV polarizations are compared and evaluated.

their devoted assistance on this dataset annotation. The authors would also like to thank G. Schwarz for many helpful hints and the reviewers for their valuable and insightful comments.

REFERENCES

- [1] R. Torres, P. Snoeij, D. Geudtner, D. Bibby, M. Davidson, E. Attema, P. Potin, B. Rommen, N. Flourey, M. Brown *et al.*, "Gmes sentinel-1 mission," *Remote Sens. Environ.*, vol. 120, pp. 9-24, 2012.
- [2] R. Torres, P. Snoeij, M. Davidson, D. Bibby, and S. Lokas, "The sentinel-1 mission and its application capabilities," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Munich, Germany, 2012, pp. 1703-1706.
- [3] S. Abdikan, F. B. Sanli, M. Ustuner, and F. Calò, "Land cover mapping using sentinel-1 sar data," *The Int. Arch. Photogram., Remote Sens. Spatial Inf. Sci.*, vol. 41, p. 757, 2016.
- [4] E. Attema, M. Davidson, N. Flourey, G. Levrini, B. Rosich-Tell, B. Rommen, and P. Snoeij, "Sentinel-1 esa's new european sar mission," in *Sensors, Systems, and Next-Generation Satellites XI*, vol. 6744. Int. Soci. Opt. Photonics, 2007, p. 674403.
- [5] Y. Ban, A. Jacob, and P. Gamba, "Spaceborne sar data for global urban mapping at 30 m resolution using a robust urban extractor," *ISPRS J. Photogram. Remote Sens.*, vol. 103, pp. 28-37, 2015.
- [6] W. Wagner, D. Sabel, M. Doubkova, M. Hornáček, S. Schlaffer, and A. Bartsch, "Prospects of sentinel-1 for land applications," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Munich, Germany, 2012, pp. 1741-1744.
- [7] A. Barra, O. Monserrat, P. Mazzanti, C. Esposito, M. Crosetto, and G. Scarascia Mugnozza, "First insights on the potential of sentinel-1 for landslides detection," *Geomatics, Natural Hazards and Risk*, vol. 7, no. 6, pp. 1874-1883, 2016.
- [8] J. Reiche, E. Hamunyela, J. Verbesselt, D. Hoekman, and M. Herold, "Improving near-real time deforestation monitoring in tropical dry forests by combining dense sentinel-1 time series with landsat and alos-2 palsar-2," *Remote Sens. Environ.*, vol. 204, pp. 147-161, 2018.
- [9] N. Torbick, D. Chowdhury, W. Salas, and J. Qi, "Monitoring rice agriculture across myanmar using time series sentinel-1 assisted by landsat-8 and palsar-2," *Remote Sens.*, vol. 9, no. 2, p. 119, 2017.
- [10] A. Gruber, W. Wagner, A. Hegyiova, F. Greifeneder, and S. Schlaffer, "Potential of sentinel-1 for high-resolution soil moisture monitoring," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Melbourne, Australia, 2013, pp. 4030-4033.
- [11] S. Quan, B. Xiong, D. Xiang, L. Zhao, S. Zhang, and G. Kuang, "Eigenvalue-based urban area extraction using polarimetric sar data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 2, pp. 458-471, 2018.
- [12] D. Xiang, T. Tang, Y. Ban, Y. Su, and G. Kuang, "Unsupervised polarimetric sar urban area classification based on model-based decomposition with cross scattering," *ISPRS J. Photogram. Remote Sens.*, vol. 116, pp. 86-100, 2016.
- [13] D. Xiang, T. Tang, Y. Ban, and Y. Su, "Man-made target detection from polarimetric sar data via nonstationarity and asymmetry," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 4, pp. 1459-1469, 2016.
- [14] D. Xiang, T. Tang, C. Hu, Q. Fan, and Y. Su, "Built-up area extraction from polar sar imagery with model-based decomposition and polarimetric coherence," *Remote Sens.*, vol. 8, no. 8, p. 685, 2016.
- [15] F. Del Frate, I. Fabrini, M. Penalver, and M. Iapaolo, "Neumapper: A neural networks software for image classification," in *Proc. ESA-EUSC-JRC Image Inf. Min. Conf. ISPR*, 2011, pp. 1-4.
