SAR Moving Target Segmentation and Removal Based on Deep Learning

Yifan Wu[1,2,4(✉)], Xiyu Qi[1,3,4], Lijia Huang[1,2], Bingchen Zhang[1,2], and Lili Yan[5]

[1] Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
[2] Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China
[3] Key Laboratory of Network Information System Technology (NIST), Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China
[4] School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100190, China
wuyifan20@mails.ucas.ac.cn
[5] China Centre for Resources Satellite Data and Application, Beijing 100094, China

Abstract. Synthetic Aperture Radar (SAR) stands as an integral part of advanced remote sensing technology. Nevertheless, practical applications experience inevitable disturbances from moving target noise, compromising both image integrity and target detection performance. This paper introduces a pioneering approach reliant on deep learning principles for the elimination of moving target noise within SAR imaging. Firstly, we use the Back-Projection (BP) algorithm to form the foundational images from echo signals. Employing the Unet network, we subsequently acquire a segmentation map of the moving target noise. By subtracting this segmented noise map from the original image, we succeed in the effective erasure of moving targets, yielding a resultant image devoid of moving target noise. Experimental validation demonstrates that our method can effectively remove moving target noise and yield SAR images that solely contain static scenes.

Keywords: Synthetic Aperture Radar (SAR) · Moving target segmentation · Deep learning

1 Introduction

Synthetic Aperture Radar (SAR) is a high-resolution imaging radar. The basic principle is to use the relative motion of the carrier platform and the target to

Y. Wu and X. Qi—Contributed equally to this work and should be considered co-first authors.

W. Yu and L. Xuan (Eds.): MLICOM 2023, LNICST 518, pp. 107–118, 2024.
https://doi.org/10.1007/978-3-031-71716-1_9

synthesize the equivalent virtual long antenna aperture, to obtain azimuth high-resolution radar images. Compared with optical sensors, SAR has all-day and all-weather working capability. It is widely used in the fields of remote sensing of the earth, resource survey, disaster forecasting, and military reconnaissance. SAR imaging algorithms have formed a series of mature algorithms, mainly consisting of three categories: time-domain algorithms, frequency-domain algorithms, and interpolation algorithms [18,23,25]. The back-projection algorithm (BP) and the fast back-projection algorithm (FBP) are the classical algorithms in the time domain [19,20]. Theoretically, time-domain algorithms impose no constraints on the motion trajectory of the radar platform, thereby accommodating a variety of observational modes [15–17,21]. They also provide the flexibility to adjust the size of the output scene [10,12,22].

Synthetic Aperture Radar (SAR) was designed primarily for static scene imaging. Consequently, moving targets are significantly susceptible to Doppler shifts, causing target defocusing along the azimuth direction [2]. Rany was the first to examine the characterization of moving target signals in SAR images [13]. As demonstrated in the same study, moving targets frequently display shifts and/or smearing effects in SAR images. Notably, when SAR observes over extended periods, moving targets can generate long and intricate motion trajectories. In realistic scenarios, the occurrence of moving target noise often results in image quality degradation, compromising the accuracy of subsequent processing and analysis. Therefore, the development of an effective algorithm to eliminate moving target noise and produce high-quality aperture images is necessary.

Numerous processing methods for the moving target noise phenomenon in Synthetic Aperture Radar (SAR) have been proposed by scholars across the globe. For multi-channel SAR systems, several methods for moving target feature removal exist, which employ the unique spatiotemporal characteristics of moving targets to attenuate moving target signals via filtering, while retaining the static scene signals [7,14]. The method was first proposed by Schulz in 1986. He later provided the basic framework of this method in [14]. In 2016, Li et al. extended this method to high-resolution wide-swath SAR systems [7]. Nonetheless, these methods necessitate a more sophisticated SAR system design and rely on multiple channels for effective implementation. For Video SAR, experts globally have proposed some moving target detection methods [3,4,24]. The study specified in [3] employed the Hough transform to identify moving targets demonstrating stripe shapes within SAR imagery. Meanwhile, the research in [24] used a threshold segmentation algorithm to extract moving targets by adjusting the background of sequential images. Finally, the work presented in [4] extended the use of the coherent transform detection method to video SAR. However, many of these research methods depend on the high degree of image registration in SAR, leading to an increase in false alarms if not effectively registered. Furthermore, some of these methods only achieve moving target detection without accomplishing moving target removal.

