

SARPointNet: An Automated Feature Learning Framework for Spaceborne SAR Image Registration

Xin Li [✉], Taoyang Wang [✉], Hao Cui [✉], Member, IEEE, Guo Zhang [✉], Qian Cheng [✉], Tiancheng Dong, and Boyang Jiang

Abstract—Accurate registration between synthetic aperture radar (SAR) images is the basis for high-precision geometric correction of SAR images. The feature points extracted by conventional feature extraction methods are unsatisfactory, which are affected by imaging geometric characteristics and speckle noise of SAR images. This article innovatively proposes a spaceborne SAR image feature learning framework to realize automatic sample generation and model training. It mainly includes two modules: The feature sample generation module based on the initial geometric information of spaceborne SAR. The initial rational polynomial coefficient (RPC) parameters of the spaceborne SAR are adopted to realize the initial positioning of the SAR image, and a variety of feature extraction operators are used to match the overlapping areas to obtain high-precision matching points, which are employed as training samples for image pairs; pseudo-Siamese feature learning network SARPointNet for SAR image feature learning. The pseudo-Siamese network is used to extract the feature points and descriptors of the sample image pairs. The feature optimization process is realized through the descriptor constraints between image pairs, which promotes the network to improve the accuracy of feature extraction. The proposed method has been tested in mountain, hilly, flatland, and urban scenarios, respectively. The results demonstrate that the correspondence points extracted by SARPointNet are evenly distributed, are in large quantities (at least ten times that of other methods), and achieve high precision (the root mean square error is less than 1 pixels), which shows great advantages over other methods.

Index Terms—Convolutional neural network (CNN), rational polynomial coefficients (RPCs), registration, synthetic aperture radar (SAR) image.

Manuscript received 17 May 2022; revised 10 July 2022; accepted 28 July 2022. Date of publication 4 August 2022; date of current version 15 August 2022. This work was supported in part by the Key Research and Development Program of the Ministry of Science and Technology under Grant 2018YFB0504905, in part by the High Resolution Earth Observation Systems National Science and Technology Major Projects under Grant GFZX0404130302, and in part by the Civil Aerospace Advance Research Project Grant under Grant D040107, in part by the Postdoctoral Innovation Talent Support Program under Grant BX2021222, and in part by LIESMARS Special Research Funding. (*Corresponding author: Hao Cui.*)

Xin Li, Guo Zhang, Tiancheng Dong, and Boyang Jiang are with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: lixin_1995@whu.edu.cn; guozhang@whu.edu.cn; 2021106190028@whu.edu.cn; jiangboyang@whu.edu.cn).

Taoyang Wang, Hao Cui, and Qian Cheng are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China (e-mail: wangtaoyang@whu.edu.cn; cuihaocasm@whu.edu.cn; qian_cheng_1994@163.com).

Digital Object Identifier 10.1109/JSTARS.2022.3196383

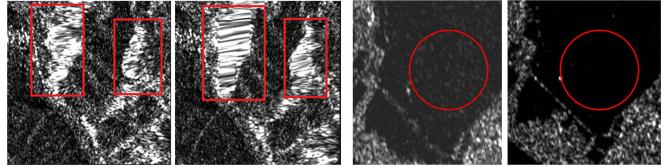


Fig. 1. Imaging differences of SAR images caused by different incident angles and different speckle noises.

I. INTRODUCTION

SPACEBORNE synthetic aperture radar (SAR), which has the advantages of all-weather, all-day, large-scale, and strong penetration for high-resolution Earth observation applications, is an active imaging sensor that uses microwaves to monitor the ground targets. As a useful supplement to optical remote sensing images, spaceborne SAR remote sensing images are widely used in topographic mapping, land and resources surveys, crop yield estimation, hydrological monitoring, and disaster assessment. However, the imaging process of SAR adopts the side-looking oblique imaging mechanism so the image has geometric features such as foreshortening, layover, and shadow [1]. In addition, imaging process is prone to clutter interference, resulting in speckle noise in images (Fig. 1). This makes it difficult to obtain reliable feature points, which brings difficulties to the high-precision registration of SAR images.

In recent years, the technology of deep learning is developing rapidly, it is completely data-driven and can automatically learn features from images. At present, feature extraction of optical remote sensing images based on deep learning is widely studied, but there are few researches on SAR images. There are two major challenges in SAR image registration, which are as follows.

- 1) Lacking of reliable open-sourced datasets. Deep learning technology relies on data driven, but currently there is a lack of sample sets for SAR image feature extraction, which cannot support the training of deep neural networks.
- 2) Lacking of feature extraction network for SAR image registration. At present, the common feature extraction networks lack the application scenarios consideration of remote sensing image features, and the network robustness is low, which unable to support SAR image registration.

In this article, a targeted spaceborne SAR feature learning framework is proposed, which fully considers the data

characteristics and application task characteristics of spaceborne SAR images. The correspondence feature points of the image pairs are automatically generated based on the initial geometric positioning information of the spaceborne SAR. Aiming at the application scenario of SAR feature points, a targeted SAR feature extraction network is designed. The convolutional neural network (CNN) is used to extract the high-dimensional features corresponding to the SAR image, and the feature descriptor is constrained by the designed contrast loss function, which promotes the network to learn high-precision feature point extraction.

The main contributions of this article include the following three aspect.

- 1) We proposed an automatic training-sample generation method based on the initial geometric information of spaceborne SAR. We use the initial RPC information of the SAR image for geometric positioning, and use three classical feature extraction operators to register the overlapping region of different SAR images after positioning, then obtain the corresponding points as the samples for network training.
- 2) We proposed a feature learning network for the SAR image registration, coined SARPointNet. This is a typical pseudo-Siamese network structure, which realizes the input and feature extraction of paired SAR images, and outputs the corresponding feature points and descriptors. We design a targeted descriptor difference loss to promote the model to learn more robust feature information.
- 3) We designed targeted experiments and conducted sufficient experimental verifications in four scenarios: mountain, hilly, flatland, and urban. The results show that our method exhibits significant advantages, extracting the most correct matching while obtaining reliable accuracy compared to the contrasting algorithms.

II. RELATED WORK

A. Image Feature Extraction of Traditional Method

The purpose of image registration is to identify corresponding identical or similar structures/content from two or more images. The ultimate goal is to geometrically warp the perceived image into a common spatial coordinate system of the reference image and align their common regions pixel-for-pixel (i.e., image registration). Remote sensing image registration methods can generally be divided into region-based methods and feature-based methods. The time complexity of the region-based method is relatively high due to the influence of speckle noise and local distortion in SAR images. Therefore, SAR image processing mostly uses feature-based methods for registration. Overall, the registration of SAR images includes three steps: feature extraction, feature description, and feature matching.

Feature point extraction is an especially designed processing method for input images through data dimensionality reduction or reorganization of existing data features to facilitate subsequent use. One of the most widely used algorithms is the scale-invariant feature transform (SIFT) [2], which is widely used in optical image registration. Different from the traditional

passive optical band imaging mode, the spaceborne SAR image adopts active microwave imaging. This imaging mode also leads to strong speckle noise and complex local distortion of the SAR image, resulting in extremely unstable point features detected by the traditional SIFT algorithm. Dellinger et al. [4] proposed a new gradient definition for SAR images, and introduced a new SIFT-like algorithm, SAR-SIFT, to implement a feature detection operator dedicated to SAR images. Ma et al. [5] adopted the gradient definition based on intensity difference to achieve feature detection, and improved the matching performance by enhancing the feature matching algorithm. Ma et al. [6] proposed to use the SAR-SIFT operator to extract feature points from the phase consistency intensity map and match them based on their spatial constraints.

