A Novel Shadow and Layover Segmentation Network for Multi-Angle SAR Images Fusion

XUEWEI LI¹, GANG ZHANG², CANBIN YIN³, YUQUAN WU¹, AND XINGCHEN SHEN¹

¹Science and Technology on Integrated Information System Laboratory, Institute of Software, Chinese Academy of Sciences, Beijing 100190, China
²Unit 95865 of PLA, Beijing 102218, China
³Department of Aerospace Science and Technology, Space Engineering University, Beijing 100084, China

Corresponding author: Gang Zhang (gangzhang1989@126.com)

This work was supported in part by the Chinese Academy of Sciences under Grant E1YD5906 and Grant E2YD590901.

ABSTRACT
Shadow and layover are geometric distortion phenomena in side-view imaging synthetic aperture radar (SAR) systems, especially in mountainous areas and densely populated urban areas. The shadow can block the target of the observation area, making it impossible to obtain the scattering characteristics of the target. The layover causes phase distortion and alters target characteristics. Shadow and layover severely hinder the interpretation of SAR images. To confront the above problems, a multi-angle fusion algorithm based on unsupervised progressive segmentation network is proposed. Firstly, inspired by mega-constellations of low earth orbit, a spaceborne SAR collaborative observation model is proposed to generate multi-angle images of fluctuant terrain. Secondly, according to the difference of echoes in the shadow and layover regions, an unsupervised progressive segmentation network is designed to sequentially segment the shadow and layover regions. Finally, to improve the contrast and brightness of the fused SAR image, a single-scale weighted fusion algorithm is designed. Experiments were conducted using the simulated multi-angle SAR images. Compared with single-angle images, the accuracy of target detection and figure-of-merit of the fused SAR image are significantly higher than those of other methods.

INDEX TERMS
Shadow and layover, synthetic aperture radar, unsupervised progressive segmentation network, multi-angle fusion.

I. INTRODUCTION
Synthetic aperture radar (SAR) [1] has been widely used in various fields based on the ability to acquire high-resolution images nearly all the time and all weather conditions. With the development of high-resolution spaceborne SAR [2], high-resolution SAR data are becoming more abundant and easier to acquire. However, shadow and layover [3] caused by the side-look imaging mode seriously affect the application of SAR images. Shadow and layover are a kind of geometric distortion. The existence of the shadow can block some or all of the targets, resulting in blind spots [4]. The existence of the layover undermines the continuity of the interference phase, resulting in unavoidable errors of filtering and unwrapping in the data processing [5].

The shadow and layover mostly exist in mountainous areas and high buildings of SAR images. They can directly degrade the performance of SAR image interpretation, especially in the task of weak and small target detection. Many scholars have studied the detection of shadow and layover, such as coherence coefficient maps [6], the threshold segmentation method based on intensity maps [7], the edge sharpening method using image filtering [8] and multi-stages layover detection method [9]. However, all of these methods have their limitations. Due to the inherent random speckle noise in SAR images [10], it is difficult to extract the shadow and layover accurately by threshold segmentation or digital image processing.

In recent years, convolutional neural networks (CNN) in deep learning have developed rapidly in the field of computer vision. The CNN can extract hierarchical features from SAR images, which is of great significance for shadow and layover...
To fuse saliency features with deep CNN features without adding feature dimension and network parameters, a new saliency fusion method is proposed.

- To improve the brightness and contrast of the fused image, a single-scale weighted fusion (SSWF) algorithm is proposed to fuse segmented MA SAR images.
- The proposed method is validated by employing visual subjective results and the performance of target object detection. The experimental results show that the proposed methods significantly improve the target recognition accuracy.

The rest of this article is organized as follows. Section II briefly reviews the related works. Section III presents the detailed structure of the proposed method. Section IV introduces the comprehensive experiments, comparison results and analysis. Finally, Section V concludes this article.

II. RELATED WORKS

This section summarizes the previous works in the related field, mainly includes principles of shadow and layover, segmentation methods of shadow and layover regions and MA fusion algorithms.

A. PRINCIPLE OF SHADOW AND LAYOVER

Shadow and Layover, the geometry distortion of range compression, which are caused by the side-look imaging mode of SAR, has a serious impact on the broad application of SAR technology. Especially, shadow obscures the target and layover causes phase distortion and alters target characteristics in mountainous areas and high buildings.