Convolutional Neural Network (CNN), as a classical algorithm in the field of machine learning image research, has received continuous attention. [27]. They have shown prominence in the field of remote sensing with applications ranging

from terrain surface classification [6,26], aerial scene classification [11], to the detection of terrain changes over time by SAR and optical satellite sensors [8,9], and the segmentation of objects from airborne imagery [1,5]. Nonetheless, no existing studies have yet attempted to address the challenge of removing moving target noise from SAR images, which continues to be an arduous task that is yet to be explored.

To address the challenges of SAR moving target noise elimination, we introduce a unique deep learning-based solution. Our strategy starts with the generation of raw images using the BP algorithm, which itself contains moving target noise. To effectively eliminate this noise, we use the Unet network. For noise segmentation, we encode the original image. The subsequent decoding step generates a segmentation map which effectively divides the noise. After noise segmentation, we subtract the original image from the segmented image to eliminate moving targets. Finally, the SAR image without moving target noise is generated.

2 Methodology

In this paper, we propose a deep-learning-based method to address the challenge of moving target noise removal, as illustrated in Fig. 1. Initially, we employ the Back-Projection (BP) algorithm to process radar echo signals, generating an original image with moving targets. Subsequently, to achieve moving target segmentation (MTS) in the image, we use the Unet network in deep learning to analyze and partition the noise induced by moving targets. Following this, we acquire the noise-free image by moving target removal (MTR).

2.1 Moving Target Imaging Analysis Based on BPA

The BP algorithm stands as a major pillar in the realm of complex target scene imaging processing. Its ability to account for the spatial distance between each point within the scenario and the radar sets the stage for its superior handling capabilities for situations involving intricate targets and backgrounds. In the wide swath of radar imaging, the target scene often encompasses a variety of components, ranging from static, stationary structures to intricate and dynamic objects. These components, each with their unique spatial coordinates and signal properties, add layers of complexity to the imaging process.

The essence of the BP imaging algorithm lies in the process of reprojecting the echo signals. This means that regardless of the complexity of the targets or the cluttered nature of the background, the BP algorithm can model the physics of wave propagation and return for each point in the scenery. Specifically, regarding each point that frames the target scene, the algorithm calculates the respective ranges to the radar. By having this range detail incorporated into the calculations, the algorithm ensures accurate image construction, factoring in the individual contributions of complex targets. As the algorithm operates through this series of projections, it ultimately constructs a comprehensive imaging outcome of the target scene.

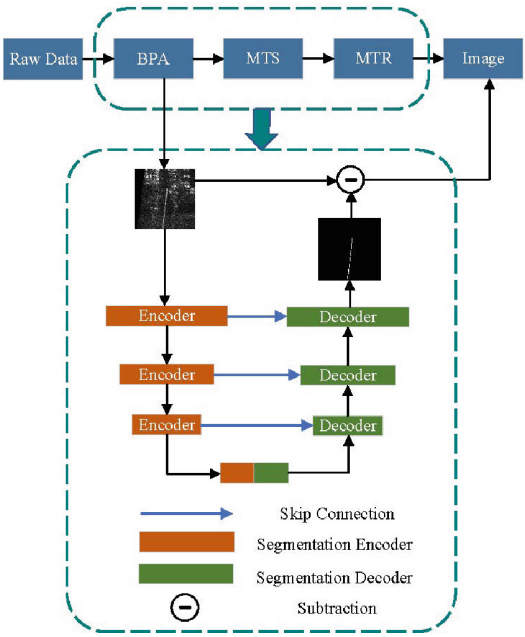

Fig. 1. The overview of our proposed SAR moving target segmentation and removal framework based on Deep Learning. The BP algorithm was utilized to image the echo signal and generate an image containing the moving target. Using the Unet network, the segmentation map is obtained for the motion target noise. Subtraction is done between the original image and the segmentation map so as to remove the moving target. Finally, the image without moving target noise is obtained.

