Analyzing the Separability of SAR Classification Dataset in Open Set Conditions

Ning Liao, *Student Member, IEEE*, Mihai Datcu, *Fellow, IEEE*, Zenghui Zhang, *Senior Member, IEEE*, Weiwei Guo, Juanping Zhao, and Wenxian Yu, *Senior Member, IEEE*

***Abstract*—In synthetic aperture radar (SAR) image classification applications, there are three categories of data, including training and benchmark data with fixed classes, as well as actual data in practical applications. A real problem comes that there exist unknown classes not included in training and benchmark data, which is defined as the open set condition. However, little work on recognizing unknown classes and analyzing the separability of SAR datasets has been developed. Motivated by this observation, this article demonstrates the difficulty of practical classification and analyzes SAR dataset separability in open set conditions. In this article, the SAR separability analyzer (SAR-SA) is proposed to model each known class as a multivariate Gaussian distribution. SAR-SA can classify the known classes and recognize the samples locating in each known distribution with low probabilities as unknown. Besides, SAR datasetwise separability index (DSI) and classwise separability index (CSI) are defined to quantify the separability in open set conditions at the dataset level and class level. DSI and CSI are effective indicators of the difficulty of SAR classification datasets. Extensive experimental results demonstrate that the DSI in open set conditions is nearly half of that in supervised conditions. Dataset with low DSI is hard to realize accurate classification in open set conditions. At the class level, even though the SAR image classes are semantically different from each other, there exists more or less overlap between the distributions of supervised known classes and unknown classes. Classes with low CSI are harder to be correctly classified and recognized.**

***Index Terms*—Classwise separability index (CSI), datasetwise separability index (DSI), multivariate Gaussian distribution, open set condition, synthetic aperture radar (SAR) image classification, SAR separability analyzer (SAR-SA).**

I. Introduction

SYNTHETIC aperture radar (SAR) is an active sensor, which works well in all-weather and all-day conditions and provides low to high-resolution remote sensing images. SAR image classification is one of the fundamental techniques with a wide range of applications in many fields over the past decades [1], including land cover classification, urban interpretation, and target recognition.

In order to support SAR image classification research, a large number of SAR image classification datasets have been published and widely used. For land cover classification and urban interpretation, Dumitru *et al.* [2] generated a high-resolution TerraSAR-X dataset and a Sentinel-1 SAR classification band dataset. Their TerraSAR-X dataset covers the urban, industrial areas, and infrastructure from all over the world. The number of classes ranges from 5 to 20 depending on the local architecture of the countries. In the Sentinel-1 dataset, the number of classes varies from 5 to 10 specific to the image location and local urban structures. Zhao *et al.* [3] presented the OpenSARUrban dataset for urban interpretation which covers areas of 21 major cities in China. There are 5 functional classes and a total of 10 classes in OpenSARUrban. For automatic target recognition, moving and stationary target acquisition and recognition (MSTAR) dataset [4] is a benchmark dataset which consists of 10 different classes of targets from X-band images with 0.3 m resolution. The OpenSARShip [5] contains ship chips covering 17 different classes from Sentinel-1 SAR images. In addition to the above datasets, other SAR classification datasets [6]–[9] have also been proposed for the SAR image classification task.

Based on the published SAR datasets, various methods [10]–[13] have been proposed for classification. Cui *et al.* [14] presented a simple yet efficient feature extraction method using bag-of-words framework. It yields a classification performance with accuracy beyond 90% on the TerraSAR-X dataset containing 15 different urban classes. Geng [15] designed a deep convolutional autoencoder, which achieved overall accuracy with nearly 90% on TerraSAR-X images consisting of 5 classes. Huang *et al.* [16] developed a convolutional neural network (CNN) combining classification, reconstruction, and feedback bypass together, which realized accuracy with 99.05% on MSTAR dataset. He *et al.* [17] integrated densely connected triplet CNNs with Fisher discrimination regularized metric learning, achieving the average recognition accuracy of 88.97% on medium resolution Sentinel-1 SAR images obtained from OpenSARShip [5].

However, the existing training and benchmark datasets only contain a limited number of classes. In practical conditions, there exist many other data of the classes that are not included in the training and benchmark data. This practical condition is called the open set condition. When these methods are applied

Manuscript received April 28, 2021; revised June 24, 2021; accepted July 23, 2021. Date of publication July 27, 2021; date of current version August 18, 2021. This work was supported in part by the National Natural Science Foundation of China under Grants 62071333 and U1830103, and in part by ESA-MOST CHINA Dragon 5 programme ID.58190. *(Corresponding author: Mihai Datcu.)*

Ning Liao, Zenghui Zhang, Juanping Zhao, and Wenxian Yu are with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: liaoning@sjtu.edu.cn; zenghui.zhang@sjtu.edu.cn; juanpingzhao@sjtu.edu.cn; wxyu@sjtu.edu.cn).

Mihai Datcu is with the Research Center for Spatial Information (CEOSapceTech), University POLITEHNICA of Bucharest (UPB), Bucharest 060042, RomaniaRemote Sensing Technology Institute (IMF), German Aerospace Center(DLR), Oberpfaffenhofen D-82234, Germany (e-mail: mihai.datcu@dlr.de).

Weiwei Guo is with the Center of Digital Innovation, Tongji University, Shanghai 200092, China (e-mail: weiweiguo@tongji.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3100342

in open set conditions, the classification performance will be greatly reduced. To handle this challenge, open set recognition (OSR) [18] methods have been developed. OSR means not only to classify the known classes seen in the training phase but also to recognize unknown classes not participated in training as unknown. Many previous works have been proposed such as reconstruction-based methods [19], [20] and prototype network-based methods [21]. Besides, statistic modeling [22], [23] is utilized for new class discovery via out-of-distribution detection. The authors in [24] and [25] proposed a study to recognize whether a sample is unknown by distance comparing. In remote sensing, a representative-discriminative learning framework [26] is designed for OSR for hyperspectral images. The framework learned representative features at first and then learn discriminative features for classification.

However, in the SAR community, there is little work aiming at SAR OSR [27]. It seems that there is no work to analyze the reasons why the classification task in SAR in open set conditions is difficult. Fundamentally, worthy insights into data can be obtained by measuring the separability. Few works focused at assessing the dataset itself by analyzing the separability of data has been developed, which indicates the difficulty of classification. The class scatter matrix [28] is widely used to measure the separability of a whole dataset combining the between-class scatter and within-class scatter together. Based on the class scatter matrix, a regularization strategy is developed to improve the separability evaluation [29]. Separability index (SI) [30] is proposed to measure the class overlap directly. Similarly, a lot of extended evaluation indicators have also been put forward and analyzed [31], [32]. The classification will be easier to achieve high performance if the data separability is large. However, these indicators only focus on the supervised classification problems.

Hence, analyzing the reasons for the difficulty of SAR classification in open set conditions is very necessary and will be of great help to carry out SAR OSR. In this article, we concentrate on evaluating the separability of SAR datasets in open set conditions, which is more helpful to the practical applications.

SAR images in different data types follow different statistical distribution models. In this article, we focus on the features extracted from the raw data. Considering the complexity of the spatial distribution of the original images and the low-dimensional nature of the semantic information of classes, we model each known class in SAR datasets as a multivariate Gaussian distribution in the learned feature space in [33]. This article splits the dataset into known classes and unknown classes to simulate the practical condition that there exist unknown classes. If a testing sample locates in the low probability space of each known distribution, it will be recognized as unknown. Successively, two separability indices are defined to measure the separability in open set conditions at the dataset level and class level. The major contributions of this article are as follows.

1) By modeling each known class as a multivariate Gaussian distribution, SAR separability analyzer (SAR-SA) is proposed for the known class classification and unknown class recognition.
2) Taking the idea of class scatter matrix, datasetwise separability index (DSI) is defined to quantify the separability of a dataset in open set conditions at the dataset level. The classification difficulty of different datasets in open set conditions can be obtained by comparing their DSI.
3) Combining the precision and recall results, classwise separability index (CSI) is defined by using the F_2 score to quantify the separability of each class at the class level in open set conditions. CSI can measure the potential of a class to be accurately classified and recognized in open set conditions.
4) We proposed an evaluation pipeline for two SAR image datasets. The experimental results demonstrate that the DSI in open set conditions is nearly half of that in the supervised condition. Classes with low CSI are harder to be correctly classified and recognized.

The rest of this article is organized as follows. In Section II, we introduce the architecture, training, and testing strategies of SAR-SA. In Section III, SAR DSI and CSI are defined. Section IV gives an introduction to the datasets prepared for experiments. Experimental design and detailed result analysis are shown in Section V. Finally, Section VI concludes this article.

II. SAR Separability Analyzer

In this section, first, the preliminaries of variational autoencoder (VAE) is briefly introduced. Then we describe the architecture of SAR-SA. The strategies of SAR-SA in the training and testing phase are introduced in the end.

A. Preliminaries of VAE

The vanilla VAE [34] is a powerful Bayesian method for learning the latent distribution of data. VAE is composed of an encoder and a decoder. The encoder network encodes the input data $\boldsymbol{x}$ to latent code $\boldsymbol{z}$. Standard Gaussian prior $p(\boldsymbol{z})$ is placed on the latent code. The decoder network maps the latent code $\boldsymbol{z}$ to the reconstructed data $\tilde{\boldsymbol{x}}$. The process of VAE is formulated as

$$\begin{aligned} \mathrm{En}(\boldsymbol{x}) &= \boldsymbol{z} \sim q_\phi(\boldsymbol{z} \mid \boldsymbol{x}) \\ \mathrm{De}(\boldsymbol{z}) &= \tilde{\boldsymbol{x}} \sim p_{\boldsymbol{\theta}}(\tilde{\boldsymbol{x}} \mid \boldsymbol{z}). \end{aligned} \tag{1}$$

In (1), $\mathrm{En}(\boldsymbol{x})$ denotes the encoder process, $\mathrm{De}(\boldsymbol{z})$ denotes the decoder process. $q_{\boldsymbol{\phi}}(\boldsymbol{z} \mid \boldsymbol{x})$ is the approximated posterior distribution specific to the input data $\boldsymbol{x}$, $p_{\boldsymbol{\theta}}(\tilde{\boldsymbol{x}} \mid \boldsymbol{z})$ is the likelihood. The parameters of encoder and decoder are denoted by $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$, respectively. VAE is optimized by maximizing the evidence lower bound (ELBO) [35] defined as

$$\begin{aligned} \mathrm{ELBO}(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{x}, \tilde{\boldsymbol{x}}) = \mathbb{E}_{q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x})} \left[\log p_{\boldsymbol{\theta}}(\tilde{\boldsymbol{x}} \mid \boldsymbol{z})\right] \\ - D_{KL}\left(q_{\boldsymbol{\phi}}(\boldsymbol{z} \mid \boldsymbol{x}) \| p(\boldsymbol{z})\right) \end{aligned} \tag{2}$$

in which $\mathbb{E}$ is expectation calculation.