- [16] M. Datcu, G. Schwarz, S. Kiemle, A. De Miguel, A. Colapicchioni, C. Rosati, A. Galoppo, A. Valente, P. Harms, B. Bilgin *et al.*, "Pims: Knowledge based image information mining providing new functionalities in the terrasars ground segment system."
- [17] S. De, L. Bruzzone, A. Bhattacharya, F. Bovolo, and S. Chaudhuri, "A novel technique based on deep learning and a synthetic target database for classification of urban areas in polar sar data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 1, pp. 154-170, 2018.
- [18] T. Xiao, Y. Xu, K. Yang, J. Zhang, Y. Peng, and Z. Zhang, "The application of two-level attention models in deep convolutional neural network for fine-grained image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, 2015, pp. 842-850.
- [19] F. Hu, G.-S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery," *Remote Sens.*, vol. 7, no. 11, pp. 14680-14707, 2015.
- [20] Z. Deng, H. Sun, S. Zhou, J. Zhao, L. Lei, and H. Zou, "Multi-scale object detection in remote sensing imagery with convolutional neural networks," *ISPRS J. Photogram. Remote Sens.*, vol. 145, pp. 3-22, 2018.
- [21] Z. Deng, H. Sun, S. Zhou, and J. Zhao, "Learning deep ship detector in sar images from scratch," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 4021-4039, 2019.
- [22] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8-36, 2017.
- [23] F. Xu, H. Wang, and Y. Jin, "Deep learning as applied in sar target recognition and terrain classification," *J. Radars*, vol. 6, no. 2, pp. 136-148, 2017.
- [24] F. Dou, W. Diao, X. Sun, Y. Zhang, and K. Fu, "Aircraft reconstruction in high resolution sar images using deep shape prior," *J. Radars*, vol. 6, no. 5, pp. 503-513, 2017.
- [25] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst. (ACM GIS)*, San Jose, CA, USA, 2010, pp. 270-279.
- [26] B. Bechtel, L. See, G. Mills, and M. Foley, "Classification of local climate zones using sar and multispectral data in an arid environment," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 7, pp. 3097-3105, 2016.
- [27] G.-S. Xia, J. Hu, F. Hu, B. Shi, X. Bai, Y. Zhong, L. Zhang, and X. Lu, "Aid: A benchmark data set for performance evaluation of aerial scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3965-3981, 2017.
- [28] P. Jin, G.-S. Xia, F. Hu, Q. Lu, and L. Zhang, "Aid++: An updated version of aid on scene classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Valencia, Spain, 2018, pp. 4721-4724.
- [29] G.-S. Xia, X. Bai, J. Ding, Z. Zhu, S. Belongie, J. Luo, M. Datcu, M. Pelillo, and L. Zhang, "Dota: A large-scale dataset for object detection in aerial images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, Utah, 2018, pp. 3974-3983.
- [30] P. Helber, B. Bischke, A. Dengel, and D. Borth, "Introducing eurosat: A novel dataset and deep learning benchmark for land use and land cover classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Valencia, Spain, 2018, pp. 204-207.
- [31] E. R. Keydel, S. W. Lee, and J. T. Moore, "Mstar extended operating conditions: A tutorial," in *Algorithms for Synthetic Aperture Radar Imagery III*, vol. 2757. Int. Soci. Opt. Photonics, 1996, pp. 228-243.
- [32] S. Chen, H. Wang, F. Xu, and Y.-Q. Jin, "Target classification using the deep convolutional networks for sar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4806-4817, 2016.
- [33] J. Pei, Y. Huang, W. Huo, Y. Zhang, J. Yang, and T.-S. Yeo, "Sar automatic target recognition based on multiview deep learning framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2196-2210, 2018.