First, range compression is applied to the demodulated SAR echo signal. After demodulation and range compression, the received signal at target $P(i,j)$ in the scene can be expressed as:

$$S_{ij}(\tau,\eta) = w(\eta)sinc\Big(B\Big(\tau - \frac{2R_{ij}(\eta)}{c}\Big)\Big)exp\Big(-j\frac{4\pi}{\lambda}R_{ij}(\eta)\Big) \tag{1}$$

where τ is the fast time, η is the slow time, B is the signal bandwidth, c is the speed of light, λ is the wavelength, $w(\eta)$ is the azimuth window, and $R_{ij}(\eta)$ is the instantaneous slant range of grid point $P(i,j)$ and can be obtained according to the Eq. (2).

$$R_\eta(i,j) = |P_{ij}Q\eta| = \sqrt{\big(X_\eta - X(i,j)\big)^2 + \big(Y_\eta - Y(i,j)\big)^2 + Z_\eta^2} \tag{2}$$

where the coordinates of the aperture point Q at azimuth time η are(X_η, Y_η, Z_η), and the coordinates of the image grid point are $\big(X(i,j), Y(i,j), 0\big)$.

Based on the slant range between the radar and each grid point at every azimuth moment (time of pulse transmitting), the two-way time delay can be

calculated. Then, the corresponding echo value for grid point $P(i,j)$ can be retrieved and represented as S_{ij}. Then, phase compensation is applied to the data at each azimuth moment, and the results are added together. The final focused BP image can be expressed as:

$$I(i,j) = \sum_{\eta} exp\Big\{ j\frac{4\pi}{\lambda} R_{ij}(\eta) \Big\} \cdot S_{ij}(\eta) \tag{3}$$

where the phase of the delay range in the time domain is compensated by $\phi = -4\pi R_{ij}(\eta)/\lambda$, and the compensation function can be expressed as follows:

$$H = exp\big\{ j4\pi R_{ij}(\eta)/\lambda \big\}. \tag{4}$$

The BP algorithm is a computationally intensive algorithm, and its major disadvantage is its high computational complexity and time cost, which leads to significant demands for computing resources and processing speed. To accelerate the running speed of the BP algorithm and improve its efficiency, a parallel computing method can be used to enhance the calculation speed. By utilizing technologies such as GPUs or distributed computing platforms, the data can be segmented into multiple sub-tasks for parallel computing, thereby speeding up the BP algorithm's processing time.

In SAR imaging systems, the process of image formation is dynamically influenced by the movement of the antenna or sensor platform. Within a target scene, stationary objects depicted through the consistent motion of the antenna provide the sharp and well-resolved image. However, complexities arise when moving targets are introduced into the scene, resulting in potential defocusing effects during the imaging process due to the differential rates of motion.

SAR imaging, by its design, is particularly adept at capturing stationary environments. This is primarily enabled by the sequential sampling and integration of radar backscatter signatures over the sensor's continuous movement, creating a synthetic aperture. However, if the target within the scene is not stationary, it can lead to an inconsistency in these collected signatures, thereby creating a defocusing effect that disrupts the imaging process.

It is crucial to understand that the position of a moving target fluctuates during the acquisition time of each sub-aperture image. This positional variance is not only a natural consequence of the target's own motion but is further compounded by the ongoing motion of the SAR platform. This dynamic interaction leads to a temporal misalignment of the acquired radar backscatter signatures.

Consequently, the presence of these moving targets can create a defocusing effect along the azimuth - the horizontal plane running perpendicular to the radar's line of sight. This unanticipated defocusing effect can deteriorate the image quality, particularly impeding the accurate representation of stationary targets in the scene.

Although SAR imaging systems are exceptionally effective tools for the visualization of stationary environmental features, the presence of moving targets introduces unique challenges. These challenges, deriving from the azimuthal

defocusing effect, necessitate the development of advanced algorithms and signal processing techniques to maintain imaging precision and thus, enhance the reliability of the SAR-derived products.

2.2 Moving Target Segmentation Based on UNet

In our pursuit to isolate moving targets within the context of SAR imagery composed mainly of stationary objects, we employ the framework of deep learning with an emphasis on the Unet network to accomplish moving target segmentation (MTS). This approach introduces a unique and effective solution to the issues associated with moving target noise in SAR imaging.

Deep learning, a subset of machine learning, is aptly suited to this undertaking due to its ability to learn and extract high-level features from data. Specifically, we make use of the Unet network, a convolutional neural network structure famed for its proficiency in image segmentation tasks.