The first term on the right side of (2) is the likelihood between the raw input data and reconstructed data, it is represented by the reconstruction loss. Maximizing the likelihood equals minimizing the reconstruction loss. The second term is the Kullback–Leibler (KL) divergence between the approximated posterior and prior distribution. It can be viewed as a regularization in

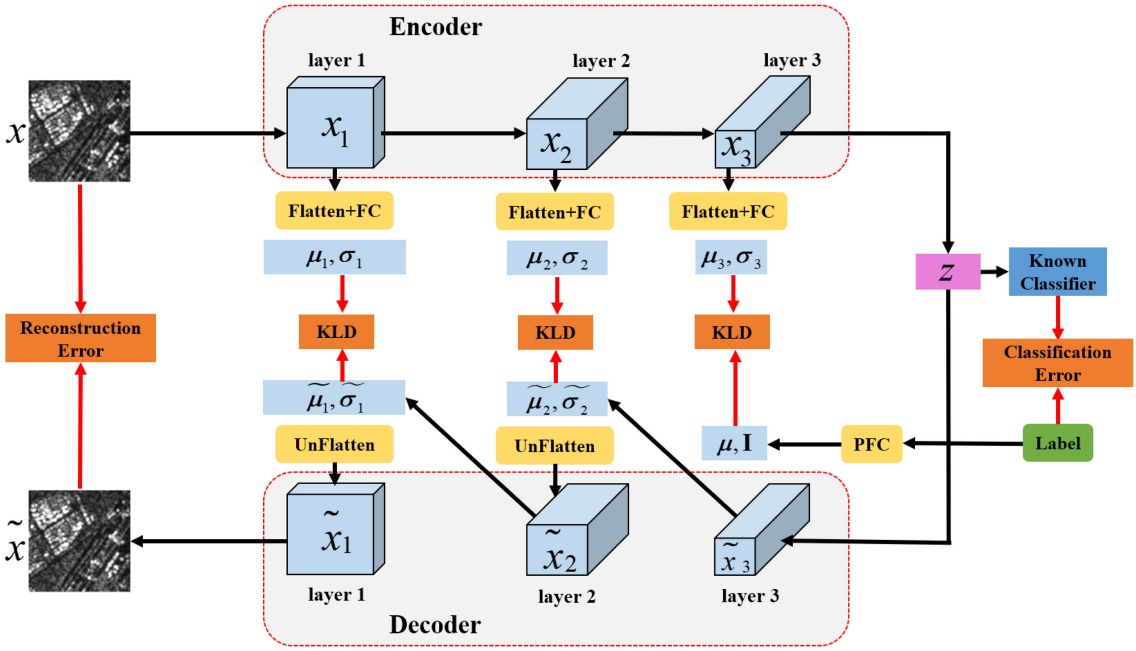


Fig. 1. Architecture of SAR separability analyzer (SAR-SA). SAR-SA is composed of an encoder, classifier, decoder, and PFC. It models the feature in each encoder and decoder layer as a multivariate Gaussian distribution. The encoder takes the raw image as input and outputs the parameters of Gaussian distribution in each encoder layer. Classifier takes the latent representation in the top encoder layer as input for classification. The decoder takes the representation as input and outputs the parameters of Gaussian distribution in each decoder layer. The raw image is reconstructed in the final decoder layer. PFC takes the one-hot label as input and outputs the mean vector of the class-specific Gaussian distribution.

the training phase. In D_{KL}, $q_{\boldsymbol{\phi}}(\boldsymbol{z} \mid \boldsymbol{x})$ is denoted as $N(\boldsymbol{\mu}, \boldsymbol{\sigma}^2\mathbf{I})$ with the learned mean and variance denoted as $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}^2$. $p(\boldsymbol{z})$ is assumed to be a standard Gaussian distribution $N(\mathbf{0}, \mathbf{I})$, in which $\mathbf{I}$ is an identity matrix. Assuming the dimension of the Gaussian distribution is d, the KL divergence is expressed as

$$\begin{aligned} &- D_{KL}\left(q_{\boldsymbol{\phi}}(\boldsymbol{z} \mid \boldsymbol{x}) \| p(\boldsymbol{z})\right) \\ &= \frac{1}{2} \sum_{i=1}^{d} \left(1 + \log\left(\sigma_i^2\right) - \mu_i^2 - \sigma_i^2\right). \end{aligned} \tag{3}$$

Hence, one goal of VAE is to minimize the discrepancy between the input and reconstructed data. It also prompts the posterior distribution to match the prior distribution.

B. SAR-SA Architecture

Referring to the conditional Gaussian distribution learning in [33], SAR-SA is proposed to learn the latent multivariate Gaussian distribution of each known class by introducing the class information into the model. To learn a more refined latent distribution of each SAR class, KL divergence in each paired encoder and decoder layer are all calculated and minimized. The architecture of SAR-SA is shown in Fig. 1. It is composed of an encoder, classifier, decoder, and a prior fully connected layer (PFC).

1) Encoder: The forward path of the encoder is designed as a series of convolution, batch normalization, and activation. Orthogonal to the forward path, each encoder layer adopts the probabilistic ladder [36] to obtain the mean and variance vector of the multivariate Gaussian specific to the encoder layer. The kth encoder layer is denoted as

$$\begin{aligned} \boldsymbol{x}_k &= \operatorname{Conv}\left(\boldsymbol{x}_{k-1}\right) \\ \boldsymbol{\mu}_k &= \operatorname{Linear}(\operatorname{Flatten}\left(\boldsymbol{x}_k\right)) \\ \boldsymbol{\sigma}_k^2 &= \mathrm{F}\left(\operatorname{Linear}\left(\operatorname{Flatten}\left(\boldsymbol{x}_k\right)\right)\right). \end{aligned} \tag{4}$$

Conv takes the feature $\boldsymbol{x}_{k-1}$ in previous layer for convolution. Flatten is adopted to flatten the 3D convolution output $\boldsymbol{x}_k$ to a 1D data. F is a nonlinear function denoted as $\log(1+\exp(\cdot))$, ensuring the variance no less than 0. Mean and variance in the kth layer, expressed as $\boldsymbol{\mu}_k$ and $\boldsymbol{\sigma}_k^2$, are learned by feeding the flattened 1D data into two linear layers.

Using the reparameterization trick, the latent representation $\boldsymbol{z}$ is obtained as $\boldsymbol{z} = \boldsymbol{\mu} + \boldsymbol{\varepsilon} \cdot \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I})$ with $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ learned in the top encoder layer, in which $\mathbf{I}$ is an identity matrix.

2) Classifier: The classifier is a linear layer followed by a softmax layer. It takes the latent representation $\boldsymbol{z}$ as input and the softmax layer outputs the probability distribution over each known class.

3) Decoder: The decoder is composed of a series of activation, transposed convolution, and batch normalization. The kth decoder layer is denoted as

$$\begin{aligned} \boldsymbol{d}_{k+1} &= \operatorname{Unflatten}\left(\operatorname{Linear}(\tilde{\boldsymbol{z}}_{k+1})\right) \\ \tilde{\boldsymbol{x}}_{k+1} &= \operatorname{TConv}\left(\boldsymbol{d}_{k+1}\right) \\ \tilde{\boldsymbol{\mu}}_k &= \operatorname{Linear}\left(\operatorname{Flatten}\left(\tilde{\boldsymbol{x}}_{k+1}\right)\right) \\ \tilde{\boldsymbol{\sigma}}_k^2 &= \mathrm{F}\left(\operatorname{Linear}\left(\operatorname{Flatten}\left(\tilde{\boldsymbol{x}}_{k+1}\right)\right)\right) \\ \tilde{\boldsymbol{z}}_k &= \tilde{\boldsymbol{\mu}}_k + \boldsymbol{\varepsilon} \cdot \tilde{\boldsymbol{\sigma}}_k^2, \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}). \end{aligned} \tag{5}$$

The 1D data $\tilde{\boldsymbol{z}}_{k+1}$ is first, fed into a linear layer and then be unflattened to a 3D data $\boldsymbol{d}_{k+1}$. TConv is a transposed convolution, it takes the 3D data as input and the size of the feature $\tilde{\boldsymbol{x}}_{k+1}$ is increased. The mean and variance in the *k*th decoder layer expressed as $\tilde{\boldsymbol{\mu}}_k$ and $\tilde{\boldsymbol{\sigma}}_k^2$ are learned through two linear layers, forming the multivariate Gaussian distribution specific to the *k*th decoder layer. In the final decoder layer, the input image is reconstructed.

4) PFC: In vanilla VAE, the prior distribution of latent code is set as standard Gaussian distribution. Differently, in order to introduce the class semantic information into the framework, PFC takes the one-hot encoding of the labels as input. The output $\boldsymbol{\mu}$ through a linear layer is the mean of the prior distribution of each sample conditioned on its class. The convariance matrix of the prior distribution is an identity matrix.

C. Strategies in Training and Testing Phase

1) Training: In the training phase, SAR-SA is optimized by minimizing the reconstruction error, classification error, and KL divergence jointly. Thus, the loss function is composed of the following three parts.

(a) The reconstruction loss L_r is the L_1 distance between the input image $\boldsymbol{x}$ and the final reconstructed image $\tilde{\boldsymbol{x}}$

$$L_r = \sum |\boldsymbol{x} - \tilde{\boldsymbol{x}}|. \tag{6}$$

(b) The classification loss L_c is the softmax cross-entropy between the prediction and the ground-truth labels

$$L_c = -\sum \boldsymbol{y}_c \log(\boldsymbol{p}_c), \tag{7}$$

in which $\boldsymbol{y}_c$ is the label for a sample. $\boldsymbol{p}_c$ is the predicted probability distribution on known classes.

(c) The KL-divergence L_{KL} in the latent space and each paired encoder–decoder layer. With the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}^2$ learned in the top encoder layer, the approximated posterior Gaussian distribution is denoted as $q_{\boldsymbol{\phi}} = N(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\sigma}^2)$. For each sample, the prior multivariate Gaussian distribution for the *l*th class is denoted as $p_{\boldsymbol{\theta}}^{(l)}(\boldsymbol{z} \mid \boldsymbol{x}) = N(\boldsymbol{z}; \boldsymbol{\mu}_l, \mathbf{I})$ with $\boldsymbol{\mu}_l$ learned from PFC. Compared with (3), assuming the dimension of the multivariate Gaussian is d, the KL divergence is formulated as

$$-D_{KL}\left(q_{\boldsymbol{\phi}}(\boldsymbol{z} \mid \boldsymbol{x}) \| p_{\boldsymbol{\theta}}^{(l)}(\boldsymbol{z} \mid \boldsymbol{x})\right) = \frac{1}{2}\sum_{i=1}^{d}\left(1 + \log\left(\sigma_i^2\right) - (\mu_i - \mu_i^{(l)})^2 - \sigma_i^2\right). \tag{8}$$

The combination of (6) and (8) forms the ELBO of vanilla VAE as described in (2). The only difference is that the combination of (6) and (8) forms the ELBO with leading the class information into the framework.

In each paired encoder–decoder layer of SAR-SA, the distributions are multivariate Gaussian distributions with different dimensions. Assuming the number of encoder or decoder layers is K, the sum of KL divergence between the posterior and class-specific prior in the top layer, and KL divergence in each paired encoder–decoder layer form the third part of the loss function as

$$L_{KL} = \frac{1}{K}\left[D_{KL}\left(q_{\boldsymbol{\phi}}(\boldsymbol{z} \mid \boldsymbol{x}) \| p_{\boldsymbol{\theta}}^{(l)}(\boldsymbol{z} \mid \boldsymbol{x})\right) + \sum_{k \neq K} D_{KL}\left(q_{\boldsymbol{\phi}}\left(\boldsymbol{x}_k \mid \boldsymbol{x}\right) \| p_{\boldsymbol{\theta}}\left(\tilde{\boldsymbol{x}}_k \mid \tilde{\boldsymbol{x}}_{k+1}\right)\right)\right]. \tag{9}$$

Thus, the training loss function is summarized as

$$L = L_r + L_c + L_{KL}. \tag{10}$$

2) Testing in Supervised Settings: SAR-SA is trained for multiple epochs for optimization in training phase, the learned parameters of SAR-SA with the lowest loss are preserved for testing, the model with these parameters is selected as the best model. After the training phase, the best model is utilized for classification directly.