Feature description is a way to describe the characteristics of pixels around a feature point, which can realize the likelihood judgment of different feature descriptors through similarity measurement, to realize the registration of two feature points. By combining the feature detected by SIFT, Fan et al. [3] improved the original SIFT algorithm to detect feature points from the second octave, skipping the dominant direction assignment process to improve the clarity of feature descriptors. Li et al. [9] introduced a radiation-invariant feature transform (RIFT) that utilizes phase-consistent maximum and minimum moments to detect feature points, and proposed a maximum-index mapping MIM method for feature description. In addition, a large number of scholars have tried to use frequency-domain information extraction for feature description. Aguilera et al. [7] proposed the use of multiscale and multidirectional log-Gabor filters to construct feature descriptors that outperform gradient-based descriptors when recording multisensor images. Liu et al. [8] proposed an affine-invariant descriptor, named the maximum stable phase congruence (MSPC) descriptor. The descriptor is constructed by a log-Gabor filter combined with the maximum stable extremal region (MSER) feature to guarantee affine invariance. Yu et al. [10] proposed the construction of a scale space based on nonlinear diffusion, and implemented a frequency-domain-based feature descriptor for the feature points extracted by SAR-SIFT through a multiscale and multidirectional log-Gabor filter.

We can see that feature-based methods require careful engineering and domain knowledge to design feature extractors whether based on spatial information or frequency domain. This makes the handcrafted features somehow specific, but less generalized. In particular, feature detection is affected by various factors such as imaging geometry for spaceborne SAR images. The handcrafted features are not robust and universal to the geometric deformation of SAR images.

B. Image Feature Extraction Based on Deep Learning

Due to its strong ability for deep feature acquisition and nonlinear expression, deep learning technology is applied to image information representation or similarity measurement, as well as parameter regression of image-pairs transformation. The relevant literature review [11], [12] provides a detailed review of learning-based methods in image registration from feature detection, feature matching, and other related algorithms and applications [13]–[15].

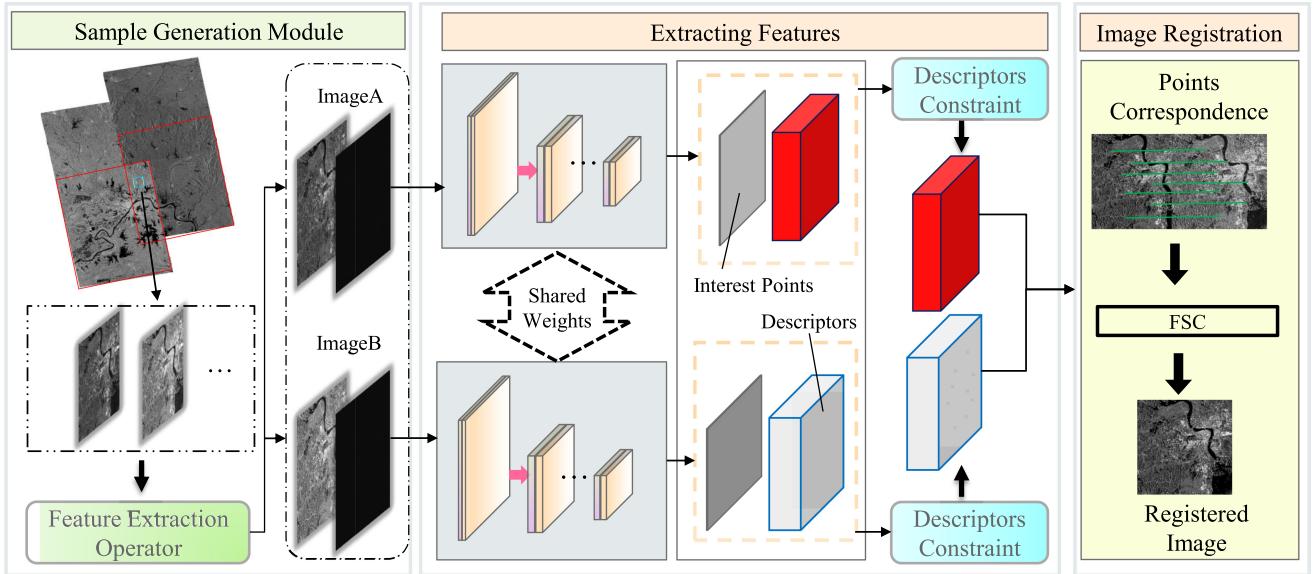


Fig. 2. Framework of the proposed SARPointNet algorithm.

For the learning method of feature detection, CNN can automatically acquire more expressive features than handcrafted feature points [16]–[21]. Learned invariant feature transform (LIFT) [22] is one of the earliest proposed methods to optimize feature detection and description together (i.e., joint learning). He et al. [23] proposed a framework for remote sensing image matching based on Siamese CNN, and verified the effectiveness of the method on different types of satellite datasets with complex background changes. Superpoint [24] implements unsupervised fully convolutional framework training keypoints and descriptors without manual annotation. Han et al. [25] developed a robust method to build more stable feature representations through the internal relationship between features. The research by Fan et al. [27] is based on DRE’s L2 normalized Siamese network (L2-Siamese), which realizes the training of special extractors and achieves subpixel registration accuracy for multiresolution optical remote sensing images. Ye et al. [28] constructed the steerable filters of first- and second-order channels (SFOC) to address the nonlinear radiometric differences. Zhou et al. [29] extracted multiorientated gradient features to depict the structure properties of images, then the gradient feature maps are convolved in a multiscale manner, which produces the multiscale convolutional gradient features (MCGFs). The CNN method of feature detection is to directly transfer the satellite image into the neural network, and obtain the image depth feature through the network. Compared with traditional handcrafted features, this approach can effectively improve the accuracy of registration tasks.

For the learning method of feature matching, many other end-to-end image level learning-based registration methods are presented [26]–[31]. MatchNet is adopted to learn the descriptor and metric simultaneously [35]. Wang et al. [36] proposed a network that learns the mapping, and realizes the matching relationship between patch image pairs and matching labels through an end-to-end structure. M et al. [37] regressed the spatial gradient of deformation and employed a 2-D transform layer to effectively warp one image to another in an end-to-end

manner, verifying the high potential of the proposed method experimentally. Ye et al. [38] proposed a multiscale framework with unsupervised learning, named MU-Net, which directly learns the end-to-end mapping from the image pairs to transformation parameters.

At present, the deep learning methods for remote sensing image registration mainly focus on the registration of optical images, and only a small number of networks have been tested on SAR images with almost no geometric deformation. Due to the special imaging mode and speckle noise of SAR images, the optical image registration network is difficult to be directly applied to SAR images, so there is still a lack of deep learning methods for spaceborne SAR image registration.

III. PROPOSED METHOD

Aiming at the demand for high-precision feature points in the SAR image registration, we propose a targeted feature learning framework, which is inspired by the recent progress of applying deep learning to interest point detection and descriptor learning. As shown in Fig. 2, our method includes two modules, which are as follows.

- 1) Feature sample generation module. We generate feature point samples based on the initial geometric information of spaceborne SAR.
- 2) Feature extraction and image matching. We design a feature learning network for the SAR image matching to realize end-to-end learning of point correspondence.

Then, we can use the trained model to extract the feature points of the image pair for image registration.