1) PRINCIPLE OF SHADOW

When the back slope angle of the terrain $\alpha$ is larger than the slope angle $\theta$, radar signals are blocked by slope. Therefore the blocked area cannot reflect the signal, resulting in the shadow phenomenon of SAR image [21]. The principle of shadow in the SAR image is shown in Fig. 1, where $\beta$ means look angle of SAR systems. The area of $ABC$ represents the fluctuant terrain. The shadow regions will become blind spots of SAR observations, resulting in reducing the target detection accuracy.

![FIGURE 1. The principle of shadow in SAR images. (The region of BCD has no echo in the imaging plane.)](image)
2) PRINCIPLE OF LAYOVER
When the slope angle of the fluctuant terrain $\theta$ is larger than the look angle of the SAR $\beta$, the echo from the top of the slope is received earlier than that from the bottom. The top of the fluctuant terrain appears before the bottom in the SAR image, resulting in the layover phenomenon of targets [22]. The principle of the layover in SAR images is shown in Fig. 2. The layover region can cause the phase distortion and alters target characteristics, resulting in affecting SAR image applications.

B. SEGMENTATION METHODS OF SHADOW AND LAYOVER REGIONS
Recently, the deep learning CNN have developed rapidly in the domain of computer vision. The CNN can extract hierarchical features from SAR images, which is of great significance for shadow and layover detection. The deep CNN methods can solve the problems of artificial auxiliary requirements and threshold settings in traditional approaches. Robson et al. [23] proposed a rock glacier segmentation network combining deep CNN and object based image analysis methods to extract and segment rock glaciers (layover) in the remote sensing images. Tiwari et al. [12] designed the CLSTMISS to achieve the segmentation of phase-stable (homogeneous) regions and unstable (layover) regions in multi-temporal interferometric SAR (InSAR) images. Cai et al. [11] proposed the MF2AM-CNN for extracting layover regions. MF2AM-CNN adopted ResNet101, atrous spatial pyramid pooling (ASPP), semantic embedding module and multi-level feature fusion module to improve the effect of edge segmentation. Chen et al. [24] presented a complex-valued convolutional and multi-feature fusion network (CVCMFFNet) specifically for building semantic segmentation of InSAR images. CVCMFFNet can effectively segment the layover, shadow and background on both the simulated InSAR building images and the real airborne InSAR images.

Compared with optical images, SAR images are the grayscale distribution images, which lack details of targets shape and texture. Meanwhile, deep CNN models also need amount training samples. However the acquisition way of SAR images is simple and expensive. In addition, annotations of the shadow and layover still require prior knowledge of the terrain distribution. Therefore, in this article, according to the pixel difference of the shadow and layover regions in SAR images, an novel unsupervised progressive segmentation network (UPSNet) is proposed.

C. MULTI-ANGLE FUSION ALGORITHMS OF SAR IMAGES
The multi-angle (MA) fusion algorithms of SAR image is to fuse the complementarity information from multiple look angles. The fused image can significantly improve the description ability of the target, resulting in improving the accuracy of SAR applications. Linear average and maximum fusion algorithms [14] are the simplest M2O fusion algorithms, which have the advantages of simple implementation and fast computation. However, the fused SAR image becomes blurry. Meanwhile, shadow and layover regions in SAR images also have a great impact on fusion results. To confront the this issues, Zhu et al. [15] proposed a MA fusion algorithm based on visibility classification of non-layover regions. Although the algorithm effectively reduces the influence of the layover regions, there are very few studies on MA image fusion algorithms. However, MA SAR images mainly were applied to automatic target recognition (ATR). Huan et al. [18] first used feature-level (principal component analysis, wavelet transform and support vector machines) and decision-level (Bayesian algorithm) fusion strategies for ATR. Zhang et al. [17] proposed a bidirectional long short time memory (Bi-LSTM) recurrent neural network, which utilizes the Bi-LSTM to fuse handcraft features of MA SAR images. Zhao et al. [19] designed a ATR CNN, which proposed a random Fourier transform layer to fuse deep features of MA SAR images. Zhao et al. [19] used bidirectional gated recurrent units to fuse MA deep features by extracting from EfficientNet [20].