3) Testing in Open Set Settings: After the training phase, the mean and variance of the latent representations of all the correctly classified training samples are obtained. The parameters of the class-specific multivariate Gaussian distribution are denoted as $\boldsymbol{m}_l$ and $\boldsymbol{\sigma}_l^2$ for the *l*th class. The multivariate Gaussian distribution of the *l*th known class is modeled as $f_l(\boldsymbol{z}) = N(\boldsymbol{z}; \boldsymbol{m}_l, \boldsymbol{\sigma}_l^2)$. Besides, the mean and standard deviation of the reconstruction loss of all the correctly classified training samples are acquired.

In the testing phase, the results are, first, predicted as in supervised settings. Regardless of whether there exist unknown classes in the testing data, the model will only choose one of the known classes that it thinks best matches as the prediction. If there exist unknown classes in the testing data, the unknown data will all be falsely classified. In order to make a tradeoff between known class classification and unknown class recognition, results are calibrated following two steps. The first calibration step is based on the reconstruction error. The second calibration step is to judge whether the probability of the latent representation locating in each known class distribution is lower than the threshold.

In the first calibration step, we set the upper bound of the reconstruction error as

$$\text{rec}_{\text{upb}} = \text{mean}(\text{rec}_{\text{train}}) + \lambda \cdot \text{std}(\text{rec}_{\text{train}}) \tag{11}$$

where mean and std represent calculating mean and standard deviation of the reconstruction errors $\text{rec}_{\text{train}}$. λ is the weight parameter of the standard deviation.

If the reconstruction error of a testing sample is larger than the threshold in (11), which means that this sample is not well reconstructed under the feature extraction and reconstruction framework of known classes. Thus, the testing sample will be recognized as unknown for the first.

Second, in terms of the remaining testing samples whose reconstruction errors are less than the upper bound. The probability of the extracted latent representation $\boldsymbol{z}$ locating in the *l*th known distribution $f_l(\boldsymbol{z}) = N(\boldsymbol{z}; \boldsymbol{m}_l, \boldsymbol{\sigma}_l^2)$ with dimension as d is calculated as

$$P_l(\boldsymbol{z}) = 1 - \int_{m_1 - |z_1 - m_1|}^{m_1 + |z_1 - m_1|} \cdots \int_{m_d - |z_d - m_d|}^{m_d + |z_d - m_d|} f_l(\boldsymbol{t}) d\boldsymbol{t} \tag{12}$$

in which the subscripts of m and z denote the dimension index.

If the probability calculated in (12) for each known class is lower than the threshold we set, which means the latent representation of the testing sample locates in the low probability space of each known distribution, the testing sample will be recognized as unknown.

III. SAR Separability Index

In this section, SAR separability index in datasetwise and classwise are introduced. SAR DSI takes the idea in [28] by using the scatter-matrix-based measure. SAR CSI adopts the F_2 score [37] which puts more significance on the recall rate than the precision. The two indices are introduced as follows.

A. SAR Datasetwise Separability Index

Taking the idea of scatter-matrix [28], the within-class scatter matrix $\boldsymbol{S}_W$, between-class scatter matrix $\boldsymbol{S}_B$ and total scatter matrix $\boldsymbol{S}_T$ are obtained as

$$\begin{aligned}\mathbf{S}_W &= \sum_{i=1}^{C}\left[\sum_{j=1}^{n_i}\left(\mathbf{v}_j^{(i)}-\mathbf{m}_i\right)\left(\mathbf{v}_j^{(i)}-\mathbf{m}_i\right)^{\top}\right]\\ \mathbf{S}_B &= \sum_{i=1}^{C} n_i\left(\mathbf{m}_i-\mathbf{m}\right)\left(\mathbf{m}_i-\mathbf{m}\right)^{\top}\\ \mathbf{S}_T &= \sum_{i=1}^{C}\left[\sum_{j=1}^{n_i}\left(\mathbf{v}_j^{(i)}-\mathbf{m}\right)\left(\mathbf{v}_j^{(i)}-\mathbf{m}\right)^{\top}\right]\\ &= \mathbf{S}_W+\mathbf{S}_B. \end{aligned} \tag{13}$$

In (13), C is the number of classes in a dataset. n_i is the sample number of the *i*th class. $\mathbf{v}_j^{(i)}$ denotes the feature vector of the *j*th sample in the *i*th class. The mean of the feature vectors of the *i*th class is expressed with $\mathbf{m}_i$. The mean of the feature vectors of all classes is expressed with $\mathbf{m}$. In supervised conditions, the separability of a dataset is measured by the ratio of the traces of between-class scatter matrix and within-class scatter matrix. The scatter matrix measure is depicted as $tr(\mathbf{S}_B)/tr(\mathbf{S}_W)$. A large scatter matrix measure value means a large between-class scatter and small within-class scatter.

Matching the idea of analyzing the separability of the dataset in open set conditions, each class is set unknown alternatively. $\mathbf{v}_j^{(i)}$ represents the extracted latent representation. A set of separability matrices can be obtained in each open set setting. DSI in open set conditions is defined as

$$\mathrm{DSI}=\frac{1}{C}\sum_{i=0}^{C-1}\frac{tr(\mathbf{S}_B^{(i)})}{tr(\mathbf{S}_W^{(i)})} \tag{14}$$

where the between-class scatter matrix $\mathbf{S}_B^{(i)}$ and the within-class scatter matrix $\mathbf{S}_W^{(i)}$ are acquired when the *i*th class is set unknown, combining the latent representations of the $C-1$ known classes and one unknown class.

B. SAR Classwise Separability Index

In classification tasks, four indices are defined. They are true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) [38]. The precision, recall rate, and F_β-score are calculated as

$$\begin{aligned}\text{Precision} &= \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}\\ \text{Recall} &= \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}\\ F_\beta &= (1+\beta^2)\frac{\text{Precision}\cdot\text{Recall}}{(\beta^2\cdot\text{Precision}+\text{Recall})}. \end{aligned} \tag{15}$$

Precision represents the proportion of the number of samples that are truly predicted to be positive to the number of samples that are predicted to be positive. Recall rate indicates the ratio of the number of truly predicted positive samples to the number of actual positive samples. The precision can be regarded as a reflection of the impact other classes cause. The recall can be viewed as the uniqueness of a class.

Precision and recall rate are mutually restrictive. When precision is high, recall rate tends to be low. Precision tends to be lower when recall rate is high. Combining the two indices and emphasizing the uniqueness, recall rate is considered to be of more significance than precision. CSI for the *j*th class in open set conditions is defined using F_2 score with $\beta=2$

$$\begin{aligned}\mathrm{CSI}_j &= 0.7\cdot F_{2j(i=j)}^{(i)}+0.3\cdot\frac{1}{C-1}\sum_{i\neq j}F_{2j}^{(i)}\\ &\quad (j=0,1,\ldots,C-1)\end{aligned} \tag{16}$$

in which $F_{2j}^{(i)}$ is the F_2 score of the *j*th class in the condition when the *i*th class is set unknown.

The first term on the right side of (16) indicates the F_2 score of recognizing the *i*th class as unknown when it is set as unknown. The second term indicates the F_2 score of classifying the *j*th class when it is known. With the purpose of emphasizing the performance of correctly recognizing each class as unknown when it is set unknown, the weight proportion of these two terms is set as 7:3.

IV. Datasets

Two SAR image datasets were prepared for analyzing the separability in experiments. The first dataset is the OpenSARUrban [3] acquired from Sentinel-1 satellite [39] for urban interpretation. The second dataset is a ship dataset with images acquired from TerraSAR-X satellite [40]. Besides, as a completely different dataset from the two SAR datasets, the MNIST dataset [41] is utilized for comparative experiments helping validating unknown discovery and demonstrating the separability. Details of these datasets are introduced in the following sections.

A. OpenSARUrban Dataset

The OpenSARUrban [3] dataset was collected from 19 Sentinel-1 [39] images with a relatively low resolution of about 20 m. The patch size is 100×100. In the dataset, there are 5

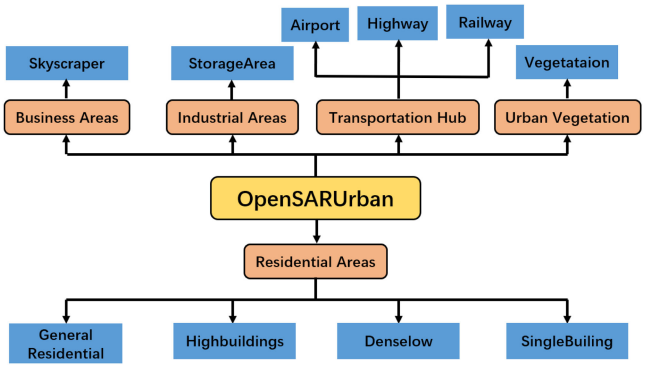


Fig. 2. Structure of the OpenSARUrban dataset [3], including 5 functional classes and totally 10 urban classes.

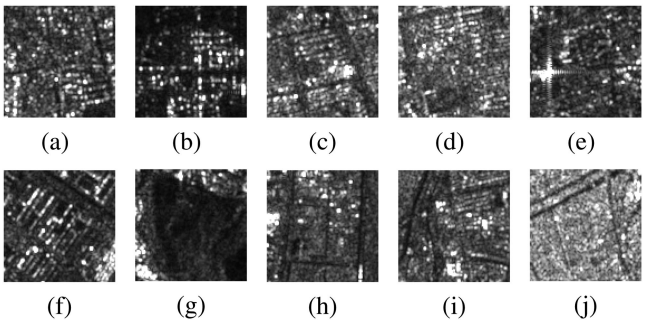


Fig. 3. Patches in the OpenSARUrban dataset [3]. (a) Denselow. (b) General Residential. (c) Highbuildings. (d) SingleBuilding. (e) Skyscraper. (f) StorageArea. (g) Vegetation. (h) Airport. (i) Railway. (j) Highway.

functional classes and a total of 10 different urban classes which cover the areas of 21 major cities in China. The structure of the OpenSARUrban dataset is shown in Fig. 2. The 5 functional classes include business areas, industrial areas, transportation hub, urban vegetation, and residential areas. The total 10 classes are skyscraper, storage area, airport, highway, railway, vegetation (Veg), general residential (Gen.Res), highbuildings (Highbuil), denselow, and single building (SingleBuil).

In our experiments, images of VV polarization are analyzed. Patches of each class are shown in Fig. 3. The difference between various classes of images is not obvious. In addition, the imbalanced distribution of data volume of each class is expressed in Fig. 4.

The data volumes in classes of highbuildings, storage area, and general residential take the upper hand. Followed by denselow, single building, skyscraper, and vegetation. By comparison, the airport, railway, and highway classes show their weakness in data volume.

B. TerraSAR-X Ship Dataset

The ship dataset was acquired from TerraSAR-X satellite [40], it covers 8 ports with patch size ranging from 5×16 pixels to 237×225 pixels. There are four classes of ships in the dataset, including cargo, container, hooker, and tanker. Patches of each class are shown in Fig. 5. The patches of each class are all composed of extremely strong scatter points and low backscatter.