A. Feature Sample Generation Module

The biggest difference between satellite imagery and natural pictures is that satellite imagery has geographic significance and location attributes because of its accompanying RPC geometric positioning model [39]. RPC is essentially a rational function model (RFM), it can be understood as a camera model with

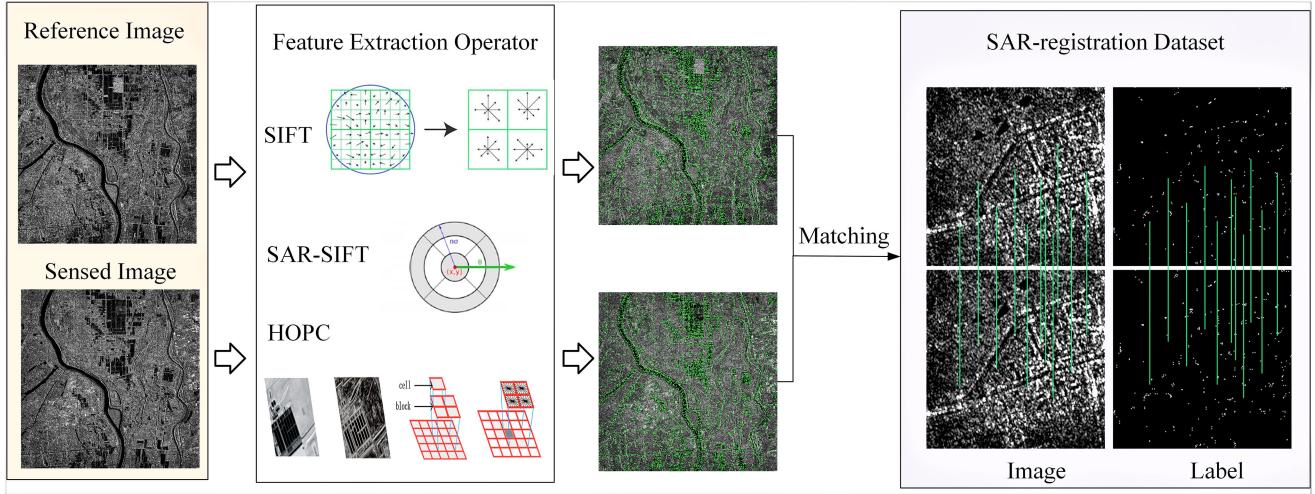


Fig. 3. Method flow for the feature sample generation module.

distortion parameters, which describes the transformation relationship from 3-D geographic coordinates to 2-D image coordinates, generally called from the object side to the image side, which is also called projection. Conversely, inverse projection is a process from image side to object side.

$$(X, Y) = \text{RPC}_{\text{image}}(B, L, H) \quad (1)$$

$$(B, L, H) = \text{RPC}_{\text{image}}^-(X, Y, h) \quad (2)$$

where (X, Y) is the image coordinate; (B, L, H) is the geographic coordinate; h is the elevation value at (X, Y) , and the initial value can be read from the RPC parameters in inverse projection; $\text{RPC}_{\text{image}}()$ represents the projection; and $\text{RPC}_{\text{image}}^-()$ represents the inverse projection.

The training of neural network models for SAR image registration requires image pairs that have been registered in different scenarios. Currently, there is a lack of benchmark datasets for deep learning of SAR images. Therefore, an automatic method for making feature samples based on the initial geometric information of spaceborne SAR is proposed in order to solve the problem of shortage of training data.

In this study, the RPC parameters of spaceborne SAR images are used to estimate the initial position of the overlapping area between SAR images [40]. As shown in Fig. 3, the feature points set of the SAR image is extracted by multiple feature detection operators such as the feature and phase of the overlapping region, and then, the image is matched. The successfully matched points are used as the sample points in the deep network training.

The detailed process of the feature sample generation is as follows.

- 1) The overlapping regions of different images are obtained based on the initial RPC geometric positioning model of SAR images.
- 2) Traditional image feature extraction operators are used to extract multisource features for the overlapping image pairs. In this article, SIFT [2], SAR-SIFT [4], and

TABLE I
SAR-REGISTRATION DATASETS PARAMETER DETAILS

Item	Parameters
Data source	GF3 Fine Straight II mode
Data resolution	10 m
Type	TIFF+RPC
Feature Point Source	SIFT、SAR-SIFT、HOPC
Correspondence points accuracy	Less than 1 pixel
Sample size	256 x 256
Sample data format	Tiff
Number of sample sets	9000

histogram of orientated phase congruency (HOPC) [41] are utilized for feature extraction.

- 3) All the extracted feature points are coupled into two comprehensive point sets, and the points correspondence is established through the matching algorithm. The fast sample consensus (FSC) [43] algorithm realizes the elimination of mismatched points and generates correspondence point pairs with high precision.
- 4) Finally, the images are cropped into 256×256 image pairs adapted to the training of the network, respectively. The method of generating the label corresponding to the image is to automatically generate the registered point pairs according to the transformation relationship of the image cropping.

The SAR-registration dataset covers different scenarios in Hubei Province, such as mountains, hilly, flatland, and urban, as shown in Fig. 4. The terrain of Hubei Province is roughly surrounded by mountains in the East, West, and North, and the middle is low and flat, showing an incomplete basin slightly open to the south. In the total area of the province, the proportion of mountains and hills is more than 80%, and the proportion of flat areas and lakes is less than 20%. Table I lists the detailed description of the SAR-registration dataset, including data source, data resolution, feature point source, sample parameters, etc.

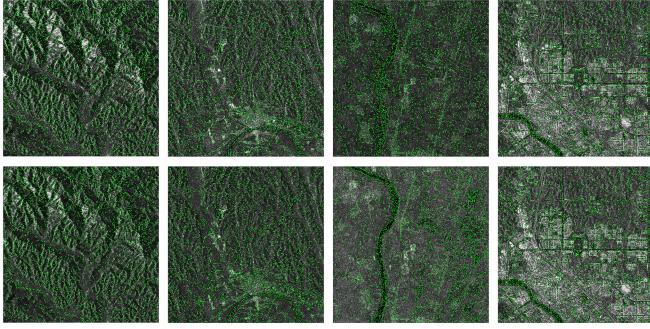


Fig. 4. Distribution of feature points set in SAR image pairs. The four scenarios are mountain, hilly, flatland, and urban from left to right.

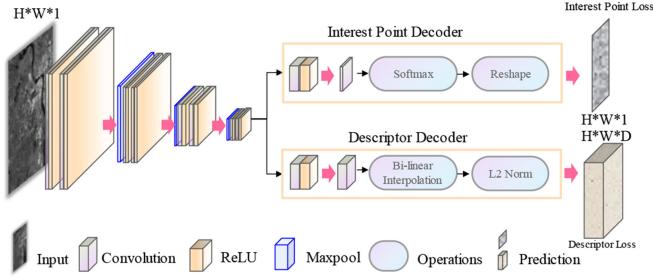


Fig. 5. Schematic diagram of the encoder and decoder.

B. SARPointNet Architecture

The basic components for learning-based feature extraction are inspired by existing feature point extraction networks [24] in our SARPointNet architecture. Taking into account the data characteristics and task requirement of spaceborne SAR images registration, the input of this network is a grayscale image and the output are keypoint heatmap and descriptor. The network model has a single shared encoder to process and reduce the dimensionality of the input image. When completing encoding, the network is deconstructed into two decoder branches based on decoding usage—one for interest point detection and the other for interest point description, both branches learn their own task-specific weights. Network parameters are shared between the two tasks, improving network efficiency by sharing the ability to compute and represent.

As shown in Fig. 5, VGG style is used as a conventional image processing encoder to reduce image dimensionality. The encoder consists of a convolutional layer, a pooling layer, and a nonlinear activation function. The encoder has a total of eight convolutional layers, the convolution kernel is 3×3 , and the size of the convolutional layer is 64-64-64-64-128-128-128-128. There is a 2×2 max pooling layer every two layers. The decoder is divided into two branches, one for feature point detection and the other for feature point description. Each decoder starts with a 256-dimensional 3×3 convolutional layer followed by a 1×1 convolutional layer with 65 units and 256 units for interest point detectors and descriptors, respectively. The computational load of the model is reduced by introducing an explicit decoder [42] in the feature point detection branch. The convolutional layers in the network are followed by ReLu and Batch Norm for normalization.