UNet is a deep neural network specifically designed for image segmentation tasks. It consists of an encoder and a decoder, which together perform the downsampling and upsampling operations, respectively. The downsampling operation is responsible for extracting high-dimensional features from the input image, while the upsampling operation aims to map the features back to the original image size to generate a probability segmentation map for each pixel.

The UNet architecture can be visualized as shown in the following diagram:

The encoder part of UNet comprises a series of convolutional and pooling layers. The initial convolutional layers capture local feature maps that help compute neighbor relationships between the pixels, aiding in the downsampling process. Subsequent convolutional and pooling layers allow the network to extract higher-dimensional features, such as edges and object shapes. These features are crucial for accurate segmentation.

Convolution operation in the i-th layer in the encoder part can be expressed as:

$$h_i = ReLU(Conv(x_i, W_i) + b_i) \tag{5}$$

where, x_i refers to the input of the i-th layer, W_i and b_i are the convolutional kernel weights and biases of the i-th layer, $Conv$ represents the convolution operation, and $ReLU$ denotes the rectified linear unit activation function.

Pooling operation in the i-th layer in the encoder part can be expressed as:

$$p_i = MaxPool(h_i) \tag{6}$$

where, $MaxPool$ represents the max pooling operation, typically using a 2×2 pooling window.

The decoder part of UNet consists of upsampling layers and transposed convolutional layers. Upsampling operation in the i-th layer in the decoder part can be expressed as:

$$u_i = Upsample(p_i) \tag{7}$$

where $Upsample$ represents the convolution operation. $Upsample$ is usually performed using interpolation or transposed convolution techniques to upsample the feature maps to the original image size.

The key feature of UNet is the skip connections. At each step in the decoder, skip connections are introduced, which concatenate the feature maps from the corresponding encoder step. These connections allow the network to retain detailed information from the encoder while also utilizing the upsampled features from the decoder. By doing this, UNet can effectively combine both low-level and high-level features, resulting in improved segmentation results. Skip connection in the i-th layer in the decoder part can be expressed as:

$$c_i = Concat(h_i, u_i) \tag{8}$$

where, $Concat$ denotes concatenating the feature maps along the depth dimension.

To perform segmentation, UNet uses the softmax function and computes the logarithmic loss scores. During training, optimization algorithms like stochastic gradient descent (SGD) or adaptive algorithms are commonly employed to update the network weights. Regularization techniques, such as dropout or batch normalization, can also be applied to avoid overfitting the model to the training data. Output segmentation result operation can be expressed as:

$$o = Softmax(Conv(c_i, W_o) + b_o) \tag{9}$$

where, W_o and b_o are the convolutional kernel weights and biases of the output layer, and $Softmax$ is the activation function used to compute probabilities.

In the context of our study, the Unet network is utilized to analyze and partition noise attributed to the presence of moving targets within the SAR imagery, a crucial task in moving target segmentation. The implementation of the Unet network allows for the effective separation of signal noise associated with moving targets from the primary stationary object data within the image.

2.3 Moving Target Removal

In order to extract static scenes from the presence of moving targets in SAR images, we propose a method of subtracting noise segmentation plots from images. The image contains comprehensive information about the scene, including stationary and moving constituents. The noise-segmented map, on the other hand, isolates noise components, primarily induced by the moving targets.

The procedure of subtraction is designed to eliminate the components associated with the moving targets. This is executed by directly subtracting the segmented map, which essentially represents the moving targets and associated noise, from the original image. This subtraction operation eliminates the components common to both- the primary moving targets and the associated noise while preserving the other elements, which essentially represent the static scene within the SAR image. The subtraction operation can be expressed as:

$$\hat{I}(i,j) = \bar{I}(i,j) - o(i,j) \tag{10}$$

where$\hat{I}(i,j)$ represents the images without moving target noise, $\bar{I}(i,j)$ denotes the raw images containing the moving target, o represents the noise-segmented map, and (i,j) denotes the position of pixel points.

The result of this subtraction operation is an image that contains information solely about the static scene, clearing the view of any defocusing effects caused by the presence of moving targets. The moving targets and any associated noise or distortion are effectively erased, leaving behind an optimized image representing the static landscape.

3 Experiments

In this section, the experimental results of the proposed moving target removal method are presented.