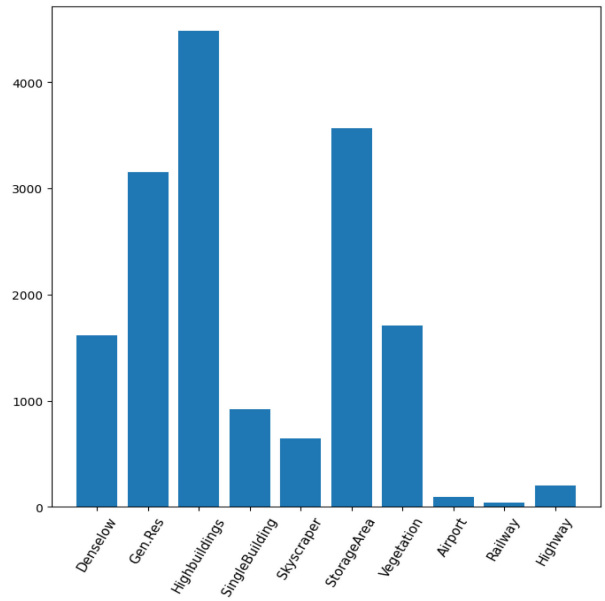


Fig. 4. OpenSARUrban dataset [3] distribution among classes. The amount of data in highbuildings, storage area, and general residential are the most. Denselow, single building, skyscraper, and vegetation have a medium data volume. The data volume of airport, railway, and highway are the least.

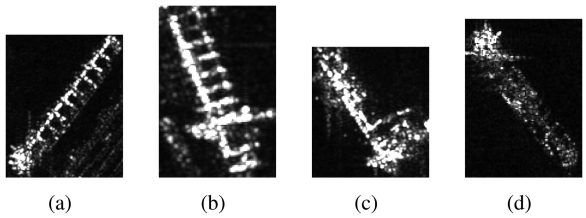


Fig. 5. Images of each class in the TerraSAR-X ship dataset. (a) Cargo. (b) Container. (c) Hooker. (d) Tanker.

TABLE I
NUMBER OF EACH CLASS IN THE TERRASAR-X SHIP DATASET

Classes	Cargo	Container	Hooker	Tanker
Number	2,281	257	3	307

Besides, the ship dataset is much more imbalanced than the OpenSARUrban dataset, numbers of each class are shown in Table I.

In terms of the data volume, cargo is the largest class. Followed by tanker and container. Hooker is the least class.

C. MNIST Dataset

The MNIST dataset [41] is a large dataset of handwritten digits. The images are normalized to fit into a 28×28 pixel bounding box. MNIST is absolutely different from the other two SAR image datasets. Images of each digit class in MNIST are shown in Fig. 6.

Images in MNIST dataset are much simpler and more recognizable than images in the other SAR datasets.

TABLE II
F_2 SCORES (%) AND MEAN RESULTS (%) OF THE OPENSARURBAN DATASET IN THE SUPERVISED SETTINGS

Classes	Unknown Classes										All Known
	Denselow	Gen.Res	Highbuil	SingleBuil	Skyscraper	StorageArea	Veg	Airport	Railway	Highway	
Denselow	—	99.82	99.86	**99.91**	99.29	99.86	99.38	99.80	97.96	99.06	98.51
Gen.Res	99.88	—	**99.96**	99.92	99.39	99.88	99.17	98.86	98.09	99.68	99.37
Highbuil	99.82	99.89	—	**99.95**	99.71	99.91	99.06	99.84	99.27	99.73	99.75
SingleBuil	99.25	99.62	99.97	—	99.03	99.69	99.56	99.04	93.13	**99.91**	99.28
Skyscraper	99.65	**100.0**	**100.0**	99.55	—	99.82	99.56	99.11	98.67	95.13	99.55
StorageArea	99.82	99.42	**99.92**	99.86	99.88	—	99.55	99.90	97.59	99.45	99.28
Veg	**99.95**	99.78	99.72	99.92	99.28	**99.95**	—	99.53	94.84	99.83	99.01
Airport	**100.0**	**100.0**	96.97	97.35	**100.0**	99.69	93.75	—	87.31	98.46	95.02
Railway	**100.0**	**100.0**	88.23	87.59	97.12	**100.0**	97.12	97.12	—	98.54	**100.0**
Highway	**100.0**	99.42	95.82	98.70	**100.0**	99.84	99.57	99.42	95.54	—	98.56
m-Precision	**99.82**	**99.82**	98.59	98.59	99.61	99.74	99.20	99.36	97.32	99.49	99.40
m-Recall	**99.82**	99.76	97.70	97.99	99.23	99.74	98.38	99.14	95.54	98.52	98.70
m-F2 score	**99.82**	99.77	97.83	98.08	99.30	99.74	98.53	99.18	95.82	98.86	98.83

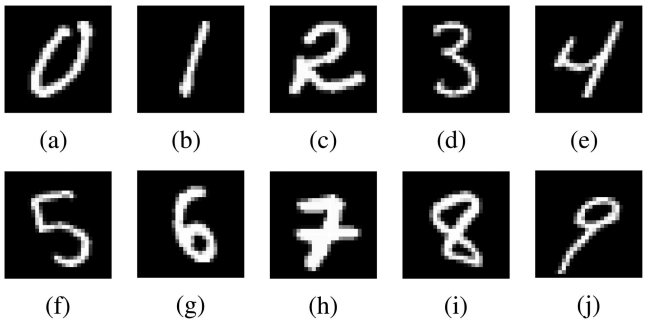


Fig. 6. Images of each digit class in MNIST dataset [41]. (a) Number 0. (b) Number 1. (c) Number 2. (d) Number 3. (e) Number 4. (f) Number 5. (g) Number 6. (h) Number 7. (i) Number 8. (j) Number 9.

V. EXPERIMENTS AND RESULT ANALYSIS

In this section, we conducted experiments in three different settings, including supervised settings, open set settings with in-dataset class unknown, and open set settings with MNIST dataset unknown. In the two open set settings, one of the classes in the dataset was selected in turn not to participate in the training process of the model. In supervised settings, except for the same design as in the two open set settings, another supervised experiment was carried out on the whole dataset with all classes seen in the training process.

A. Supervised Settings on the OpenSARUrban Dataset

In each supervised setting, the images of known classes are split into a training set and testing set with the proportion of 7:3. SAR-SA is utilized for supervised classification following the strategy of testing in supervised settings. The performance is measured by F_2 score. Besides, mean results including mean precision (m-Precision), mean recall (m-Recall), and mean F_2 score (m-F2score) are presented in addition in Table II.

In each row of Table II, the highest scores of a class in all cases are indicated in bold. As can be seen from the results, most of the F_2 scores for each class are greater than 97.00%. The mean F_2 score in each case is no less than 95.82%. For convenience, the ground truth and the classification results of the latent representations in the setting of "all known" are visualized by using t-SNE [42] shown in Fig. 7. Little difference between the ground truth and the classification results can be found from the visualization. The latent representations extracted from each known class can cluster together. The latent representation clusters of different classes are clearly separated.

The experimental results show that under supervised conditions, the degree of intraclass aggregation of the OpenSARUrban is very high, and the difference between the classes is obvious. Therefore, by modeling each SAR class in the OpenSARUrban dataset as a multivariate Gaussian distribution in latent space, the confusion of the data is very low. SAR-SA owns excellent classification capability under supervised conditions.

B. Open Set Settings With In-Dataset Class Unknown on the OpenSARUrban Dataset

In this setting, the OpenSARUrban dataset is split into known classes and unknown classes. Each class in the OpenSARUrban dataset is set unknown in turn. The remaining 9 known classes are split into a training set and testing set with a proportion of 7:3. SAR-SA is utilized for known class classification and unknown new class discovery for the first.

In order to balance the performance of known class classification and unknown class recognition, an ablation study on the value setting of λ in (11) and the probability threshold of (12) was carried out. The results in open set settings with in-dataset class unknown of "Denselow unknown," "Skyscraper unknown," and "Vegetation unknown" are shown in Fig. 8. According to the results, in the case of ensuring that the classification performance of the known classes is not too low, we choose the setting with a higher recognition performance for the unknown class. Thus we set $\lambda = 2$ in (11) and the probability threshold of (12) as 0.5 for tradeoff.

Second, the proposed SAR DSI and CSI are calculated for quantifying the separability of the OpenSARUrban in open set conditions. Table III shows the results, including F_2 scores, m-Precision, m-Recall, m-F2score, and the values of scatter matrix measure $tr(S_B)/tr(S_w)$.

In Table III, bold values on the diagonal line are the F_2 scores of recognizing the unknown class correctly as unknown. The remaining values in each column are the F_2 scores of classifying known classes. In order to make it easier to compare with the results in the supervised setting of "all known" in Table II, the supervised results are listed in the rightmost column of Table III.

1) Known Class Classification: As shown in Table III, comparing the results of known classes in each open set setting with

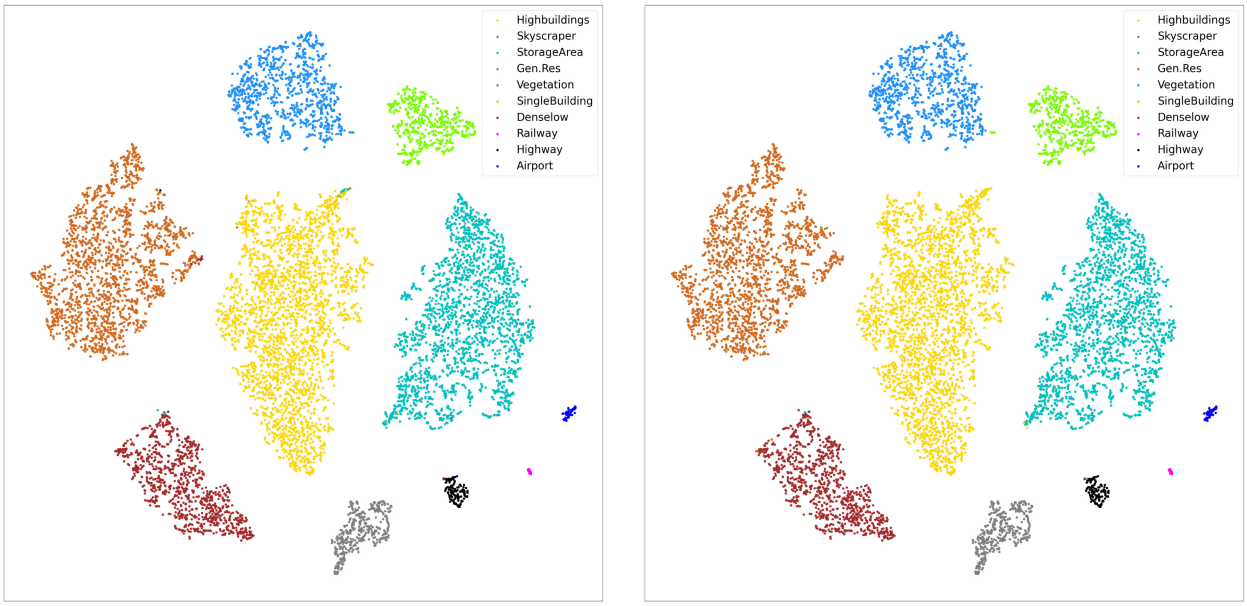


Fig. 7. Ground truth and the classification results of the latent representations of the OpenSARUrban dataset in "all known" supervised setting. Little difference between (a) and (b) exists, latent representations of each class cluster together, clusters of different classes separate from each other clearly. (a) "All Known" Ground Truth. (b) "All Known" Classification Result.