In the learning stage of feature point extraction, the input of the pseudo-Siamese network architecture is the image pair ImageA, ImageB $\in R^{H \times W}$, and the corresponding points from label. After feature extraction, the decoder outputs the keypoints heatmaps ImageA_det $\in R^{H \times W \times 1}$, ImageB_det $\in R^{H \times W \times 1}$, and descriptors: ImageA_desc $\in R^{H \times W \times D}$, ImageB_desc $\in R^{H \times W \times D}$ for ImageA and ImageB.

In the image registration stage, we obtain initial correspondences from keypoints and their descriptors. To obtain keypoints, nonmaxima suppression (NMS) and thresholding on the heatmap is utilized to filter out redundant candidates. Descriptors are sampled from ImageA_desc and ImageB_desc using bilinear interpolation. The nearest neighbor matching process is performed using two sets of keypoints and descriptors to form N correspondences, which are used as the input for subsequent fine registration.

Based on the obtained two sets of initial registration points, we apply the FSC elimination operator to reduce the proportion of the mismatched points. Similar to the outlier detection method random sample consistency algorithm (RANSAC), FSC estimates a model from observed data through a random sampling and voting scheme. However, FSC can get more correct matches in fewer iterations than RANSAC, effectively improving the accuracy of the results.

C. Loss Function

The final loss function is the sum of three losses: Loss_p for the feature point detector, Loss_d for the descriptor, and contrast loss function (Loss_{cp}) is an especially designed descriptor optimization for the SAR image feature as descriptor constraints. We use paired images as input to the pseudo-Siamese network, which allows us to optimize three losses simultaneously, as shown in Fig. 2, we balance the final loss with λ and α , respectively

$$\begin{aligned} \text{TotalLoss} = & \text{Loss}_p(X, Y) + L_p(X', Y') \\ & + \lambda \text{Loss}_d(D, D') - \alpha \text{Loss}_{cp} \end{aligned} \quad (3)$$

where (X, Y) and (X', Y') are the detected points and the true values in the corresponding labels, respectively. (D, D') are the feature descriptors of the corresponding points, respectively.

The full convolutional cross entropy is used to calculate Loss_p for feature point detector loss, where $x_{hw} \in X, y_{hw} \in Y$, are the minimum convolution units of the input image, respectively.

$$\text{Loss}_p(X, Y) = \frac{1}{H_c W_c} \sum_{\substack{h=1 \\ w=1}}^{H_c, W_c} l_p(x_{hw}, y_{hw}) \quad (4)$$

where

$$l_p(x_{hw}, y) = -\log \left(\frac{\exp(x_{hw})}{\sum_{n=1}^{65} \exp(x_{hwn})} \right). \quad (5)$$

The descriptor Loss_d is computed on all corresponding points of the input image pair, where $d_{hw} \in D$ comes from the reference image, $d_{h'w'} \in D'$ comes from the sensed image. A weighting term of $\lambda_d = 250$ is used to maintain the learning

balance between ImageA_desc and ImageB_desc. The positive margin m_p and negative margin m_n was adopted as described in [24]. The descriptor loss is defined as

$$\text{Loss}_d(D, D'; \text{if_match}) = \frac{1}{(H_c W_c)^2} \sum_{\substack{h=1 \\ w=1}}^{H_c, W_c} l_d(d_{hw}, d'_{hw}; \text{if_match}) \quad (6)$$

where

$$\begin{aligned} l_d(d_{hw}, d'_{hw}; \text{if_match}) &= \lambda_d * \text{if_match} * \max(0, m_p - d^T d') \\ &+ (1 - \text{if_match}) * \max(0, d^T d' - m_n). \end{aligned} \quad (7)$$

The descriptor difference loss Loss_{cp} calculates the difference of descriptors point by point in D and D' , respectively. The contrast loss function is defined as

$$\text{Loss}_{cp} = \sum_n^{i,j} (D_i^{\text{desc}} - D_j^{\text{desc}})^2. \quad (8)$$

According to the number of feature points extracted from SAR image pairs, the Loss_{cp} for descriptors $\text{ImageA_desc} \in R^{H \times W \times D}$ and $\text{ImageB_desc} \in R^{H \times W \times D}$ are solved one by one within the 256 dimensions vector. For example, ImageA_desc represents the descriptor vector of 3000 feature points extracted from A image, then $n = 3000$, D_i^{desc} in (8) is a $1 \times 1 \times 256$, which represents the descriptor of the point i . The 256×256 image is solved point by point in sequence so that the Loss_{cp} of the A image can be calculated. The calculation method of B images is the same, and the sum of the two is the fourth item in formula (3).

The purpose of the descriptors constraint by the descriptor difference loss is to ensure the uniqueness of the SAR image feature descriptor. Spaceborne SAR images are slant range images, resulting in an inconsistent resolution in azimuth and range. The influence of coherent speckle noise makes the edge of the target blur and the sharpness decreases. In addition, when the incident angle of SAR image imaging changes, the SAR image will also reflect a large imaging difference. We design the loss function to constrain descriptors, on the one hand, it can eliminate the interference points caused by speckle noise in SAR images. On the other hand, the wrong feature points in the overlapping, knurling, and other areas shown in SAR images with different incident angles will also be effectively removed. Therefore, the loss function Loss_{cp} is designed as descriptors constraint in order to minimize the influence of these factors on the SAR image feature extraction and ensure the accuracy of feature descriptors.

IV. EXPERIMENTS

A. Datasets and Experimental Details

1) *Experimental Design:* For the automatic spaceborne SAR image feature learning framework SARPointNet proposed in

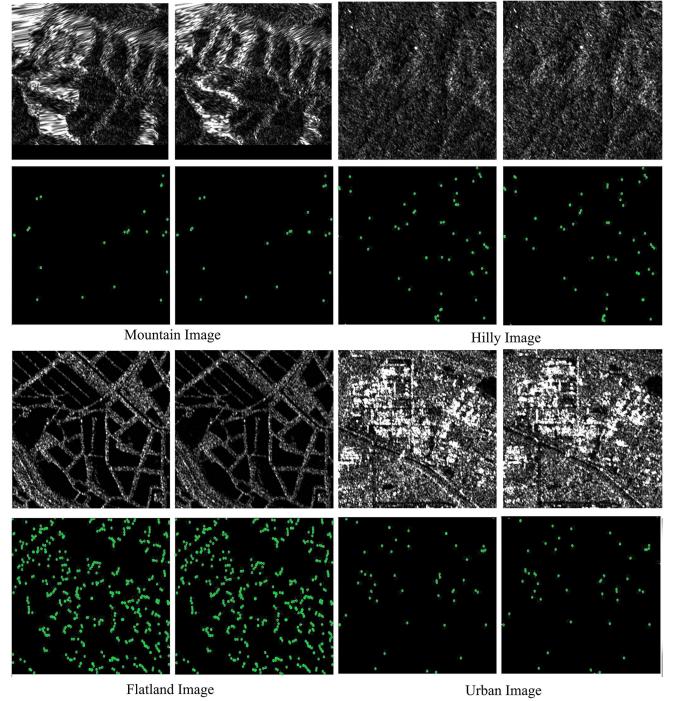


Fig. 6. Examples of SAR-registration datasets in different scenarios.