In the existing SAR systems, it is very difficult to obtain MA SAR images of the fluctuant terrain at the same time. The MA SAR images at different times have the difficulty of de-correlation. Meanwhile, due to the lack of sufficient annotated training samples, deep CNN of M2O fusion methods have not been studied. Therefore, to ensure the speed of fusion, the SSWF is proposed. SSWF can protect the brightness and contrast of fused images.

III. PROPOSED METHOD
The detailed framework of MAFUPSNet is displayed in Fig. 3, where the SAR-CO model means cooperative observation model. MA images represents the multi-angle images. UPSNet is the unsupervised progressive segmentation network and SSWF is the single-scale weighted fusion algorithm.

The MAFUPSNet can reduce the impact of the shadow and layover by fusing MA images, resulting in improving the target detection performance. Firstly, the SAR-CO model is
used to generate MA SAR images of fluctuant terrain at the same time, which reduces the difficulty of temporal decorrelation of each images. Secondly, according to the difference of pixels in shadow and layover regions, UPSNet sequentially completes the segmentation of shadow and layover regions. Finally, according to the segmented MA images, the single-scale weighted fusion algorithm is designed to protect the contrast and brightness of fused SAR image.

A. SAR COOPERATIVE OBSERVATION MODEL
To obtain MA SAR images of fluctuant terrain at the same time, inspired by mega-constellations of low earth orbit, a spacebone SAR cooperative observation (SAR-CO) model is constructed. The geometric principle of the SAR-CO model is shown in Fig. 4, which $N$ represents the number of SAR satellites. $O$ and $O'$ are the centers of the target and satellite, respectively. $H_{\text{star}}$ and $V_{\text{star}}$ represent the height and speed of the satellite, respectively. $A_p$ is the area of fluctuant terrain. $R_{\text{np}}$ represents the distance between SAR satellite and target. Each SAR satellite of the SAR-CO model works independently.

The signal of SAR-CO satellites is the discrete code signal [25], rather than chirp signal. There are two reasons. Firstly, the sidelobe energy distribution of the chirp signal is wider than discrete code signal, resulting in the strong scattering target covers the adjacent weak scattering target. Secondly, the SAR-CO model is composed of multiple SAR satellites working together. The coherent or incoherent accumulation of different signal energies can exacerbate the obscuring problem. The discrete code signal is a kind of quadrature coded signal with excellent auto-correlation and cross-correlation, which can solve the interference of echo.

In the SAR-CO model, the discrete code signal ($Q$) and carrier frequencies ($F$) are defined as:

$$
Q = \{q_1, q_2, \ldots, q_N\} \\
F = \{f_1, f_2, \ldots, f_N\},
$$

where $q_i$ and $f_i$ are the signal and carrier frequency of the $i$-th SAR satellite, respectively. For example, the signal of the $\text{Sat}_n$ is $s_{n,T}(t)$, which is defined as:

$$
s_{n,T}(t) = e^{j2\pi f_n t} \cdot \sum_{l=1}^{L} [P_l(t) \cdot e^{j(2\pi f_l t + \theta_l)}],
$$

where $L$ represents number of sub-pulses and $\theta_l$ is the phase of the $l$-th sub-pulse. $P_l(t)$ is defined as:

$$
P_l(t) = A_l \cdot \text{rect}(t/T_l),
$$

where $A_l$ represents a constant and takes the value 0 or 1. $T_l$ represents the $l$-th sub-pulse period. $\text{rect}(\cdot)$ stands for the
rectangle function, which can be written as:

\[
\text{rect}(t) = \begin{cases} 
1, & \text{if } 0 < t < T_1, \\
0, & \text{otherwise.}
\end{cases}
\] (4)

After the radar signal is reflected by the fluctuant terrain, the echo received by \( S_{at_n} \). It is expressed as:

\[
s_{n,R}(t) = \sigma_p \cdot e^{2\pi f_0(t-\tau)} \cdot \sum_{l=1}^{L} [P_l(t-\tau) \cdot e^{j(2\pi f_l(t-\tau)+\theta_l)}],
\] (5)

where \( \tau \) represents delay time of signal. And \( \tau \) can be written as:

\[
\tau = \frac{2R_{np}}{c},
\] (6)

where \( c \) is the speed of light. Assuming that the number of \( A_p \) has \( P \) targets, the echoes of the fluctuant terrain received by \( S_{at_n} \) is written as:

\[
S(t) = \sum_{p=1}^{P} s_{n,R}(t).
\] (7)

The echo of the fluctuant terrain is processed by the fast back projection (FBP) algorithm [26] and the single-look complex SAR image is obtained. The FBP algorithm has two advantages. First, the FBP algorithm is a fast imaging algorithm that can process radar echoes quickly. Second, since the FBP algorithm can project multiple SAR images into the same coordinate system, it can avoid the process of image registration, which can reduce difficult of MA SAR image fusion.