3.1 Datasets

The experimental validation utilized the Gotcha dataset, which comprises a substantial volume of high-resolution SAR echo data. To generate SAR images, the raw data first needs to undergo imaging using a SAR imaging algorithm. Each image within the dataset contains discernible targets and complex background information. In addition to various vehicles and civilian facilities, the SAR Gotcha dataset encompasses a diverse array of targets imaged from different angles, resulting in a considerable number of distinct SAR images. This dataset serves a vital purpose in numerous critical SAR image processing tasks, including target detection, identification, and classification. Consequently, it holds immense value for researchers and engineers involved in developing and studying SAR image processing algorithms. This dataset contains some moving targets, to which we have added more moving targets. We conduct research and experiments on the moving targets in the dataset to verify the effectiveness of the proposed method.

3.2 Implementation Details

To increase the diversity of the image background and provide a wide range of training sets for subsequent network training, we select different regions for imaging scenes. This process also serves to enhance the robustness of the network. The imaging grid interval is set at $0.4m \times 0.4m$ with a grid size of 256×256, resulting in an image composed of 256×256 pixel points. We employ the BP algorithm for the imaging process, interpolating the range direction ten times to improve imaging accuracy.

In cases where no moving targets are present in the imaging region, a moving target will be artificially introduced into the image. Simultaneously, a binary segmentation map of the moving target is acquired for subsequent network training. A group is formed consisting of the original image and its corresponding segmentation map. Within each group, the image datasets are divided into separate

training and test sets, which are utilized for network training and testing purposes, respectively.

To perform our experiments, we utilize UNet with 6x downsampling. Throughout the training phase, we employ the PyTorch framework and train the model end-to-end using the SGD optimizer. A mini-batch size of 4 is utilized, and the initial learning rate is set to 0.0001. Data augmentation is applied by randomly performing vertical or horizontal flipping, as well as rotation operations on the input images, resulting in a size of 256×256 pixels. All experiments are conducted on NVIDIA GeForce RTX 3080 Ti GPUs, and the training process is carried out for 30,000 epochs. Optimal performance in terms of accuracy and efficiency is achieved in our experiments by leveraging suitable pretrained backbones and finely tuning hyperparameters.

3.3 Main Results

Our experimental results are depicted in Fig. 2. The segmentation effect of the moving target in two different scenes is observed. Scene I contains complex feature targets. Scene II contains moving targets with a more pronounced contrast to the background. Plots in the first and third columns depict the moving targets, while the second and fourth columns display moving target segmentation plots. Our experimental results demonstrate that we can achieve better segmentation of the moving targets, whether in a simple scene or a complex scene. We subtracted the corresponding moving target noise segmentation maps from the original images. This process resulted in obtaining the images after the removal of the moving target noise.

3.4 Performance Analysis

The strength of the BP algorithm lies in its attention to detail, specifically, its computational ability to consider the specific range between each point within the target scene and the radar. The use of the Unet network within a deep learning framework presents a powerful and effective tool for moving target segmentation within SAR imagery. The subtraction methodology offers a streamlined and potent approach for isolating static target data from Synthetic Aperture Radar (SAR) imagery. The proposed method effectively addresses the intrinsic complications induced by moving targets and concomitant noise within the image information. The derivation of static scene images subsequent to the removal of the moving target affirms the competencies of this approach. This proposed method underscores the integration of deep learning methodologies with traditional SAR imaging operations to contribute to the advancement of radar image processing techniques.

However, the proposed method has some limitations. Currently, we have only validated it on simulated moving targets and it still needs to be evaluated on actual moving targets. Future efforts will be geared towards enhancing the proposed method.