TABLE III
F_2 SCORES (%), MEAN RESULTS (%) AND SCATTER MEASURE RESULTS OF THE OPENSARURBAN DATASET IN OPEN SET SETTINGS WITH IN-DATASET CLASS UNKNOWN AND "ALL KNOWN" SETTING

Classes	Unknown Classes										All Known
	Denselow	Gen.Res	Highbuil	SingleBuil	Skyscraper	StorageArea	Veg	Airport	Railway	Highway	
Denselow	**6.20**	82.03	90.08	96.08	86.85	89.48	91.90	93.75	52.42	80.13	**98.51**
Gen.Res	93.20	**43.37**	87.71	84.93	91.35	67.31	91.03	86.13	86.57	80.15	**99.37**
Highbuil	93.13	92.78	**2.64**	93.00	87.82	71.29	91.48	94.52	89.36	84.90	**99.75**
SingleBuil	92.48	86.99	89.02	**5.90**	92.46	82.27	86.89	81.31	93.79	86.86	**99.28**
Skyscraper	89.71	92.16	74.62	61.25	**24.16**	80.03	76.78	86.94	72.79	89.35	**99.55**
StorageArea	89.83	88.73	82.56	94.72	91.63	**30.24**	88.65	89.33	86.69	61.79	**99.28**
Veg	79.97	76.49	68.24	81.53	79.86	70.31	**44.35**	76.32	74.39	77.99	**99.01**
Airport	76.30	62.29	65.58	77.40	76.55	72.73	59.29	**17.32**	72.13	70.95	**95.02**
Railway	88.23	12.93	94.21	84.56	78.95	87.59	94.21	100.0	**5.15**	93.53	**100.0**
Highway	88.32	88.57	79.60	87.89	90.18	82.97	83.21	88.02	86.92	**17.59**	**98.56**
m-Precision	82.10	81.47	67.62	83.82	88.75	72.21	95.27	90.13	89.77	89.42	**99.40**
m-Recall	79.46	72.50	76.76	75.48	78.67	73.95	85.53	84.91	74.91	76.36	**98.70**
m-F2 score	79.74	72.73	73.42	76.72	79.98	73.42	86.35	81.36	72.02	74.32	**98.83**
$tr(S_B)/tr(S_w)$	7.21	4.08	2.68	11.62	14.16	2.99	28.73	37.81	27.36	23.07	**25.70**

the results in the "all known" setting, the F_2 scores of known classes all decreased. For example, when the single building is set unknown, the F_2 score of classifying denselow slightly decreased from 98.51 to 96.08%. However, when general residential is set unknown, the F_2 score of classifying railway significantly dropped from 100.0 to 12.93%. The confusion matrices in the "all known" setting and "general residential unknown" setting are visualized in Fig. 9.

It can be seen from the confusion matrices that when general residential is set unknown, compared with the results in the "all known" setting, several samples of each known class are wrongly recognized as unknown. In terms of the airport, though the amount of falsely classified samples is less than others, the number of samples of airport is much less than other classes, resulting in the low F_2 score.

Observing from the reduction of the F_2 scores of the known classes in each setting, the impact of introducing testing samples belonging to an unknown class is revealed. When there exist testing samples of unknown classes, SAR-SA will calibrate the results according to the probabilities locating in each known class. If the known class has a large within-class scatter in the setting, the distribution formed by using training samples is of low generalization. The latent representations of testing samples belonging to the class tend to locate in its whole probability space nearly uniformly. Many samples locate in its distribution with probabilities lower than the threshold, thus will be wrongly recognized as unknown. As a result, if the amount of data in a certain class is small, the distribution obtained from this class will be more difficult to have good generalization performance than that of a large class. Thus in open set conditions, classes with low data volume are easily to be confused with other classes.

2) Unknown Class Recognition: Considering the bold values on the diagonal line, which indicate the F_2 scores of recognizing the unknown class correctly as unknown. In comparison with the

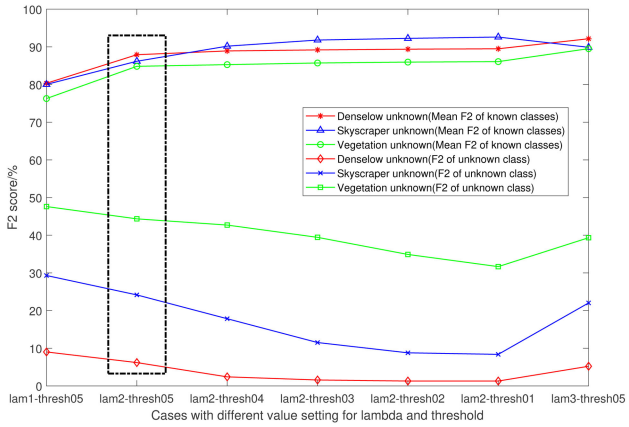


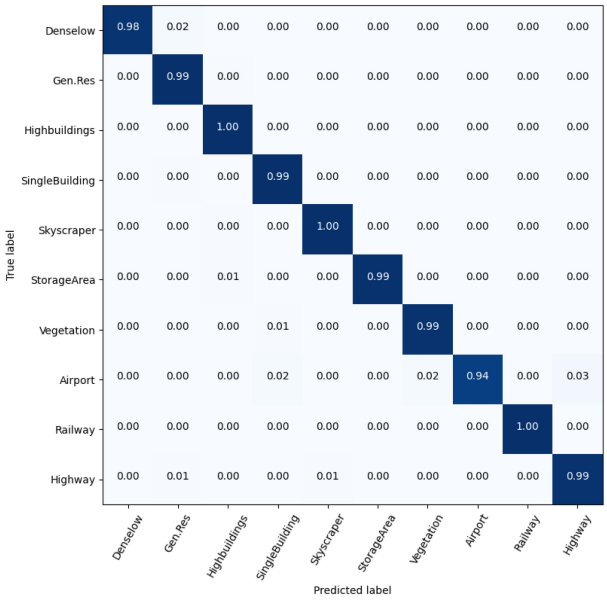
Fig. 8. Ablation study in open set settings with in-dataset class unknown of "Denselow unknown," "Skyscraper unknown," and "Vegetation unknown." On the x-axis, "lam2-thresh05" means setting λ as 2 and the threshold as 0.5, the others are similar.

F_2 scores in the "all known" condition, all results greatly decreased from almost 100.0% to less than 50.00%. The dramatic reduction in F_2 scores shows the great difficulty of recognizing a class as unknown correctly when it is out of supervision. By way of illustration, the ground truth and predicted distribution of the latent representations in cases of setting highbuildings and vegetation as unknown are displayed in Fig. 10.

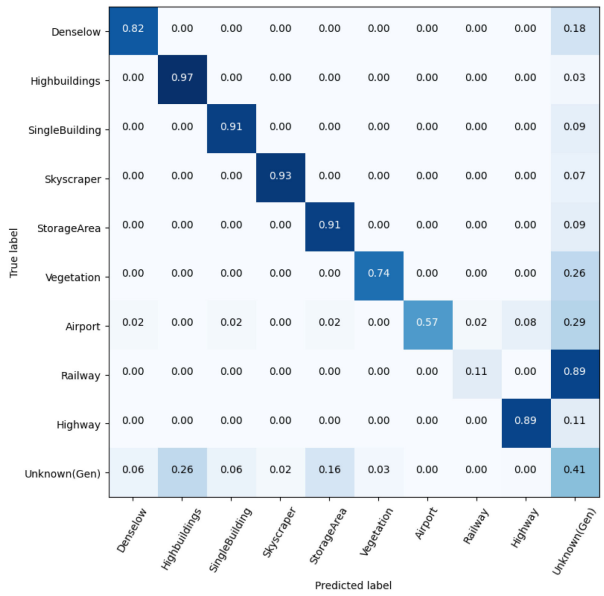
First of all, by observing the pictures in Fig. 10 as a whole, we can find the phenomenon that when a class is set unknown, the latent representations of unknown samples cannot cluster together and are scattered across the distributions of known classes. In the case that vegetation is set unknown, the F_2 score of recognizing it dropped from 99.01% in the "all known" setting to 44.35%. As shown in Fig. 10(c) and (d), the latent representations locating deeply in the clusters of known classes are all wrongly classified as belonging to known classes. The reason is that the probabilities of the samples locating deeply in known distributions are higher than the threshold. These samples have a high degree of similarity and overlap with known classes, leading to the false classification. In contrast, the remaining latent representations locating on the edges of known classes can be correctly recognized as unknown.

In the other case that when highbuildings is set unknown, the F_2 score reduced the most, from 99.75 to 2.64%. According to Fig. 4 and Fig. 10(a) and (b), highbuildings is of the greatest amount in the OpenSARUrban dataset. When it is set unknown, almost all samples have a high degree of similarity with known classes. The latent representations of highbuildings distribute in the known clusters nearly uniformly with probabilities higher than the threshold. Thus almost all the unknown samples are wrongly classified as belonging to known classes.

Hence, the conclusion comes that due to the lack of strong supervision signals on unknown classes, the latent representations of unknown samples are scattered across distributions of known classes with high probabilities. Consequently, the overlap between the distributions of the known classes and unknown classes results in the false classification and decrease of F_2 scores. The more samples of unknown class are wrongly classified as known classes, the greater similarity between the unknown classes and other known classes exists.



(a)



(b)

Fig. 9. Confusion matrices in "all known" setting and "General Residential Unknown" setting. (a) All Known. (b) General Residential Unknown.

3) Analyzing the Datasetwise Separability Using DSI: To analyze the separability at the dataset level in open set settings, two aspects should be taken into consideration. The first is the within-class scatter, the second is the between-class scatter. Combining the two aspects together, results are shown in the last row of Table III, ranging from 2.68 to 37.81.