B. UNSUPERVISED PROGRESSIVE SEGMENTATION NETWORK

Soergel et al. [7] had analyzed echo intensities of shadow and layover regions. The echo intensity difference between the shadow region and the homogeneous region is about 100 times. The echo intensity of the layover region is only about 3 times higher than that of the homogeneous region. Meanwhile, due to the lack annotations of shadow and layover areas, inspired by unsupervised learning, an unsupervised progressive segmentation network (UPSNet) is designed to sequential segment shadow and layover areas.

The UPSNet includes the following three stages: preprocessing, unsupervised segmentation network (Un-SegNet) and post-processing. The detailed structure of UPSNet is shown in Fig. 5, where \( x \) and \( y \) are the SAR intensity image and the segmented image, respectively. The \( x_d, x_l, x_m \) and \( x_s \) represent the degraded image, the output of Un-SegNet, the upsampled mask and removed shadow image, respectively. The \( \text{Upsample} \) represents image upsampling, and the upsampling mode is the nearest. Compared with transposed convolution, the nearest neighbor upsampling can avoid the problem of blurred boundary. \( \text{CB}_1-\text{CB}_N \) and \( \text{Argmax} \) are convolutional blocks and classification layer of unsupervised segmentation network (Un-SegNet). The loss function of Un-SegNet is softmax loss. The \( \text{SpSA} \) is is a superpixel segmentation algorithm proposed by P. F. Felzenszwalb and D. P. Huttenlocher, which called Felz algorithm [27]. The Felz algorithm is a graph-based greedy clustering algorithm, which uses a weighted graph to represent images. The Felz
algorithm has the advantages of simple implementation and high speed. Compared with the simple linear iterative clustering (SLIC) algorithm [28], the Felz algorithm can more accurately separate the boundaries of the shadow and other regions. Pseudo Label represents the pseudo-label of the input image that generated by using the Felz algorithm.

The pre-processing is used to remove speckle noise and image degradation. The pre-processing includes the following four parts: despeckling, multi-look processing, Gaussian smoothing filtering and morphological filtering. The despeckling can reduce the impact of noise on segmentation performance. The multi-look processing is applied to reduce images size and the Gaussian smoothing filtering is used to reduce the difference between adjacent pixels of the input image. The morphological filtering is used to convert discontinuous scattered points into continuous regions. The Un-SegNet stage consists of convolutional blocks and a classification layer. The convolutional blocks are used to extract deep CNN features from SAR images. The classification layer is used to classify each pixel of the image. The post-processing stage, which consists of adaptive threshold segmentation (ATS) algorithm [28], the Felz algorithm can more high speed. Compared with the simple linear iterative clustering layer is used to classify each pixel of the image. The classification process is applied to reduce the contrast and brightness. The reason is that there are too smooth, resulting in destroying the contrast and brightness. But the fused SAR image is too smooth, resulting in destroying the contrast and brightness. The multi-scale fusion algorithms used multiple scales of image information to improve image details (i.e. edges and textures). However the multi-scale fusion algorithms require accurate decomposition level or direction. This is very difficult in practical application. In addition, the segmented MA SAR images is composed of reserved (non-shadow and non-layover) regions and 0-pixel (shadow and layover) regions. The 0-pixel regions exacerbate the above problems.

Therefore, a single-scale weighted fusion (SSWF) algorithm is designed for fusing segmented MA SAR images. The input of SSWF is segmented MA images, which come from the UPSNet. The detailed flow of SSWF algorithm is shown in Algorithm 1, where UPSNet(·) represent the unsupervised progressive segmentation network. H and W are the height and weight of image x, respectively. \( f_m(\cdot) \), \( f_f(\cdot) \) and \( fuse(\cdot) \) represent remove, find and fuse functions, respectively. The \( f_f(\cdot) \) is used to find 0-pixel in set P and the \( f_m(\cdot) \) is used to remove 0-pixel element. The \( fuse(\cdot) \) is defined as:

\[
fuse(P) = \sum_{m=1}^{M} \frac{p_m}{\sum_{m=1}^{M} \exp(p_m)},
\]

where \( M \) represents the number of elements in the pixel set P, which has removed 0-pixel element.