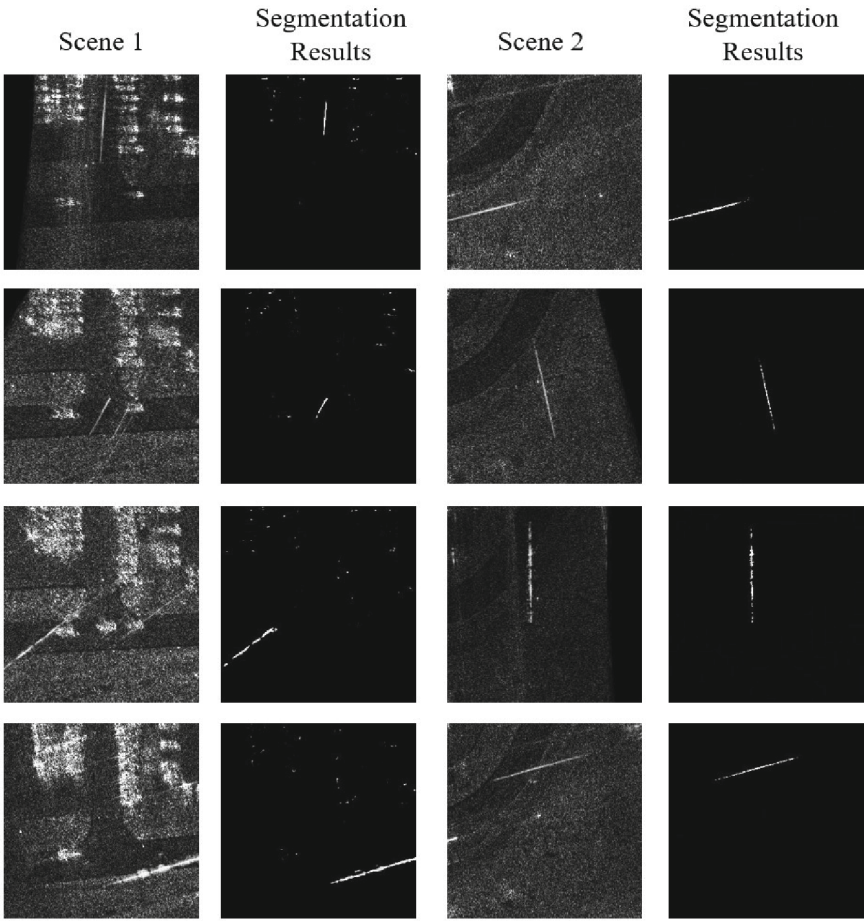

Fig. 2. Moving target segmentations of images for two different scenes.

4 Conclusion

In this study, we propose a novel method for removing moving target noise in SAR images based on deep learning. Firstly, we implement the BP algorithm to construct preliminary images. To further enhance the noise removal outcome, we introduce the Unet network. The Unet network provides robust image processing capabilities and effective feature extraction. We use the subtraction methodology to isolate static target data from SAR images. Through the removal of moving target noise, we successfully obtain SAR images without moving targets. Experimental results demonstrate that our method effectively reduces moving target noise in SAR images, leading to improved image clarity and detail. The significance of this research lies in the application of deep learning techniques to SAR

image processing, the introduction of an innovative method for removing moving target noise, and the experimental validation of its efficacy.

References

1. Bianchi, F.M., Espeseth, M.M., Borch, N.: Large-scale detection and categorization of oil spills from SAR images with deep learning. Remote Sens. **12**(14), 2260 (2020)
2. Gou, L., Li, Y., Zhu, D., et al.: A real-time algorithm for circular video SAR imaging based on GPU. Radar Sci. Technol. **17**(5), 550–556 (2019)
3. Jin, L., Wu, F., Yang, Y., Li, D., Sun, J.: Study of airborne video synthetic aperture radar. J. Microwaves **36**(1), 45–48 (2020)
4. Jinshan, D.: Focusing algorithms and moving target detection based on video SAR **9**(2), 321–334 (2020)
5. Kampffmeyer, M., Salberg, A.B., Jenssen, R.: Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1–9 (2016)
6. Kampffmeyer, M., Salberg, A.B., Jenssen, R.: Urban land cover classification with missing data modalities using deep convolutional neural networks. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. **11**(6), 1758–1768 (2018)
7. Li, X., Xing, M., Xia, X.G., Sun, G.C., Liang, Y., Bao, Z.: Simultaneous stationary scene imaging and ground moving target indication for high-resolution wide-swath SAR system. IEEE Trans. Geosci. Remote Sens. **54**(7), 4224–4239 (2016)
8. Luppino, L.T., Bianchi, F.M., Moser, G., Anfinsen, S.N.: Remote sensing image regression for heterogeneous change detection. In: 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1–6. IEEE (2018)
9. Luppino, L.T., Bianchi, F.M., Moser, G., Anfinsen, S.N.: Unsupervised image regression for heterogeneous change detection. arXiv preprint arXiv:1909.05948 (2019)
10. Meng, D., Hu, D., Ding, C.: Precise focusing of airborne SAR data with wide apertures large trajectory deviations: a chirp modulated back-projection approach. IEEE Trans. Geosci. Remote Sens. **53**(5), 2510–2519 (2014)
11. Penatti, O.A., Nogueira, K., Dos Santos, J.A.: Do deep features generalize from everyday objects to remote sensing and aerial scenes domains? In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 44–51 (2015)
12. Ran, L., Liu, Z., Li, T., Xie, R., Zhang, L.: An adaptive fast factorized back-projection algorithm with integrated target detection technique for high-resolution and high-squint spotlight SAR imagery. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. **11**(1), 171–183 (2017)
13. Raney, R.K.: Synthetic aperture imaging radar and moving targets. IEEE Trans. Aerosp. Electron. Syst. **3**, 499–505 (1971)
14. Schulz, F.: Suppressing moving target artifacts in multi-channel stripmap SAR images by Space-Doppler filtering. IEEE J. Sel. Topics Signal Process. **5**(3), 494–503 (2011)
15. Sun, G.C., Wu, Y., Yang, J., Xing, M., Bao, Z.: Full-aperture focusing of very high resolution spaceborne-squinted sliding spotlight SAR data. IEEE Trans. Geosci. Remote Sens. **55**(6), 3309–3321 (2017)