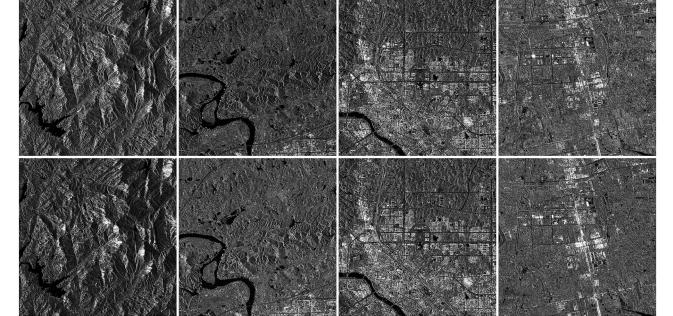


Fig. 7. SAR image registration experiment test dataset.

this article, it is compared and evaluated with the existing advanced methods. The contrast loss function designed in this article is used as a descriptor constraint to verify its effectiveness through a set of ablation experiments.

a) *Comparative experiment:* Several existing traditional image registration methods are adopted for experiments: SIFT, SAR-SIFT, and HOPC. For SIFT, SAR-SIFT, and HOPC, matching is performed after feature extraction by the Euclidean distance ratio between the nearest neighbor and the next nearest neighbor of the corresponding feature, and the threshold is set to 0.7. In addition, we also increased the comparison of the Superpoint framework based on the CNN. Supervised training is performed using the SAR-registration dataset based on the pretrained models, training methods, and parameter settings provided by the author. In order to ensure the consistency of experimental conditions, we also use the aforementioned matching method to achieve points correspondence after extracting features and descriptors.

TABLE II
COMPARATIVE EXPERIMENTAL RESULTS OF SPACEBORNE SAR IMAGE REGISTRATION IN DIFFERENT SCENARIOS

Test images	Evaluation Factor	SIFT	SAR-SIFT	HOPC	Superpoint	SARPointNet
Mountain	Keypoint	156	330	197	158	1665
	RMSE (pixel)	0.864	0.853	1.307	12.589	0.711
	Times(s)	22.2	35.529	6.1	12.0	9.0
Hilly	Keypoint	114	379	181	241	1627
	RMSE (pixel)	0.871	0.779	1.166	35.613	0.505
	Times(s)	22.4	32.2	6.1	12	9
Flatland	Keypoint	104	262	100	6	748
	RMSE (pixel)	0.870	0.846	1.313	130.677	0.861
	Times(s)	21.2	28.7	5.8	12	9
Urban	Keypoint	78	176	140	8	1627
	RMSE (pixel)	0.612	0.879	1.313	9.005	0.981
	Times(s)	28.4	34.5	6.7	12	9

The best results are shown in bold.

b) Ablation experiment: This group of experiments is implemented by setting the presence or absence of descriptor constraints. Both groups of experiments use the training dataset in Section IV-A2 for model training and the image registration adopts the method in this article.

For the evaluation of image registration, both the comparative experiment and the ablation experiment are verified by the testing dataset in Section IV-A2, and finally, analyze the two sets of experimental results through the relevant evaluation standards. The experimental results of the two groups were analyzed according to the relevant evaluation criteria.

2) Datasets:

a) Training dataset: We randomly reshuffle the SAR-registration dataset for training and validation. The final training dataset consists of 8000 image pairs for training and 1000 image pairs for validation. Fig. 6 shows the dataset form of training samples in different scenarios.

b) Testing dataset: For the designed comparative experiment, we carefully selected four pairs of SAR images from different scenarios for testing. The data size is 2000×2000 and the data source is the SAR image data of GF3 Fine Straight II imaging mode [40] in Hubei Province, China. We named four different scene data as: mountain images, hilly images, flatland images, and urban images. As shown in Fig. 7.

3) Parameters Setting: The experimental parameters are set as follows: the model is trained for 18 000 iterations in 6 h; the learning rate is 0.0001 and the detection threshold of the feature detector is 0.015. For the loss function balance parameter mentioned in Section III-C, a weighting term of $\lambda_d = 250$ is used to maintain the learning balance between ImageA_desc and ImageB_desc, where the descriptor dimension $D = 256$. The balance parameter $\lambda = 1$ for the overall descriptor, and the parameter $\alpha = 0.1$ for the contrast loss function. All training was done with a batch size of 8.

4) Evaluation Criterion: In this article, the evaluation methods of spaceborne SAR image registration mainly include the registration accuracy and the number of correct matching points, named Keypoints number.

a) Registration accuracy: Registration accuracy was assessed by the root mean square error (RMSE) criterion. The trained SARPointNet model automatically extracts N corresponding point pairs from the reference image and the sensing image, where the corresponding point coordinates are $\{(x_i, y_i), (x'_i, y'_i)\}$. The Keypoints are carefully selected and refined with outlier removal methods to minimize residuals. Therefore, these point pairs are used as a reference to verify the accuracy of affine transformation model parameters. RMSE is calculated according to the following formula:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - x'_i)^2 + (y_i - y'_i)^2}. \quad (9)$$

b) Keypoints number: The number of correct correspondences is used as a criterion [44] to evaluate the robustness of the proposed method.

B. Experimental Results

Due to the large differences in the imaging characteristics of spaceborne SAR images, there are still a large number of mismatched points after brute force matching. In the comparative experiments setup in this article, RANSAC [43] is uniformly used to eliminate the mismatched points.

Note: Since the original SIFT algorithm cannot extract effective Keypoints in the experiments, we use the improved SIFT algorithm in [5] for experimental verification.

Table II shows the evaluation results of image registration in four scenarios. In the experiment of mountain images, the proposed framework had the most Keypoints number of 1 665 and the RMSE was 0.711 pixels, which is better than other methods. In the experiment of hilly images, the proposed framework still extracts the most Keypoints number of 1 627, and the RMSE is still optimal in several methods. In the experiment of flatland image, the proposed method extracted 748 pairs of Keypoints with 0.861 for RMSE, which was better than the result of SIFT and slightly lower than the RMSE of SAR-SIFT. In the experiments of urban images, SIFT algorithm obtains the best

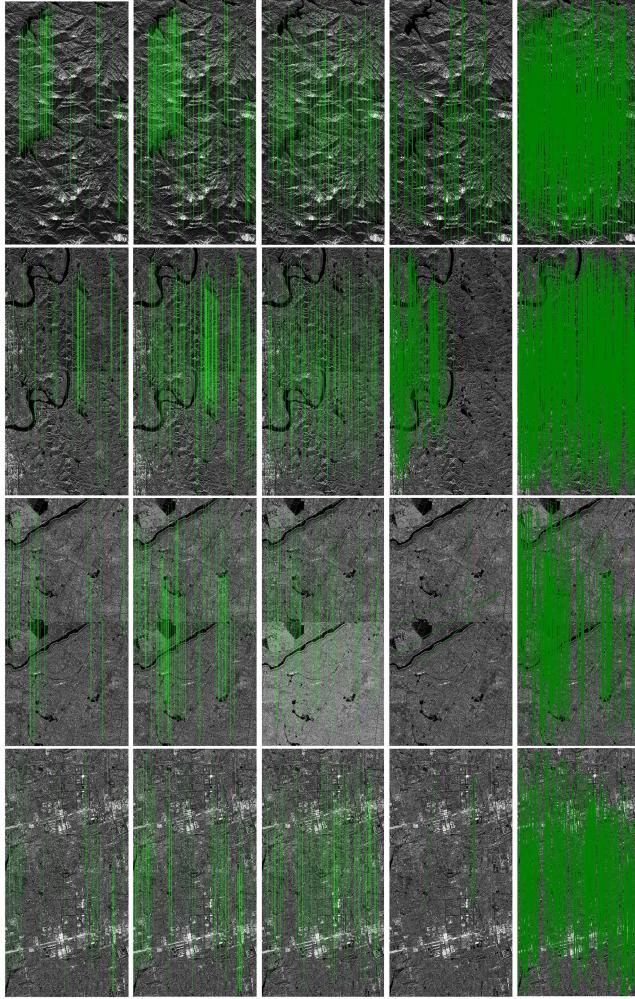


Fig. 8. Correct matching connection diagram of the five methods in different scenarios. From top to bottom, there are four scenarios: Mountain, hilly, flatland, and urban. Each row represents the registration results of SIFT, SAR-SIFT, HOPC, Superpoint, and SARPointNet from left to right.