Compared with previous single-scale fusion algorithms, the proposed SSWF algorithm not only has the advantages of simple design and low computational complexity, but also can protect contrast and brightness. The reason is that there are a large number of 0-pixel regions in the segmented images. They will reduce the weight of non-zero regions, resulting in changing the brightness and contrast of the fused image. Compared with the previous multi-scale fusion algorithms, the proposed SSWF algorithm can significantly improve the fusion efficiency and reduce the computational complexity.

### IV. EXPERIMENTAL RESULTS

#### A. IMPLEMENTATION DETAILS

1) DATASETS

To generate the MA SAR images of real fluctuant terrain, we use 8 SAR satellites to construct the SAR-CO model. The detailed parameters of the SAR-CO model are listed in Table 1, where \( f_c, B, \Delta f, f_j \) represent carrier frequency, bandwidth, frequency interval and sampling frequency, respectively. \( H_{sat} \) and \( V_{sat} \) are the altitude and velocity of the satellite, respectively. The \( SR \) is the spatial resolution. \( \theta_l \) and \( \theta_r \) mean the look angle and pitch angle of the SAR, respectively.
Firstly, we use fractal Brownian surfaces [29] to generate fluctuant terrain. Meanwhile, 200 targets are added to fluctuant terrain for verifying the performance of the proposed method. The 3D and 2D structure of simulated fluctuant terrain and 200 targets are displayed in Fig. 6. Then, according to the simulated fluctuant terrain, the echos of SAR-CO model are analyzed and calculated. Finally, the MA echos are processed by using FBP algorithm. The obtained MA SAR image of the simulated fluctuant terrain are shown in Fig. 7. The look angles of the 8 SAR satellites are shown in Table 1. The shadow and layover regions in MA SAR images with different look angles are different, resulting in demonstrating the sensitivity of the SAR system to the look angle. Meanwhile, blocked/covered targets in the fluctuant terrain are also different. Therefore, the influence of shadow and layover phenomenon can be significantly suppressed by fusing MA images under different look angles.

### 2) EVALUATION METRICS

To verify that the proposed method, target detection accuracy $p_d$, missed detection rate $p_u$ and figure-of-merit ($FoM$) are selected as evaluation metrics. The $p_d$, $p_u$ and $FoM$ are defined as:

$$ p_d = \frac{N_{td}}{N_{gt}}, \quad (10) $$

$$ p_u = \frac{N_{ud}}{N_{gt}}, \quad (11) $$

$$ FoM = \frac{N_{td}}{N_{fd} + N_{gt}}, \quad (12) $$

where $N_{td}$ is the number of the detected targets and $N_{gt}$ denotes the number of all targets. $N_{fd}$ denotes the number of false detected targets and $N_{ud}$ is the number of missing detected targets.

### 3) COMPARISON METHODS

K-means [30], superpixel-based fast fuzzy C-means clustering (SFFCM) [31] and maximum entropy threshold (MET) segmentation algorithms [32] are selected as comparison algorithms. The K-means image segmentation algorithm uses K-Means clustering algorithm to find similar pixels in the

| Parameters | Value |
|------------|-------|
| $f_c$      | 350MHz |
| $B$        | 616MHz |
| $\Delta f$ | 4MHz  |
| $f_s$      | 819MHz |
| $H_{sat}$  | 1100km |
| $V_{sat}$  | 7.53km/s |
| $A_p$      | 200 x 200 |
| $SR$       | 0.3m x 0.3m |

**TABLE 1.** The detailed parameters of cooperative observation SAR satellites.
image according to the divided K centers. The cluster results are the similar distances pixels set with cluster center. The metric of K-means is the Euclidean distance. The SFFCM segmentation algorithm uses the multi-scale morphological gradient reconstruction algorithm and histogram clustering to complete the image segmentation. The MET segmentation algorithm is a threshold selection method based on image entropy, and the image is divided into foreground and background by the size of the selected threshold.