- [34] X. Yuan, T. Tang, D. Xiang, Y. Li, and Y. Su, "Target recognition in sar imagery based on local gradient ratio pattern," *Int. J. Remote Sens.*, vol. 35, no. 3, pp. 857-870, 2014.
- [35] C. O. Dumitru and M. Datcu, "Information content of very high resolution sar images: Study of feature extraction and imaging parameters," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 8, pp. 4591-4610, 2013.
- [36] C. O. Dumitru, G. Schwarz, and M. Datcu, "Sar image land cover datasets for classification benchmarking of temporal changes," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1571-1592, 2018.
- [37] L. Huang, B. Liu, B. Li, W. Guo, W. Yu, Z. Zhang, and W. Yu, "Opensarship: A dataset dedicated to sentinel-1 ship interpretation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 1, pp. 195-208, 2018.
- [38] M. Schmitt, L. H. Hughes, and X. X. Zhu, "The sen1-2 dataset for deep learning in sar-optical data fusion," [Online] Available: <https://arxiv.org/abs/1807.01569>, 2018.
- [39] F. Ye, W. Luo, M. Dong, H. He, and W. Min, "Sar image retrieval based on unsupervised domain adaptation and clustering," *IEEE Geosci. Remote Sens. Lett.*, 2019.
- [40] L. Jiao, X. Tang, B. Hou, and S. Wang, "Sar images retrieval based on semantic classification and region-based similarity measure for earth observation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 3876-3891, 2015.
- [41] X. Tang, L. Jiao, and W. J. Emery, "Sar image content retrieval based on fuzzy similarity and relevance feedback," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 1824-1842, 2017.
- [42] X. Tang and L. Jiao, "Fusion similarity-based reranking for sar image retrieval," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 2, pp. 242-246, 2016.
- [43] J. Kim, D. Han, Y.-W. Tai, and J. Kim, "Salient region detection via high-dimensional color transform," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, Ohio, 2014, pp. 883-890.
- [44] L. Xiao, J.-w. ZHENG, Z.-l. ZHANG, S. Rong-tao, and C. ZHANG, "Application of google earth software in remote sensing geology teaching," *CHINESE GEOLOGICAL EDUCATION*, vol. 52, no. 2, 2017.
- [45] A. D. Johnson, R. E. Handsaker, S. L. Pulit, M. M. Nizzari, C. J. O'donnell, and P. I. De Bakker, "Snap: a web-based tool for identification and annotation of proxy snps using hapmap," *Bioinf.*, vol. 24, no. 24, pp. 2938-2939, 2008.
- [46] O. Isoguchi, K. Ishizuka, T. Tadono, T. Motohka, and M. Shimada, "Effect of faraday rotation on l-band ocean normalized radar cross section and wind speed detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 2019.
- [47] D. Cerra and M. Datcu, "A fast compression-based similarity measure with applications to content-based image retrieval," *J. Visual Commun. Image Repres.*, vol. 23, no. 2, pp. 293-302, 2012.
- [48] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *J. Machine Learning Research*, vol. 9, no. Nov, pp. 2579-2605, 2008.
- [49] X. Qin, H. Zou, S. Zhou, and K. Ji, "Region-based classification of sar images using kullback-leibler distance between generalized gamma distributions," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 8, pp. 1655-1659, 2015.
- [50] W. Yao and M. Datcu, "study cases on fast compression distance based data visualization," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, accepted, Yokohama, Japan, 2019.
- [51] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM TIST*, vol. 2, no. 3, p. 27, 2011.
- [52] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, Hawaii, 2017, pp. 4700-4708.
- [53] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, Nevada, 2016, pp. 770-778.
- [54] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size," [Online] Available: <https://arxiv.org/abs/1602.07360>, 2016.
- [55] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," [Online] Available: <https://arxiv.org/abs/1409.1556>, 2014.
- [56] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [57] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Proc. Eur. Conf. Comput. Vis.* Springer, Prague, Czech Republic, 2004, pp. 469-481.
- [58] J. Arrospe and L. Salgado, "Log-gabor filters for image-based vehicle verification," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2286-2295, 2013.
- [59] S. Cui, C. O. Dumitru, and M. Datcu, "Ratio-detector-based feature extraction for very high resolution sar image patch indexing," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1175-1179, 2013.
- [60] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao, "Wld: A robust local image descriptor," *IEEE Trans. Pattern Ana. Machine Intell.*, vol. 32, no. 9, pp. 1705-1720, 2010.
- [61] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1. IEEE Computer Society, Los Alamitos, Calif, 2005, pp. 886-893.
- [62] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *J. Educat. Psych.*, vol. 24, no. 6, p. 417, 1933.
- [63] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deepplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 40, no. 4, pp. 834-848, 2017.
- [64] M. Dong, D. Liu, Z. Xiong, C. Yang, X. Chen, Z.-J. Zha, G. Bi, and F. Wu, "3d cnn-based soma segmentation from brain images at single-neuron resolution," in *IEEE Int. Conf. Image Process. (ICIP)*, Athens, Greece, 2018, pp. 126-130.
- [65] B. Demir and L. Bruzzone, "Histogram-based attribute profiles for classification of very high resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2096-2107, 2015.
- [66] B. Demir, L. Minello, and L. Bruzzone, "An effective strategy to reduce the labeling cost in the definition of training sets by active learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 79-83, 2013.
- [67] J. Hu, P. Ghamisi, and X. Zhu, "Feature extraction and selection of sentinel-1 dual-pol data for global-scale local climate zone classification," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 9, p. 379, 2018.