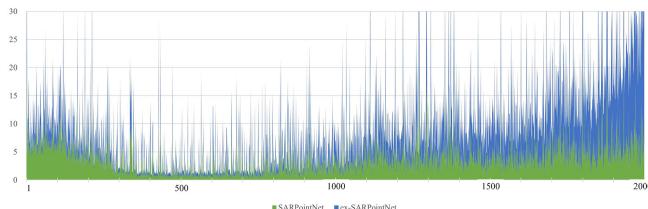


Fig. 9. Comparison of registration errors between ex-SARPointNet and SARPointNet.

RMSE value of 0.612 and the proposed framework achieves 1627 pairs of Keypoints. The registration results represent that the registration performance of our proposed method is comparable to other methods in terms of accuracy in the four scenarios. Moreover, our SARPointNet is more prominent in mountain and hilly scenarios. Although the RMSE in flatland and urban scenarios is not optimal, the registration accuracy of 1 pixel can also guarantee subsequent high-precision processing.

As shown in Fig. 8, it is the correct matching connection diagram of the five methods in different scenarios. By analyzing

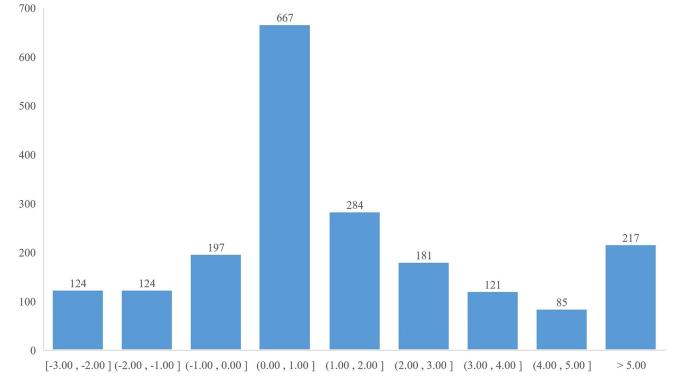


Fig. 10. RMSE difference between SARPointNet and ex-SARPointNet.

the accuracy results of the image registration in Table II and the visualized registration results in Fig. 8. In mountain and hilly scenes, the SIFT cannot achieve enough Keypoints due to the influence of geometric deformation, while flatland and urban scenes are more affected by speckle noise. The RMSE of HOPC is lower than the results of other methods mainly due to the geometric and noise differences between the reference image and sensed image, and it is matched by the block-by-block principle, which leads to the loss of many effective features. Since Superpoint is a network designed for natural images, the scale and characteristics of SAR images and natural images are quite different. Although the network has been trained on the SAR-registration dataset, it performs poorly on registration accuracy and Keypoints number due to the lack of critical feature constraints. The framework proposed in this article exhibits robustness to spaceborne SAR image registration, especially for scenes with varying elevations similar to the results of SAR-SIFT. Our SARPointNet pipeline has obvious advantages over other methods in Keypoints. Except for the flatland, the number of Keypoints obtained from SARPointNet is more than 10 times that of the other methods in the other three scenarios. Compared with the SIFT and SAR-SIFR algorithms, the Keypoints distribution of the HOPC algorithm is relatively uniform. While the method in this article ensures the uniform distribution of extracted features, the feature accuracy rate is higher so as to achieve the most Keypoints number.

V. DISCUSSION

A. Ablation Experiment

To evaluate the effectiveness of the contrast loss function proposed in this article, a set of ablation experiments are designed by setting the presence or absence of descriptor constraints. The train data mentioned in Section IV-A2 were used as training set and validation set in both experiments for experimental verification. We named the model trained without descriptor constraints as “ex-SARPointNet” and the other as “SARPointNet.”

During model training, the inputs of both methods are image pairs and matching point pairs, and the model parameters are set as detailed in Section IV-A3. After training, image registration experiments were performed on the same test datasets (2000 image pairs to be registered).

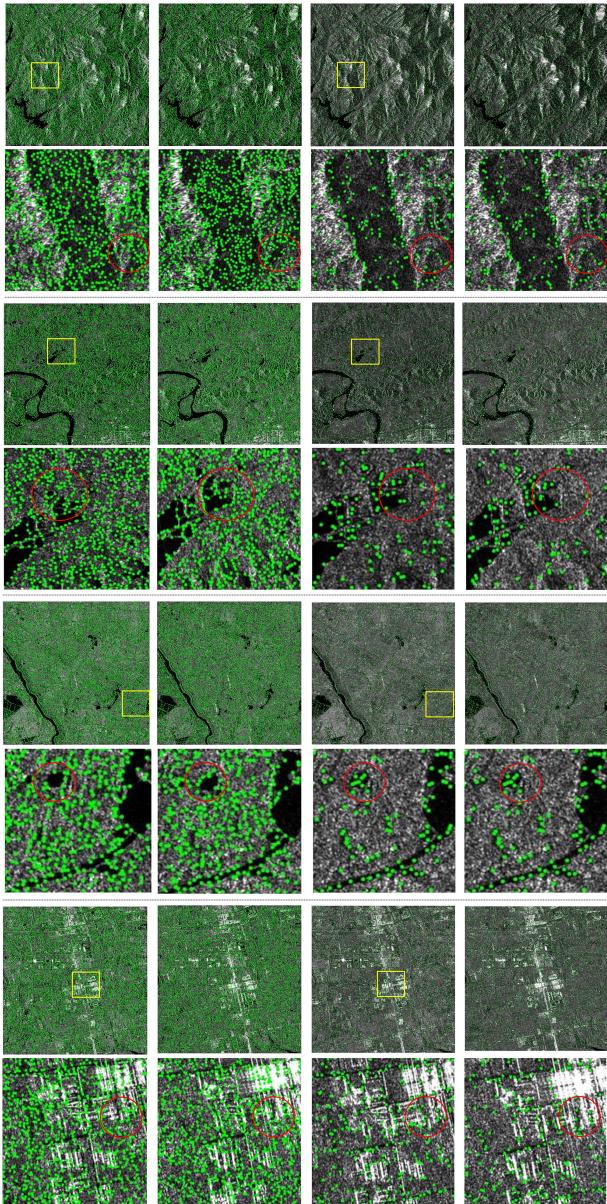


Fig. 11. Comparison of ablation experiment Results. The left two columns are the feature extraction results of ex-SARPointNet, the left two columns are the feature extraction results of SARPointNet. From top to bottom, there are four experimental scenarios: mountain, hilly, flatland, and urban. The second row is the enlarged detail of the yellow box of the full-size image in the first row. The red boxes of the left and right image pairs are marked as the difference results of the two methods.

As shown in Fig. 9, the registration accuracy of SARPointNet is improved by different margins compared to ex-SARPointNet on the test dataset. In Fig. 10, it can be seen that there are 888 image pairs with a registration accuracy improvement of more than 1 pixel, and 667 images pairs with an accuracy improvement of between 0 and 1 pixels. What is unsatisfactory is that there are still 445 image pairs whose registration accuracy has not been effectively improved.

In addition, the loss function of maximization difference for descriptor constraints makes it possible to effectively avoid the

problem of nonfeature point misidentification due to speckle noise during the feature extraction stage. We count the performance of SARPointNet in the four scenarios of the test dataset in Section IV-A.2, which has far fewer initial feature points than ex-SARPointNet to achieve more correct matches. As shown in the Fig. 11, the number of initial feature points extracted by SARPointNet is reduced by an average of 7580 points compared with ex-SARPointNet. In addition, the image registration accuracy is improved by an average of 0.28 pixels.