The linear Average (Avg), linear Maximum (Max), Laplacian pyramid (LP) [33], non-subsampled surfacelet transform (NSCT) [34], curvelet transform (CT) [35], dual tree complex wavelet transform [36] and pixel significance discrete wavelet transform [37] were chosen for the comparison fusion algorithms of the MA SAR images. The Avg fusion method takes the average pixel value as the pixel value of the fused image. The Max value fusion method takes the maximum pixel value as the fusion pixel value. The LP fusion algorithm applies Laplacian pyramid features to handle image fusion. The NSCT fusion method is a translation-invariant fusion method that uses non-subsampled pyramid features and non-subsampled directional filter banks to obtain a multi-scale multi-directional decomposition of the source image. The CT fusion algorithm is a kind of texture structure that uses wavelet transform and multi-channel filtering to protect the image. The DTCWT is a multi-scale-based fusion rule that selects a set of optimal coefficients for each level of sub-image, ensuring intra-scale and inter-scale consistency. The PSDWT is a multi-resolution pixel-level fusion algorithm based on the use of wavelet and contourlet transforms.

4) IMPLEMENTATION DETAILS

To reduce the training difficulty, we use 4 convolution blocks and a classification layer in the UPSNet. Except the last convolution block, each convolution block contains a convolution layer, layer normalization layer (LN) and non-linear activation function layer. The detailed parameters of UPSNet are shown in Table 2, where \( k \), \( s \) and \( p \) represent the kernel size, stride and padding of the convolutional layer, respectively. The optimizer of UPSNet is a stochastic gradient descent (SGD) algorithm. The initial learning rate and momentum are set to 0.05 and 0.9, respectively. The number of training iterations is 256. The learning rate dynamic adjustment strategy is Reduce-on-Plateau, and the adjustment ratio is 0.95 times the current learning rate.

| Name | Input | Output | \( Conv(k \times k, s, p) \) | \( LN_p \) |
|------|-------|--------|-----------------|--------|
| CB1  | 512×512 | 128×128 | 7×7, 4, 3 | 64 |
| CB2  | 128×128 | 128×128 | 1×1, 1, 0 | 32 |
| CB3  | 128×128 | 128×128 | 3×3, 1, 1 | 64 |
| CB4  | 128×128 | 128×128 | 1×1, 1, 0 | 32 |
| Argmax | 2 | - | - | - |

The sliding window size is set to 11×11 in the SSWF algorithm. In addition, all experiments are performed on a PC with Intel Xeon(R) CPU E5-2620v3, NVIDIA Quadro M6000 24GB GPU and 48 GB CPU.

B. EXPERIMENTAL RESULTS AND ANALYSIS

1) VISUAL COMPARISON BETWEEN THE UPSNet AND OTHER SEGMENTATION METHODS

Due to the lack of annotations of shadow and layover regions, the visual evaluation is another way to qualitatively evaluate the segmentation performance of different algorithms. The visual comparison of different segmentation algorithms is shown in Fig. 8. In the experiment, the cluster centers of K-Means segmentation algorithm and the SFFCM segmentation method are both 3.

FIGURE 8. The segmented results of different algorithms.

In Fig. 8, the MET algorithm completes the segmentation of the shadow region, and retains most targets and layover region. The SSFCM algorithm is very close to the MET and K-Means algorithms, but the SSFCM cannot effectively suppress noise in the image. The UPSNet can effectively segment shadow and layover regions, and has the best performance among all segmentation algorithms.

2) VISUAL COMPARISON OF THE SSWF AND OTHER FUSION ALGORITHMS

Fig.9 and Fig.10 show the results of direct fusion (D-*) and indirect fusion (ID-*), respectively. In the red region in each image, local details of the fused image are shown. Direct fusion is directly fuse MA SAR images and the indirect
fusion is to directly fuse segmented MA SAR images by the UPSNet.

In Fig. 9, to some extent the fused images of different algorithms effectively reduce the impact of shadow and layover regions. However, due to the existence of layover regions, part of the targets are still covered, resulting in low object detection performance. From the local details, it can be found that the contrast of the directly fused images is very low, resulting in poor visual effect. In addition, Avg, CT, DTCWT and NSCT algorithms lead to the fused image to be too smooth. The proposed SSWF algorithm preserves image contrast to some extent, but it is not ideal.