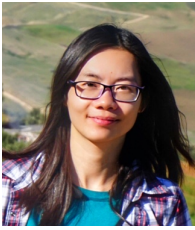
Juanping Zhao received the B.Sc. degree in electric information engineering from Xidian University, Xi'an, China, in 2014. Since then, she has been pursuing her Ph.D. degree in the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China. Since 2018, she has been a joint Ph.D. student in the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany.

Her research interests include SAR and PolSAR image interpretation, pattern recognition, and machine learning.



Zenghui Zhang (M'13) received the B.Sc. degree in applied mathematics, the M.Sc. degree in computational mathematics, and the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, China, in 2001, 2003, and 2008, respectively.

From 2008 to 2012, he was a Lecturer with the Department of Mathematics and System Science, NUDT. He is currently an Associate Professor with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China. His main research interests include radar image interpretation, radar signal processing, and compressed sensing theory and applications.



Wei Yao received the B.Eng. degree in biomedical engineering and the M.Sc. degree in pattern recognition and intelligent systems from Huazhong University of Science and Technology (HUST), Wuhan, China, in 2008 and 2011, respectively. She received the Ph.D. degree in electrical engineering at the University of Siegen, Siegen, Germany, in 2017.

She has been a researcher with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany, since 2014. Currently, she is working as a scientist at

DLR, involved in several projects in the frame of the European H2020 and Germany HGF Programs for data mining, data fusion, satellite imagery time series processing, and data visualization. Her research interests include image feature extraction, SAR image statistical modeling, hierarchical clustering, deep learning applied in remote sensing and data visualization.



Mihai Datcu (SM'04-F'13) received the M.S. and Ph.D. degrees in electronics and telecommunications from the University Politehnica Bucharest (UPB), Bucharest, Romania, in 1978 and 1986, respectively.