Therefore, the loss function of maximization difference for descriptors designed in this article effectively improves the accuracy of feature extraction, thereby improving the registration accuracy of spaceborne SAR images. The aforementioned quantitative and qualitative assessments demonstrate the power of our proposed method on satellite SAR remote sensing images, especially in complex and content-rich scenarios. We believe this advantage comes from the ability of neural networks to learn more semantic features.

B. Limitation

Although SARPointNet implements the application of CNNs to spaceborne SAR image registration tasks, the experimental results illustrate that there are still some unsatisfactory aspects of the network architecture.

- 1) There are still mismatched points extracted by the network architecture and removal outliers need to be executed. The accuracy of corresponding points from the network needs to be further improved.
- 2) The method in this article does not perform well in urban and flat areas with small terrain fluctuations. It is necessary to further optimize the network structure and enhance the robustness to multitype scenarios.

VI. CONCLUSION

In view of the demand for high-precision feature points in SAR image registration, we innovatively proposed an automated feature learning framework. A feature sample generation module based on the initial geometric information of spaceborne SAR is constructed, and then, a pseudo-Siamese feature learning network SARPointNet for SAR image feature learning is utilized for sample training. We demonstrate that our performance is on par with classic methods, both qualitative and quantitative results are included in this article to support the claim. The future work mainly focuses on the fusion expression ability of the network for global and local features, and further improving the generalization ability of the network.

REFERENCES

- [1] H. Wang et al., "Layover compensation method for regional spaceborne SAR imagery without GCPs," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8367–8381, Oct. 2021, doi: [10.1109/TGRS.2020.3045505](https://doi.org/10.1109/TGRS.2020.3045505).
- [2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004, doi: [10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94).

- [3] B. Fan, C. Huo, C. Pan, and Q. Kong, "Registration of optical and SAR satellite images by exploring the spatial relationship of the improved SIFT," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 657–661, Jul. 2013, doi: [10.1109/LGRS.2012.2216500](https://doi.org/10.1109/LGRS.2012.2216500).
- [4] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A SIFT-like algorithm for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 453–466, Jan. 2015, doi: [10.1109/TGRS.2014.2323552](https://doi.org/10.1109/TGRS.2014.2323552).
- [5] W. Ma et al., "Remote sensing image registration with modified SIFT and enhanced feature matching," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 3–7, Jan. 2017, doi: [10.1109/LGRS.2016.2600858](https://doi.org/10.1109/LGRS.2016.2600858).
- [6] W. Ma et al., "Remote sensing image registration based on phase congruency feature detection and spatial constraint matching," *IEEE Access*, vol. 6, pp. 77554–77567, 2018, doi: [10.1109/ACCESS.2018.2883410](https://doi.org/10.1109/ACCESS.2018.2883410).
- [7] C. A. Aguilera, A. D. Sappa, and R. Toledo, "LGHHD: A feature descriptor for matching across non-linear intensity variations," in *Proc. IEEE Int. Conf. Image Process.*, 2015, pp. 178–181, doi: [10.1109/ICIP.2015.7350783](https://doi.org/10.1109/ICIP.2015.7350783).
- [8] X. Liu et al., "A novel affine and contrast invariant descriptor for infrared and visible image registration," *Remote Sens.*, vol. 10, no. 4, 2018, Art. no. 658, doi: [10.3390/rs10040658](https://doi.org/10.3390/rs10040658).
- [9] J. Li, Q. Hu, and M. Ali, "RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform," *IEEE Trans. Image Process.*, vol. 29, pp. 3296–3310, 2020, doi: [10.1109/TIP.2019.2959244](https://doi.org/10.1109/TIP.2019.2959244).
- [10] Q. Yu and D. Ni, "Universal SAR and optical image registration via a novel SIFT framework based on nonlinear diffusion and a polar spatial-frequency descriptor," *ISPRS J. Photogrammetry Remote Sens.*, vol. 171, pp. 1–17, 2021.
- [11] J. Ma et al., "Image matching from handcrafted to deep features: A survey," *Int. J. Comput. Vis.*, vol. 129, pp. 23–79, 2021, doi: [10.1007/s11263-020-01359-2](https://doi.org/10.1007/s11263-020-01359-2).
- [12] Y. Wu et al., "Computational intelligence in remote sensing image registration: A survey," *Int. J. Autom. Comput.*, vol. 18, pp. 1–17, 2021, doi: [10.1007/s11633-020-1248-x](https://doi.org/10.1007/s11633-020-1248-x).
- [13] B. Zhao et al., "Deep spatial-temporal joint feature representation for video object detection," *Sensors*, vol. 18, 2018, Art. no. 774, doi: [10.3390/s18030774](https://doi.org/10.3390/s18030774).
- [14] C. Deng et al., "Learning dynamic spatial-temporal regularization for UAV object tracking," *IEEE Signal Process. Lett.*, vol. 28, pp. 1230–1234, 2021, doi: [10.1109/LSP.2021.3086675](https://doi.org/10.1109/LSP.2021.3086675).
- [15] C. Deng et al., "FAR-Net: Fast anchor refining for arbitrary-oriented object detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 6505805, doi: [10.1109/LGRS.2022.3144513](https://doi.org/10.1109/LGRS.2022.3144513).
- [16] Y. Verdie et al., "TILDE: A temporally invariant learned DEtector," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 5279–5288, doi: [10.1109/CVPR.2015.7299165](https://doi.org/10.1109/CVPR.2015.7299165).
- [17] L. Zhang and S. Rusinkiewicz, "Learning to detect features in texture images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 6325–6333, doi: [10.1109/CVPR.2018.00662](https://doi.org/10.1109/CVPR.2018.00662).
- [18] A. B. Laguna and K. Mikolajczyk, "Key.Net: Keypoint detection by handcrafted and learned CNN filters revisited," *IEEE Trans. Pattern Anal. Mach. Intell.*, to be published, doi: [10.1109/TPAMI.2022.3145820](https://doi.org/10.1109/TPAMI.2022.3145820).
- [19] Y. Ono et al., "LF-Net: Learning local features from images," in *Proc. 32nd Int. Conf. Neural Inf. Process. Syst.*, Curran Associates Inc., Red Hook, NY, USA, 2018, pp. 6237–6247.
- [20] X. Shen et al., "RF-Net: An end-to-end image matching network based on receptive field," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 8124–8132, doi: [10.1109/CVPR.2019.00832](https://doi.org/10.1109/CVPR.2019.00832).
- [21] L. Tang et al., "A scale-aware pyramid network for multi-scale object detection in SAR images," *Remote Sens.*, vol. 14, no. 4, 2022, Art. no. 973, doi: [10.3390/rs14040973](https://doi.org/10.3390/rs14040973).
- [22] K. M. Yi et al., "LIFT: Learned invariant feature transform," in *Computer Vision – ECCV 2016* (Lecture Notes in Computer Science), B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Berlin, Germany: Springer, 2016, pp. 467–483, doi: [10.1007/978-3-319-46466-4_28](https://doi.org/10.1007/978-3-319-46466-4_28).
- [23] H. He et al., "Matching of remote sensing images with complex background variations via Siamese convolutional neural network," *Remote Sens.*, vol. 10, no. 2, 2018, Art. no. 355, doi: [10.3390/rs10020355](https://doi.org/10.3390/rs10020355).
- [24] D. DeTone et al., "SuperPoint: Self-supervised interest point detection and description," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2018, pp. 337–33712, doi: [10.1109/CVPRW.2018.00060](https://doi.org/10.1109/CVPRW.2018.00060).
- [25] C. Deng et al., "High-performance visual tracking with extreme learning machine framework," *IEEE Trans. Cybern.*, vol. 50, no. 6, pp. 2781–2792, Jun. 2020, doi: [10.1109/TCYB.2018.2886580](https://doi.org/10.1109/TCYB.2018.2886580).
- [26] Y. Han et al., "Adaptive feature representation for visual tracking," in *Proc. IEEE Int. Conf. Image Process.*, 2017, pp. 1867–1870, doi: [10.1109/ICIP.2017.8296605](https://doi.org/10.1109/ICIP.2017.8296605).
- [27] R. Fan, B. Huo, J. Liu, J. Yang, and Z. Hong, "Registration of multiresolution remote sensing images based on L2-Siamese model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 237–248, 2021, doi: [10.1109/JSTARS.2020.3038922](https://doi.org/10.1109/JSTARS.2020.3038922).
- [28] Y. Ye, T. Tang, B. Zhu, C. Yang, B. Li, and S. Hao, "A multiscale framework with unsupervised learning for remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5622215, doi: [10.1109/TGRS.2022.3167644](https://doi.org/10.1109/TGRS.2022.3167644).
- [29] L. Zhou, Y. Ye, T. Tang, K. Nan, and Y. Qin, "Robust matching for SAR and optical images using multiscale convolutional gradient features," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4017605, doi: [10.1109/LGRS.2021.3105567](https://doi.org/10.1109/LGRS.2021.3105567).
- [30] Y. Tian, B. Fin, and F. Wu, "L2-Net: Deep learning of discriminative patch descriptor in Euclidean space," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 6128–6136, doi: [10.1109/CVPR.2017.649](https://doi.org/10.1109/CVPR.2017.649).
- [31] X. Li et al., "Dual-resolution correspondence networks," in *Proc. 34th Int. Conf. Neural Inf. Process. Syst.*, Curran Associates Inc., Red Hook, NY, USA, 2020, pp. 17346–17357.
- [32] H. Sokooti et al., "Nonrigid image registration using multi-scale 3d convolutional neural networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted. Intervention*, 2017, pp. 232–239.
- [33] O. Poursaeed et al., "Deep fundamental matrix estimation without correspondences," in *Proc. Eur. Conf. Comput. Vis. Workshop*, 2018, pp. 1–13.
- [34] D. DeTone et al., "Deep image homography estimation," 2016, *arXiv:1606.03798*.
- [35] X. Han et al., "MatchNet: Unifying feature and metric learning for patch-based matching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 3279–3286, doi: [10.1109/CVPR.2015.7298948](https://doi.org/10.1109/CVPR.2015.7298948).
- [36] S. Wang et al., "A deep learning framework for remote sensing image registration," *ISPRS J. Photogrammetry Remote Sens.*, vol. 145, pp. 148–164, 2018, doi: [10.1016/j.isprsjprs.2017.12.012](https://doi.org/10.1016/j.isprsjprs.2017.12.012).
- [37] M. Papadomanolaki et al., "Unsupervised multistep deformable registration of remote sensing imagery based on deep learning," *Remote Sens.*, vol. 13, no. 7, 2021, Art. no. 1294, doi: [10.3390/rs13071294](https://doi.org/10.3390/rs13071294).
- [38] Y. Ye et al., "A robust multimodal remote sensing image registration method and system using steerable filters with first-and second-order gradients," *ISPRS J. Photogrammetry Remote Sens.*, vol. 188, pp. 331–350, 2022, doi: [10.1016/j.isprsjprs.2022.04.011](https://doi.org/10.1016/j.isprsjprs.2022.04.011).
- [39] G. Zhang et al., "Analysis and test of the substitutability of the RPC model for the rigorous sensor model of spaceborne SAR imagery," *Acta Geodaetica et Cartographica Sinica*, vol. 39, pp. 264–270, 2010.
- [40] T. Wang et al., "Large-scale orthorectification of GF-3 SAR images without ground control points for China's land area," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5221617, doi: [10.1109/TGRS.2022.3142372](https://doi.org/10.1109/TGRS.2022.3142372).
- [41] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust registration of multimodal remote sensing images based on structural similarity," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2941–2958, May 2017, doi: [10.1109/TGRS.2017.2656380](https://doi.org/10.1109/TGRS.2017.2656380).
- [42] W. Shi et al., "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1874–1883, doi: [10.1109/CVPR.2016.207](https://doi.org/10.1109/CVPR.2016.207).
- [43] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM Readings Comput. Vis.*, vol. 24, no. 6, pp. 726–740, 1987.
- [44] Y. Wu, W. Ma, M. Gong, L. Su, and L. Jiao, "A novel point-matching algorithm based on fast sample consensus for image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 43–47, Jan. 2015, doi: [10.1109/LGRS.2014.2325970](https://doi.org/10.1109/LGRS.2014.2325970).
- [45] Q. Li, S. Qi, Y. Shen, D. Ni, H. Zhang, and T. Wang, "Multispectral image alignment with nonlinear scale-invariant keypoint and enhanced local feature matrix," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 7, pp. 1551–1555, Jul. 2015, doi: [10.1109/LGRS.2015.2412955](https://doi.org/10.1109/LGRS.2015.2412955).