For example, compared with the "all known" setting, the decrease of F_2 scores of each class and scatter matrix measure values when setting general residential and airport unknown are shown in Table IV. The total decrease of F_2 scores when

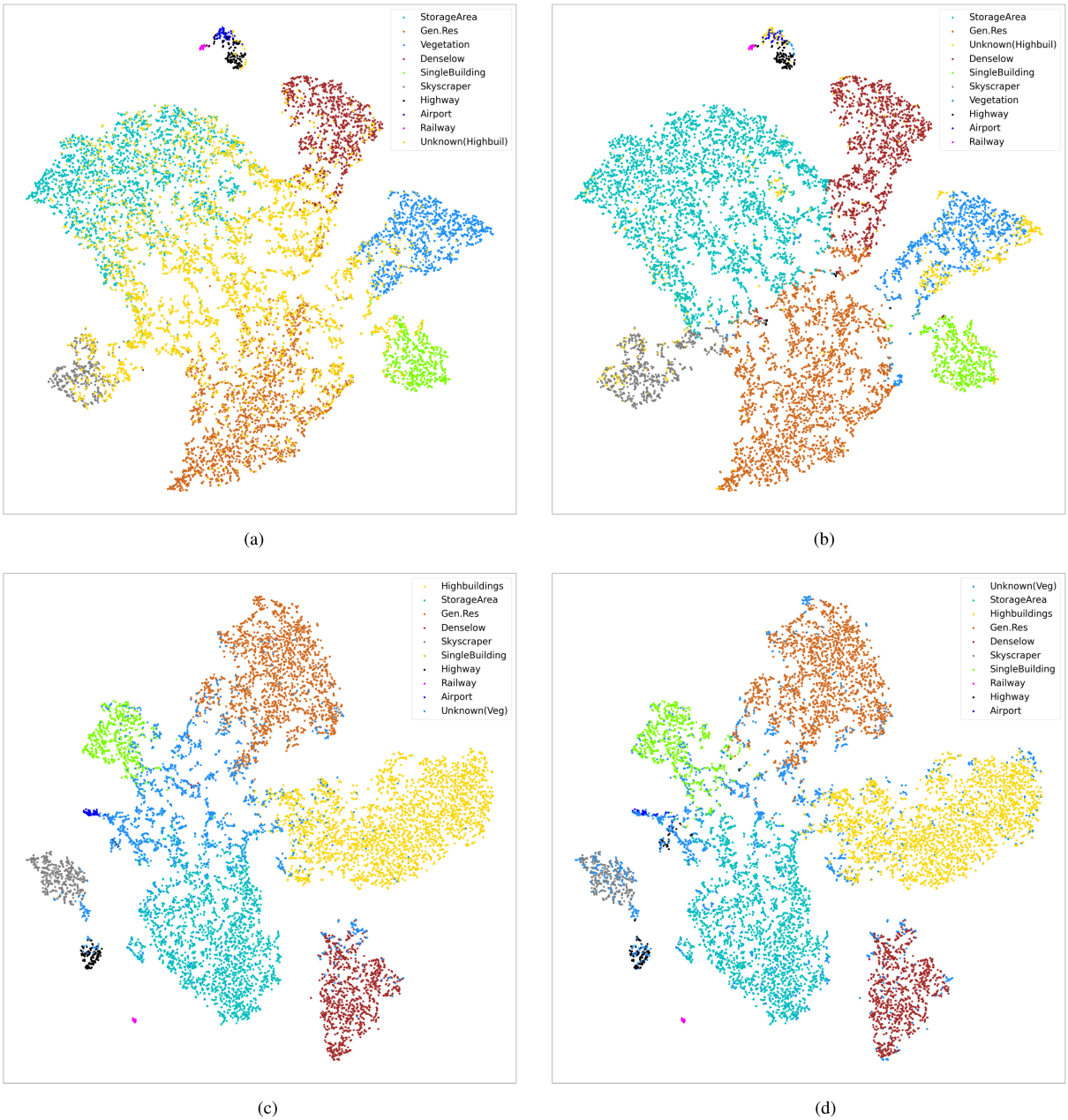


Fig. 10. Ground truth and results of the latent representations of the OpenSARUrban dataset in open set settings. (a) (b) correspond with highbuildings unknown, the unknown recognition F_2 score is 2.64%; (c) (d) correspond with vegetation unknown, the unknown recognition F_2 score is 44.35%. (a) Highbuildings Unknown-Ground Truth. (b) Highbuildings Unknown-Result. (c) Vegetation Unknown-Ground Truth. (d) Vegetation Unknown-Result.

general residential unknown is 262.01%, larger than that when airport unknown with 174.70%. When general residential is set unknown, its impact on the overall classification and recognition performance is much greater than when airport unknown. The scatter matrix measure when general residential unknown is 4.08, less than that when airport unknown with 37.81. It indicates that the higher between-class scatter and lower within-class scatter happens when airport is unknown, in comparison with the case when general residential is unknown.

From the results we can see that the values of scatter matrix measure in different open set settings are different from each other, though on the same dataset. It shows that the degree of aggregation of data within a class and the degree of dispersion of data between different classes vary with the setting of the unknown class.

To quantify the separability at the dataset level in open set settings with in-dataset class unknown, the DSI of the OpenSARUrban dataset is calculated combining all cases using (14)

$$\mathrm{DSI} = \frac{1}{10}\sum_{i=0}^{9}\frac{tr(\mathbf{S}_B^{(i)})}{tr(\mathbf{S}_W^{(i)})} = 13.83. \tag{17}$$

The DSI in the open set condition is 13.83, which is nearly half of the separability in the “All Known” setting with 25.70.

TABLE IV
F_2 SCORE DECREASE (%) AND SCATTER MATRIX MEASURE COMPARISON OF EACH CLASS ON THE OPENSARURBAN DATASET

Classes	Gen.Res Unknown	Airport Unknown
Denselow	16.48	4.76
Gen.Res	56.00	13.24
Highbuildings	6.97	5.23
SingleBuilding	12.29	17.97
Skyscraper	7.39	12.61
StorageArea	10.55	9.95
Vegetataion	22.52	22.69
Airport	32.73	77.70
Railway	87.07	0.00
Highway	9.99	10.54
SUM of Decrease	**262.01**	**174.70**
$tr(S_B)/tr(S_w)$	**4.08**	**37.81**

TABLE V
CSI (%) IN OPEN SET CONDITIONS OF EACH CLASS IN THE OPENSARURBAN DATASET

Classes	CSI	Classes	CSI
Denselow	29.76	**StorageArea**	46.96
Gen.Res	55.97	**Vegetation**	53.88
Highbuildings	28.46	**Airport**	33.23
SingleBuilding	30.53	**Railway**	28.08
Skyscraper	41.03	**Highway**	38.02

It demonstrates that in open set conditions, the between-class scatter is larger and the within-class scatter is lower than in the supervised setting. The confusion among different classes increases greatly. It reveals the serious challenge for the classification with unknown class existing.

4) Analyzing the Classwise Separability Using CSI: According to (16), the CSI of each class in open set conditions is calculated and presented in Table V. It can be seen from the table that general residential and vegetation own the highest CSI. Railway, highbuildings, and denselow are of the lowest CSI.

An intuitive comparison can, first, be discovered from the images themselves in Fig. 3. In Fig. 3, patches of (a) denselow, (c) highbuildings, and (d) single building are all composed of strong scatter points reflected from the surface of the densely arranged buildings. It is even hard for experts to classify them correctly because of the great similarities among them. Thus the unknown recognition F_2 scores and CSI of these three classes are all low. In contrast, lower scatter intensity appears in the patches of Fig. 3(b) general residential and Fig. 3(g) vegetation, which is fully different from other classes with more strong scatter points. They own higher unknown recognition F_2 scores and CSI.

In addition, more detailed analysis can be obtained from Fig. 10. Taking vegetation for instance, when vegetation is known to SAR-SA as shown in Fig. 3(a) and (b), the difference between the ground truth and classification results of vegetation is less than most of the other known classes. The number of samples in vegetation that are wrongly classified as other known classes or recognized as unknown class, and samples in other classes falsely classified as vegetation are lower than those of other known classes. It shows the robust separability of vegetation when there exist unknown classes. When vegetation is set unknown corresponding to Fig. 3(c) and (d), though part of samples belonging to vegetation are classified as other known classes incorrectly, the proportion of the falsely classified samples in vegetation is lower than that of other classes when other classes are unknown, indicating its lower similarity with other known classes.

The same analysis can be applied to other situations. Conclusion is drawn that in open set settings, general residential and vegetation are classes with the highest CSI, followed by storage area and skyscraper. Railway, highbuildings, and denselow own the lowest CSI. In open set conditions, the class with higher CSI is much easier to be classified when it is known and recognized as unknown when it is unknown to the model.

C. Open Set Settings With MNIST Unknown on the OpenSARUrban Dataset

As a completely different dataset from SAR, MNIST is used in this part for further comparison and validation. Huge differences between MNIST and OpenSARUban dataset can be found in Figs. 3 and 6. In each open set setting, the unknown classes are replaced by MNIST. The parameters of the training model in each setting are kept unchanged for a fair comparison. Results are shown in Table VI.

In Table VI, the bold values on the diagonal line are the F_2 scores of recognizing MNIST as unknown class correctly. Comparing with the results in Table III, the unknown recognition performance has been greatly improved. For example, when highbuildings is unknown to SAR-SA, the F_2 score of recognizing it as unknown class is 2.64%. After replacing highbulidings with MNIST, the F_2 score of recognizing MNIST as unknown greatly increased to 90.01%. This also happens in other cases. Another important result worth noting is that the recall rate of recognizing MNIST as unknown is 100.0% in all cases. None of the samples in MNIST is wrongly classified as belonging to other known classes.

Under the settings that known classes without denselow and known classes without airport, the latent representations of ground truth, and predicted results are visualized in Fig. 11. Samples belonging to MNIST are expressed with red points.

It is apparent from Fig. 11 that though MNIST was not part of the training process of SAR-SA and out of supervision, the latent representations extracted from MNIST can still cluster together and separate from other clusters of known classes clearly. As for MNIST itself, there is no difference between the distributions of ground truth in Fig. 11(a) and (c), and classification results in Fig. 11(b) and (d). Besides, in terms of the latent representations of known classes, probabilities of some representations which locate on the edge of the known clusters are lower than the threshold, thus these samples are wrongly recognized as unknown.

Moreover, as for the data separability in each setting, comparing the values of scatter matrix measure in the two tables, most of the values in Table VI are greater than those in Table III. The conclusion is reached that the between-class scatter when MNIST unknown is larger than that when in-dataset class unknown. The within-class scatter when MNIST unknown is less than that when in-dataset class unknown.

This comparative experiment reveals that due to the large difference between MNIST and SAR images, MNIST is easy to be recognized as unknown. In terms of SAR datasets, high degree of similarity and confusion between various classes leads

TABLE VI
F_2 SCORES (%), MEAN RESULTS (%) AND SCATTER MEASURE RESULTS OF THE OPENSARURBAN DATASET IN OPEN SET SETTINGS WITH MNIST UNKNOWN

Classes	Unknown Classes									
	Denselow	Gen.Res	Highbuil	SingleBuil	Skyscraper	StorageArea	Veg	Airport	Railway	Highway
Denselow or MNIST	**89.09**	84.87	97.40	96.92	87.17	95.23	92.78	93.78	52.42	80.18
Gen.Res or MNIST	96.84	**86.70**	96.73	86.77	91.97	70.64	92.44	86.15	86.58	80.18
Highbuil or MNIST	95.65	97.60	**90.01**	93.73	89.06	75.68	92.14	94.53	89.40	84.94
SingleBuil or MNIST	94.42	92.52	91.81	**84.43**	92.52	87.09	92.83	81.31	93.82	87.00
Skyscraper or MNIST	91.26	94.40	88.02	61.49	**82.36**	83.99	77.21	86.94	72.79	89.35
StorageArea or MNIST	92.79	92.40	95.94	96.26	92.49	**74.46**	91.03	89.39	86.73	61.96
Veg or MNIST	82.31	77.72	71.86	85.02	80.14	75.43	**85.33**	76.44	74.39	78.14
Airport or MNIST	76.55	62.29	65.79	78.62	76.55	75.71	62.29	**80.52**	72.13	70.95
Railway or MNIST	88.23	12.93	94.21	84.56	78.95	88.23	94.21	100.0	**73.17**	93.53
Highway or MNIST	88.97	90.51	82.09	87.89	90.31	84.46	83.21	88.15	86.92	**68.75**
m-Precision	95.54	92.68	94.23	94.49	94.39	91.20	95.27	94.49	93.26	92.57
m-Recall	88.90	78.40	86.54	84.85	85.56	80.97	85.53	87.57	78.41	79.58
m-F2 score	89.61	79.19	87.39	85.57	86.15	81.09	86.35	87.72	78.83	79.50
$tr(S_B)/tr(S_w)$	15.80	18.89	8.84	14.06	19.55	13.18	28.73	29.92	11.10	23.53

TABLE VII
F_2 SCORES (%), MEAN RESULTS (%) AND SCATTER MEASURE RESULTS OF THE TERRASAR-X SHIP DATASET IN OPEN SET SETTINGS WITH IN-DATASET CLASS UNKNOWN AND "ALL KNOWN" SETTING

Classes	Unknown Classes				All Known
	Cargo	Container	Hooker	Tanker	
Cargo	**85.07**	56.20	65.17	28.48	**86.19**
Container	18.46	**85.96**	41.94	52.72	**58.42**
Hooker	0.00	44.91	**00.00**	43.38	**41.41**
Tanker	17.23	3.72	52.37	**30.26**	**50.40**
m-Precision	43.92	68.78	31.41	61.91	**70.23**
m-Recall	31.97	47.57	43.39	41.40	**58.94**
m-F2 score	30.19	47.69	39.87	38.71	**59.10**
$tr(S_B)/tr(S_w)$	0.16	0.78	0.67	1.19	**1.28**

to the great difficulty of SAR classification in actual conditions with unknown classes exiting.