From 1992 to 2002, he had a longer Invited Professor assignment with the Swiss Federal Institute of Technology, ETH Zurich. He has held Visiting Professor appointments with the University of Oviedo, Spain; the University Louis Pasteur, Strasbourg, France; the International Space University, Strasbourg, France; University of Siegen, Germany; University of Innsbruck, Austria; University of Alcalá, Spain; University Tor Vergata, Rome, Italy; Universidad Pontificia de Salamanca, campus de Madrid, Spain; University of Camerino, Italy; and the Swiss Center for Scientific Computing, Manno, Switzerland. Since 1981, he has been a Professor with the Department of Applied Electronics and Information Engineering, Faculty of Electronics, Telecommunications and Information Technology, UPB, working on image processing and Electronic Speckle Interferometry. Since 1993, he has been a Scientist with the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. He is developing algorithms for model-based information retrieval from high-complexity signals and methods for scene understanding from very-high resolution synthetic aperture radar (SAR) and Interferometric SAR data. He is currently a Senior Scientist and Image Analysis Research Group Leader with the Remote Sensing Technology Institute, DLR, Wessling, Germany. Since 2011, he has been leading the Immersive Visual Information Mining Research Lab, Munich Aerospace Faculty, and has been the Director of the Research Center for Spatial Information, UPB. Since 2001, he has been initiating and leading the Competence Centre on Information Extraction and Image Understanding for Earth Observation, ParisTech, Paris Institute of Technology, Telecom Paris, Paris, France, a collaboration of DLR with the French Space Agency (CNES). He has been a Professor of the DLR-CNES Chair, ParisTech, Paris Institute of Technology, Telecom Paris. He initiated the European frame of projects for Image Information Mining (IIM) and is involved in research programs for information extraction, data mining and knowledge discovery, and data understanding with the European Space Agency (ESA), NASA, and in a variety of national and European projects. He and his team have developed and are currently developing the operational IIM processor in the Payload Ground Segment systems for the German missions TerraSAR-X, TanDEM-X, and the ESA Sentinel-1 and -2. He is the author of more than 450 scientific publications; among them 80 are journal papers, and a book on number theory. He has served as a Co-organizer of international conferences and workshops, and as a Guest Editor of a special issue on IIM of the IEEE and other journals. He is involved in research related to information theoretical aspects and semantic representations in advanced communication systems. His research interests include Bayesian inference, information and complexity theory, stochastic processes, model-based scene understanding, image information mining, for applications in information retrieval, and understanding of high-resolution SAR and optical observations.

Dr. Datcu, in 2006, received the Best Paper Award, IEEE Geoscience and Remote Sensing Society Prize; in 2008, the National Order of Merit with the rank of Knight, for outstanding international research results, awarded by the President of Romania; and in 1987, the Romanian Academy Prize Traian Vuia for the development of the SAADI image analysis system and his activity in image processing. He is a Member of the ESA Working group Big Data from space (BiDS). In 1999, he received the title *Habilitation à diriger des recherches* in computer science from the University Louis Pasteur, Strasbourg, France.



Huilin Xiong (M'04) received the B.Sc. and M.Sc. degrees in mathematics from Wuhan University, Wuhan, China, in 1986 and 1989, respectively, and the Ph.D. degree in pattern recognition and intelligent control from the Institute of Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan, in 1999.

He was a Post-Doctoral Researcher with Concordia University, Montreal, QC, Canada, and the University of Kansas, Lawrence, KS, USA, from 2001 to 2006. In 2007, he joined Shanghai Jiao Tong University, Shanghai, China, where he is currently a Professor with the School of Electronic Information and Electrical Engineering. His current research interests include pattern recognition, machine learning, and bioinformatics.



Wenxian Yu was born in Shanghai, China, in 1964. He received the B.Sc. degree in radio measurement and control and data transmission, the M.Sc. degree in communication and electronic system, and the Ph.D. degree in communication and information system from the National University of Defense Technology (NUDT), Changsha, China, in 1985, 1988, and 1993, respectively.

From 1996 to 2008, he was a Professor with the College of Electronic Science and Engineering, NUDT, where he served as the Deputy Head of the College and the Assistant Director of the National Key Laboratory of Automatic Target Recognition. He was the Executive Dean of Shanghai Jiao Tong University, Shanghai, from 2009 to 2011. He is currently with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, where he is a Yangtze River Scholar Distinguished Professor and the Head of Research. His current research interests include radar target recognition, remote sensing information processing, multisensor data fusion, and integrated navigation system. In these areas, he has published over 200 research papers.