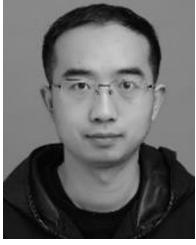
Xin Li received the master's degree in surveying and mapping engineering from Wuhan University, Wuhan, China, in 2020. He is currently working toward the Ph.D.E. degree in photogrammetry and remote sensing with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China.

His main research interests include geometry processing of spaceborne synthetic aperture radar and deep learning in remote sensing.



Qian Cheng received the B.Eng. degree in surveying and mapping in 2017 from the School of Geomatics, Liaoning Technical University, Fuxin, China, where he is currently working toward the master's degree in photogrammetry and remote sensing.

His main research interests include geometry processing of spaceborne imagery and high accuracy image matching.



Taoyang Wang received the B.E. and Ph.D.E. degrees in photogrammetry and remote sensing from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2007 and 2012, respectively. His doctoral dissertation was on the block adjustment of high-resolution satellite remote sensing imagery.

He has been working with the School of Remote Sensing and Information Engineering, Wuhan University, since 2014, where he became an Associate Research Fellow in 2015. His research interests include space photogrammetry, geometry processing of spaceborne optical/SAR/InSAR imagery, target detection, and recognition based on satellite video.



Hao Cui (Member, IEEE) received the B.S. degree in surveying and mapping engineering from Shanxi Datong University, Datong, China, in 2015, the M.S. degree in geographic information systems from Lanzhou Jiaotong University, Lanzhou, China, in 2018, and the Ph.D.E. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2021.

He is currently involved in postdoctoral research with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing.

His research interests include image quality improvement and relative radiation correction and deep learning in remote sensing.



Guo Zhang received the B.E. and Ph.D.E. degrees in photogrammetry and remote sensing from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2000 and 2005, respectively. His doctoral dissertation was on the rectification of high-resolution remote sensing imaging under lack of ground control points.

He has been working with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China, since 2005, where he became a Professor in 2011. His research interests include space photogrammetry, geometry processing of spaceborne optical/SAR/InSAR imagery, altimetry, and high-accuracy image matching.



Tiancheng Dong received the master's degree in geophysical prospecting and information technology from the China University of Geosciences, Wuhan, China, in 2020. He is currently working toward the Ph.D. degree in photogrammetry and remote sensing from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan.

His main research interests include ship detection of SAR image and preprocessing of Video-SAR.



Boyang Jiang is currently working toward the Ph.D. degree with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China.

His main research interests include geometry processing of spaceborne imagery and bundle block adjustment.