16. Sun, G.C., Xing, M., Xia, X.G., Wu, Y., Bao, Z.: Beam steering SAR data processing by a generalized PFA. IEEE Trans. Geosci. Remote Sens. **51**(8), 4366–4377 (2013)
17. Sun, G.C., et al.: Multichannel full-aperture azimuth processing for beam steering SAR. IEEE Trans. Geosci. Remote Sens. **51**(9), 4761–4778 (2013)
18. Sun, G., Jiang, X., Xing, M., Qiao, Z.J., Wu, Y., Bao, Z.: Focus improvement of highly squinted data based on azimuth nonlinear scaling. IEEE Trans. Geosci. Remote Sens. **49**(6), 2308–2322 (2011)
19. Ulander, L.M., Hellsten, H., Stenstrom, G.: Synthetic-aperture radar processing using fast factorized back-projection. IEEE Trans. Aerosp. Electron. Syst. **39**(3), 760–776 (2003)
20. Wu, Y., Huang, L., Zhang, B., Wang, X., Qi, X.: An accurate and efficient BP algorithm based on precise slant range model and rapid range history construction method for geo SAR. Remote Sens. **15**(21), 5191 (2023)
21. Zeng, L., Liang, Y., Xing, M., Huai, Y., Li, Z.: A novel motion compensation approach for airborne spotlight SAR of high-resolution and high-squint mode. IEEE Geosci. Remote Sens. Lett. **13**(3), 429–433 (2016)
22. Zhang, L., Li, H.l., Qiao, Z.J., Xing, M.D., Bao, Z.: Integrating autofocus techniques with fast factorized back-projection for high-resolution spotlight SAR imaging. IEEE Geosci. Remote Sens. Lett. **10**(6), 1394–1398 (2013)
23. Zhang, L., Qiao, Z.J., Xing, M., Li, Y., Bao, Z.: High-resolution ISAR imaging with sparse stepped-frequency waveforms. IEEE Trans. Geosci. Remote Sens. **49**(11), 4630–4651 (2011)
24. Zhang, Y., Zhu, D., Yu, X., Mao, X.: Approach to moving targets shadow detection for VideoSAR **39**(9), 2197–2202 (2017)
25. Zhang, Z., Xing, M., Ding, J., Bao, Z.: Focusing parallel bistatic SAR data using the analytic transfer function in the wavenumber domain. IEEE Trans. Geosci. Remote Sens. **45**(11), 3633–3645 (2007)
26. Zhou, Y., Wang, H., Xu, F., Jin, Y.Q.: Polarimetric SAR image classification using deep convolutional neural networks. IEEE Geosci. Remote Sens. Lett. **13**(12), 1935–1939 (2016)
27. Zhu, X.X., et al.: Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci. Remote Sens. Mag. **5**(4), 8–36 (2017)