D. DSI Analyzing on TerraSAR-X Ship Dataset

Due to the great imbalance of the ship dataset, we performed rotation, mirror, and other transformations on the small classes such as hooker to realize data augmentation. Known classes are split into training and testing set with a proportion of 8:2. Experiments in "all known" setting and open set settings with in-dataset class unknown are carried out in the same way as on the OpenSARUrban dataset. Results are shown in Table VII, bold values on the diagonal line are the F2 scores of recognizing the unknown class correctly as unknown.

In supervised settings, the performance in the "all known" setting is not as good as that on the OpenSARUrban dataset, indicating the larger difficulty in ship classification. Since SAR images only show scatter points, ship targets are all made of hard materials such as metal, they all appear with lots of strong scatter points in SAR images. Therefore, when the color information is lost and the resolution is lower than that of optical images, the difference between different types of ship targets in SAR images is greatly reduced. The appearances of ships from various classes are quite similar to each other as shown in Fig. 5.

In open set settings with in-dataset unknown, the results are similar to those in the OpenSARUrban dataset. When there exist unknown classes, the classification performance of known classes and unknown recognition performance of the unknown classes are almost completely decreased. At the dataset level, regarding the scatter matrix measure values in the bottom row, the values in open set settings are mostly less than that in the "all known" setting. It shows that the between-class scatter in open set settings is less than that in "all known" setting. The within-class scatter in open set settings is larger than that in "all known" setting. Hence, the greater degree of confusion of the whole dataset happens in open set conditions than in supervised conditions.

TABLE VIII
DSI COMPARISON IN THE OPEN SET CONDITION AND SUPERVISED DATA SEPARABILITY

Dataset	OpenSARUrban	TerraSAR-X Ship
Open Set DSI	13.83	0.70
Supervised Separability	25.70	1.28

The DSI in open set conditions of the ship dataset is obtained combining the scatter matrix measure results in all cases using (14)

$$\mathrm{DSI} = \frac{1}{4}\sum_{i=0}^{3}\frac{tr(\mathbf{S}_B^{(i)})}{tr(\mathbf{S}_W^{(i)})} = 0.70. \tag{18}$$

The DSI in open set setting is nearly half of the separability in the "all known" setting with 1.28.

For comparison, the DSI of the OpenSARUrban dataset and TerraSAR-X ship dataset in open set conditions are shown in Table VIII, together with the data separability in the "all known" condition.

As can be seen from Table VIII, both the DSI in open set conditions of the two SAR image datasets are nearly half of the supervised data separability. The comparison demonstrates that when there exist unknown classes, the separability of SAR datasets greatly reduced, the confusion among classes greatly increased. It is quite difficult to realize the accurate known class classification and unknown class recognition.

Moreover, the supervised data separability and open set DSI of TerraSAR-X dataset are far less than those of the OpenSARUrban dataset. The greater challenge of the classification under the practical conditions on the TerraSAR-X ship dataset is revealed, in comparison with the OpenSARUrban dataset.

E. CSI Analyzing on TerraSAR-X Ship Dataset

In addition, according to (16), combining the F_2 scores in all cases, the CSI of each class in the TerraSAR-X ship dataset in

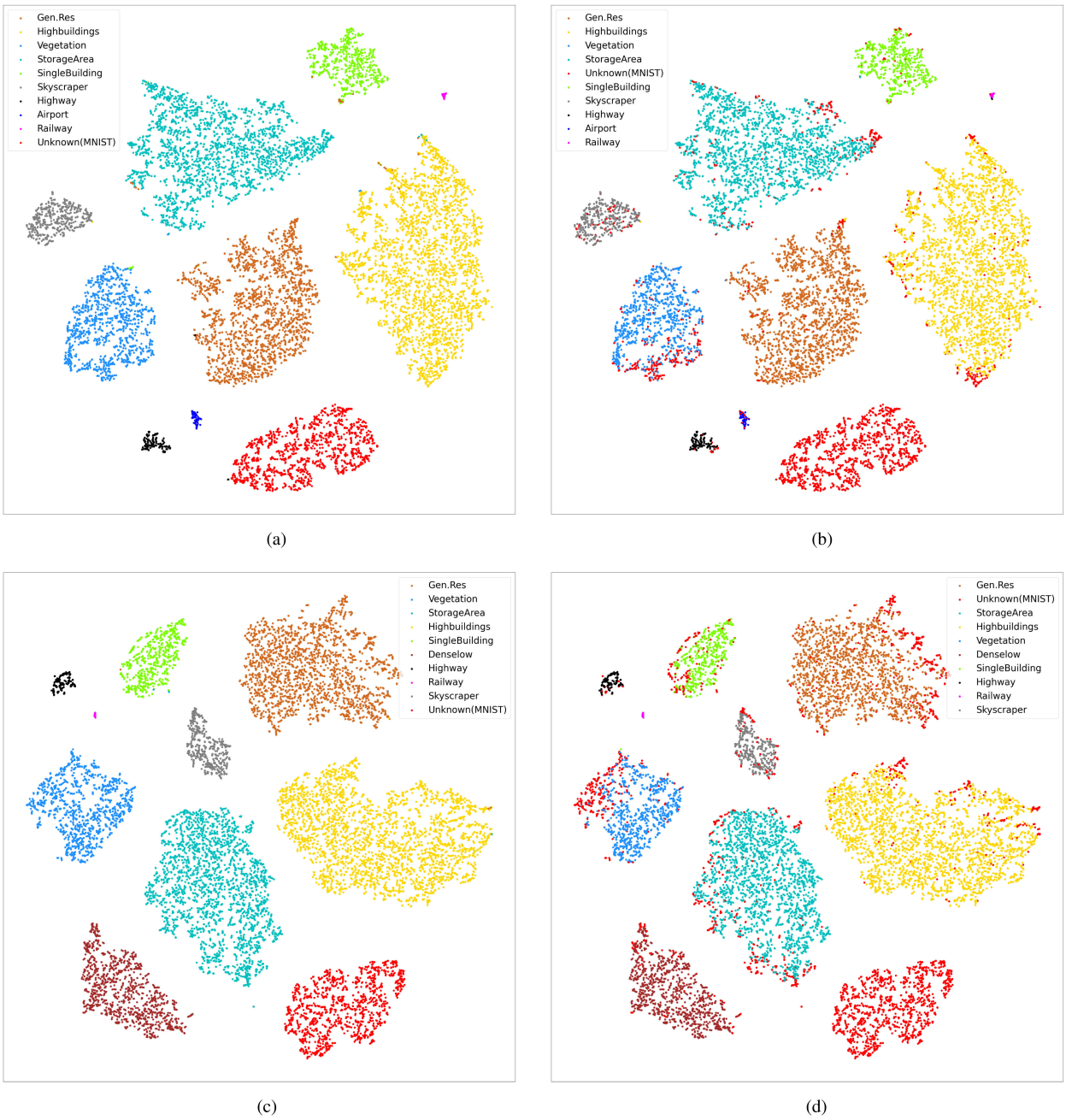


Fig. 11. Ground truth and results of the latent representations in open set settings with MNIST unknown. (a) and (b) Known classes without denselow, the unknown recognition F_2 score is 89.09%. (c) and (d) Known classes without airport, the unknown recognition F_2 score is 80.52%. (a) Without Denselow-Ground Truth. (b) Without Denselow-Result. (c) Without Airport-Ground Truth. (d) Without Airport-Result.

TABLE IX
CSI (%) IN OPEN SET CONDITIONS OF EACH CLASS IN THE TERRASAR-X SHIP DATASET

Classes	Cargo	Container	Hooker	Tanker
CSI	74.53	71.48	8.83	28.51

open set settings is presented in Table IX. CSI of cargo is the highest, hooker has the lowest CSI.

The confusion matrices in the setting when cargo and hooker unknown are delivered in Fig. 12. When cargo is unknown, though hooker is in supervision, samples of it are all wrongly recognized as unknown. While most of samples in cargo can still be recognized as unknown. It shows the large within-class scatter of hooker and low similarity between cargo and other known classes. When hooker is unknown, two-thirds of samples in cargo can be classified correctly, but none of the samples of hooker is recognized as unknown. It reveals the high degree of aggregation within cargo and the great similarity between hooker and other known classes. The results demonstrate that in open set conditions, hooker can easily be confused with other classes no matter whether it is known. The separability of hooker in

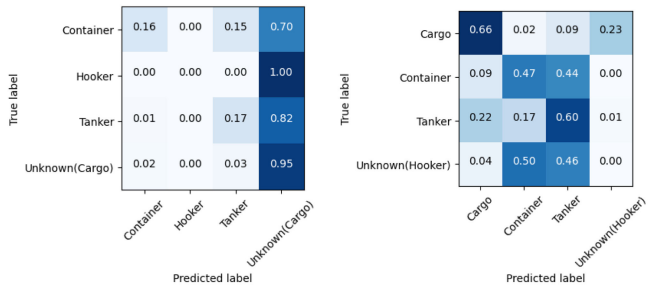


Fig. 12. Confusion matrices when cargo and hooker unknown. (a) Cargo Unknown. (b) Hooker Unknown.

open set conditions is low. In contrast, cargo is much easier to be classified when it is known and recognized as unknown when it is unknown.

The same analysis can be applied to other classes. Thus in open set conditions, cargo and container are classes with the highest CSI, followed by container. Hooker owns the lowest CSI. The class with lower CSI is more intractable to be classified correctly when it is known, and more difficult to be recognized as unknown when it is unknown.

VI. Conclusion

This article demonstrated the SAR image classification difficulty in the condition of facing unknown classes. SAR-SA is proposed first to model each known class as a multivariate Gaussian distribution. SAR-SA can not only classify the known classes seen in the training phase but also can recognize samples locating in each known distribution with probabilities lower than the threshold as unknown. DSI and CSI are defined to quantify the separability of a dataset at the dataset level and class level, respectively. The OpenSARUrban dataset and the TerraSAR-X ship dataset are prepared for experiments in supervised settings, open set settings with in-dataset class unknown and open set settings with MNIST unknown. According to the results, we find that DSI in open set conditions is nearly half of that in the supervised condition, showing the greater confusion of a dataset in open set conditions compared with in supervised conditions. The DSI of the OpenSARUrban dataset is much larger than that of TerraSAR-X ship dataset, showing that the OpenSARUrban is easier to achieve accurate classification and recognition in open set conditions than the TerraSAR-X ship dataset. At the class level, even though the SAR image classes are semantically different from each other. If a class is set unknown and out of supervision, the latent representations of unknown classes are scattered across the distribution of other known classes, resulting in the overlap with supervised known classes. Thus the classification in open set settings is more difficult. Besides, classes with low CSI are more intractable to be classified when known and recognized as unknown when unknown correctly. DSI and CSI are effective indicators to quantify the difficulty of practical classification. This method is not limited to SAR data only, it could be extended to any other kind of data.

Acknowledgment

The TerraSAR-X images used in this study were provided via ESA-MOST CHINA Dragon 5 programme ID.58190. The authors would like to thank several colleagues in the Shanghai Key Laboratory of Intelligent Sensing and Recognition for their devoted assistance on data annotation. The authors sincerely appreciate the reviewers for their insightful and valuable comments.

References

[1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 1, pp. 6–43, Mar. 2013.

[2] 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, May 2018.

[3] J. Zhao, Z. Zhang, W. Yao, M. Datcu, H. Xiong, and W. Yu, "OpenSARUrban: A Sentinel-1 SAR image dataset for urban interpretation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 187–203, 2020.

[4] T. D. Ross, S. W. Worrell, V. J. Velten, J. C. Mossing, and M. L. Bryant, "Standard SAR ATR evaluation experiments using the MSTAR public release data set," in *Proc. SPIE*, vol. 3370, pp. 566–573, 1998.

[5] L. Huang *et al.*, "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, Jan. 2018.

[6] J. Singh and M. Datcu, "SAR image categorization with log cumulants of the fractional Fourier transform coefficients," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 12, pp. 5273–5282, Dec. 2013.

[7] X. Hou, W. Ao, Q. Song, J. Lai, H. Wang, and F. Xu, "FUSAR-Ship: building a high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition," *Sci. China Inf. Sci.*, vol. 63, no. 4, 2020, Art. no. 140303.

[8] Z. Huang, M. Datcu, Z. Pan, and B. Lei, "Deep SAR-Net: Learning objects from signals," *ISPRS J. Photogram. Remote Sens.*, vol. 161, pp. 179–193, 2020.

[9] B. Li, B. Liu, L. Huang, W. Guo, Z. Zhang, and W. Yu, "OpenSARShip 2.0: A large-volume dataset for deeper interpretation of ship targets in Sentinel-1 imagery," 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA), 2017, pp. 1–5, doi: 10.1109/BIGSARDATA.2017.8124929.

[10] D. Marmanis, M. Datcu, T. Esch, and U. Stilla, "Deep learning earth observation classification using ImageNet pretrained networks," *IEEE Geosci. Remote Sens. Lett*, vol. 13, no. 1, pp. 105–109, Jan. 2016.

[11] X. Leng, K. Ji, S. Zhou, X. Xing, and H. Zou, "Discriminating ship from radio frequency interference based on noncircularity and non-Gaussianity in Sentinel-1 SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 352–363, Jan. 2019.

[12] W. Liang, Y. Wu, M. Li, and Y. Cao, "High-resolution SAR image classification using context-aware encoder network and hybrid conditional random field model," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5317–5335, Aug. 2020.

[13] G. Dong, H. Liu, and J. Chanussot, "Keypoint-based local descriptors for target recognition in SAR images: A comparative analysis," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 1, pp. 139–166, Mar. 2021.

[14] S. Cui, G. Schwarz, and M. Datcu, "Remote sensing image classification: No features, no clustering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 11, pp. 5158–5170, Nov. 2015.

[15] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, and F. Chen, "High-resolution SAR image classification via deep convolutional autoencoders," *IEEE Geosci. Remote Sens. Lett*, vol. 12, no. 11, pp. 2351–2355, Nov. 2015.

[16] Z. Huang, Z. Pan, and B. Lei, "Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data," *Remote Sens.*, vol. 9, no. 9, 2017, Art. no. 907.

[17] J. He, Y. Wang, and H. Liu, "Ship classification in medium-resolution SAR images via densely connected triplet CNNs integrating fisher discrimination regularized metric learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3022–3039, Apr. 2021.

[18] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, and T. E. Boult, "Toward open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1757–1772, Jul. 2013.

[19] R. Kassab and F. Alexandre, "Incremental data-driven learning of a novelty detection model for one-class classification with application to high-dimensional noisy data," *Mach. Learn.*, vol. 74, no. 2, pp. 191–234, 2009.
[20] H. Hoffmann, "Kernel PCA for novelty detection," *Pattern Recognit.*, vol. 40, no. 3, pp. 863–874, 2007.
[21] H. -M. Yang, X. -Y. Zhang, F. Yin, Q. Yang, and C. -L. Liu, "Convolutional prototype network for open set recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, doi: 10.1109/TPAMI.2020.3045079.
[22] D. A. Clifton, S. Hugueny, and L. Tarassenko, "Novelty detection with multivariate extreme value statistics," *J. Signal Process. Syst.*, vol. 65, no. 3, pp. 371–389, 2011.
[23] S. Ntalampiras, I. Potamitis, and N. Fakotakis, "Probabilistic novelty detection for acoustic surveillance under real-world conditions," *IEEE Trans. Multimedia.*, vol. 13, no. 4, pp. 713–719, Aug. 2011.
[24] A. Ghoting, S. Parthasarathy, and M. E. Otey, "Fast mining of distance-based outliers in high-dimensional datasets," *Data Mining Knowl. Discov.*, vol. 16, no. 3, pp. 349–364, 2008.
[25] M. K. Albertini and R. F. de Mello, "A self-organizing neural network for detecting novelties," in *Proc. ACM Symp. Appl. Comput.*, 2007, pp. 462–466.
[26] R. K. Baghbaderani, Y. Qu, H. Qi, and C. Stutts, "Representative-discriminative learning for open-set land cover classification of satellite imagery," in *Proc. 16th Eur. Conf. Comput. Vis.*, Springer, 2020, pp. 1–17.
[27] S. Dang, Z. Cao, Z. Cui, Y. Pi, and N. Liu, "Open set incremental learning for automatic target recognition," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4445–4456, Jul. 2019.
[28] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification and Scene Analysis*, vol. 3. New York, NY, USA: Wiley, 1973.
[29] L. Wang, "Feature selection with Kernel class separability," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 9, pp. 1534–1546, Sep. 2008.
[30] C. Thornton, *Truth From Trash: How Learning Makes Sense*. Cambridge, MA, USA: MIT Press, 2002.
[31] L. Mthembu and T. Marwala, "A note on the separability index," 2008, *arXiv:0812.1107*.
[32] N. Mthembu and J. Greene, "A comparison of three class separability measures," in *Proc. 15th Annu. Sympo. Pattern Recogni. Assoc.*, 2004, pp. 63–67.
[33] X. Sun, Z. Yang, C. Zhang, K. V. Ling, and G. Peng, "Conditional Gaussian distribution learning for open set recognition," in *Proc. Conf. Comput. Vision Pattern Recognit.*, 2020, pp. 13 477–13486.
[34] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2013, *arXiv:1312.6114*.
[35] M. D. Hoffman and M. J. Johnson, "ELBO surgery: Yet another way to carve up the variational evidence lower bound," in *Proc. Workshop Adv. Approx. Bayesian Inference*, 2016, vol. 1, p. 2.
[36] C. K. Sønderby *et al.*, "How to train deep variational autoencoders and probabilistic ladder networks[J]," 2016, *arXiv:1602.02282*.
[37] Y. Sasaki *et al.*, "The truth of the F-measure," *Teach Tutor Mater.* 2007.
[38] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.
[39] R. Torres *et al.*, "GMES Sentinel-1 mission," *Remote Sens. Environ.*, vol. 120, pp. 9–24, 2012.
[40] W. Pitz and D. Miller, "The TerraSAR-X satellite," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 615–622, Feb. 2010.
[41] Y. LeCun, "The MNIST database of handwritten digits," 1998. [Online]. Available: http://yann. lecun. com/exdb/mnist/
[42] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Technol.*, vol. 9, no. 11, 2008, pp. 2579–2605 .

Ning Liao (Student Member, IEEE) received the B.Sc. degree in information countermeasure technology from Xidian University, Xi'an, China, in 2019, and was awarded an outstanding graduate. He is currently working toward the Ph.D. degree in information and communication engineering with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China.

His research interests include remote sensing, synthetic aperture radar (SAR) image interpretation, open set recognition, and computer vision.

Mihai Datcu (Fellow, IEEE) 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, and the title habilitation a diriger des recherches in computer science from the University Louis Pasteur, Strasbourg, France, in 1999.

From 1992 to 2002, he had a longer Invited Professor assignment with the Swiss Federal Institute of Technology, ETH Zurich, Zürich, Switzerland. He has held Visiting Professor positions with the University of Oviedo, Oviedo, Spain; the University Louis Pasteur, Strasbourg, France; International Space University, Strasbourg, France; University of Siegen, Siegen, Germany; University of Innsbruck, Innsbruck, Austria; University of Alcala, Alcala de Henares, Spain; University Tor Vergata, Rome, Italy; Universidad Pontificia de Salamanca, campus de Madrid, Madrid, Spain; University of Camerino, 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 also been a Scientist with German Aerospace Center, Deutsches Zentrum für Luft- und Raumfahrt (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 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 with 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 has authored more than 450 scientific publications, which include 80 journal papers, and a book on number theory. 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 is a Member of the ESA Working Group and Big Data from Space (BiDS). He was the recipient of the IEEE Geoscience and Remote Sensing Society Prize Best Paper Award in 2006, the National Order of Merit with the rank of Knight, for outstanding international research results, awarded by the President of Romania, in 2008, and the Romanian Academy Prize Traian Vuia, for the development of the SAADI image analysis system and his activity in image processing, in 1987. He was a Guest Editor for a Special Issue on IIM of the IEEE and other journals. He was a Co-Organizer of international conferences and workshops.

Zenghui Zhang (Senior Member, IEEE) 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 research interests include radar signal processing, joint radar and communication system, synthetic aperture radar image interpretation, deep learning techniques, target detection and recognition.

Weiwei Guo received the B.Sc., M.Sc., and Ph.D. degrees in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2005, 2007, and 2014, respectively.

He has been a visiting Ph.D. student in the area of computer vision with School of Electronic Engineering and Computer Science, Queen Mary University of London, U.K., between 2008 and 2010, and a Postdoctoral Researcher with Shanghai Jiao Tong University, Shanghai, China, in the area of remote sensing image interpretation. He is an Assistant Professor with the Center of Digital Innovation, Tongji University, Shanghai, China. His research interests include pattern recognition and machine learning and their applications in the fields of computer vision, remote sensing image understanding, HCI.

Juanping Zhao received the B.Sc. degree in electrical engineering from Xidian University, Xian, China, in 2014, and the Ph.D. degree in information and communication engineering from the School of Electronic Information and Electrical Engineering from Shanghai Jiao Tong University, Shanghai, China, in 2020.

From 2018 to 2019, she was a Joint Ph.D. Student with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Oberpfaffenhofen, Germany. Her research interests include SAR and PolSAR image interpretation, remote sensing image processing, pattern recognition, and machine learning.

Wenxian Yu (Senior Member, IEEE) 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 the Shanghai Jiao Tong University, Shanghai, China, 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. He has authored or coauthored more than 200 research papers in the areas of his research. His research interests include radar target recognition, remote sensing information processing, multisensor data fusion, and integrated navigation system.