In Fig. 10, the fused images of segmented MA images can reduce the influence of shadow and layover areas. It can be seen from the results that there can be some black areas in the fused image. However, the existence of the black area does not affect the interpretation of targets. By comparing local details, the SSWF algorithm has the highest image contrast. The proposed SSWF algorithm preserves image contrast to some extent, but it is not ideal.

In Fig. 10, the fused images of segmented MA images can reduce the influence of shadow and layover regions. It can be seen from the results that there can be some black areas in the fused image. However, the existence of the black area does not affect the interpretation of targets. By comparing local details, the SSWF algorithm has the highest image contrast. The proposed SSWF algorithm preserves image contrast to some extent, but it is not ideal.

3) TARGET DETECTION PERFORMANCE OF SINGLE-ANGLE IMAGE AND FUSED IMAGE

Table 3 lists the targets detection performance of single-angle (SA) images and fused images. The D-Ours and ID-Ours represent the target detection of direct fusion and indirect fusion, respectively. The best results are marked in bold. In addition, the two-parameter CFAR (2p-CFAR) detector [38] is used as the detection algorithm. The local window of the 2p-CFAR detector is set to $10 \times 10$ and the false alarm rate ($p_{fa}$) is set to $10^{-5}$. The protect window of 2p-CFAR is set to $21 \times 21$.

In Table 3, the following conclusions can be drawn. Firstly, the detection performance of the SA image varies greatly. The best detection performance is $Sat_6$, with a detection accuracy ($p_d$) of 73%. The worst detection performance is $Sat_7$, with a detection accuracy of 3.5%. It can be concluded that the quality of the SAR image is very sensitive to the look angle. In the SAR image of $Sat_7$, there are amount of shadow and layover areas. The visual performance of the layover areas is very similar to the target, which leads to misidentification of the real target. Secondly, in the detection results of the D-Ours and ID-Ours, it can be seen that the target detection performance of fused images has been significantly improved. Compared with the detection performance of the best SA image ($Sat_6$), they are improved by 12% (D-Ours) and 17% (ID-Ours), respectively. Thirdly, the proposed UPSNet can effectively improve the detection performance. Compared with the detection results of D-Ours and ID-Ours, the $p_d$ of the indirectly fused images is increased by 5%. Fourthly, in metric of FoM, the fused images can significantly reduce the number of false targets. Compared with the $Sat_6$, direct fusion and indirect fusion improve by 7.79% and 23.24%, respectively. Finally, the UPSNet and SSWF algorithms can significantly improve the target detection performance.

Table 3. The comparison of results between SA images and fused images.

| Image | $N_{td}$ | $N_{ud}$ | $p_d(\%)$ | $p_{fa}(\%)$ | FoM |
|-------|---------|---------|----------|-------------|-----|
| $Sat_1$ | 25 | 75 | 62.5 | 37.5 | 0.3750 |
| $Sat_2$ | 57 | 143 | 28.5 | 71.5 | 0.2227 |
| $Sat_3$ | 54 | 146 | 27.0 | 73.0 | 0.2077 |
| $Sat_4$ | 71 | 129 | 35.5 | 64.5 | 0.2011 |
| $Sat_5$ | 115 | 85 | 57.5 | 43.5 | 0.3979 |
| $Sat_6$ | 146 | 54 | 73.0 | 27.0 | 0.5271 |
| $Sat_7$ | 7 | 193 | 3.5 | 96.5 | 0.0330 |
| $Sat_8$ | 45 | 155 | 22.5 | 77.5 | 0.1613 |
| D-Ours | 170 | 30 | 85.0 | 15.0 | 0.6050 |
| ID-Ours | 180 | 20 | 90.0 | 10.0 | 0.7595 |

To better compare the performance of SA images and fused images, the receiver operating characteristic (ROC) curve is given in Fig. 11. It can be seen from the curves that our proposed method performs better than other SA SAR image. Whether it is the direct fusion result (D-Ours) or the indirect fusion result (ID-Ours), their performance has been greatly improved. When $p_{fa}$ is as high as $10^{-1}$, the $p_d$ of the fused image is 100%. And when $p_{fa}$ decrease to $10^{-5}$, the $p_d$ of direct fusion and indirect fusion appears different, namely ID-Ours still keeps good detection performance. This is because the UPSNet and SSWF not only remove layover and shadow regions to reduce the negative influence, but also enhances the contrast of the targets.
4) TARGET DETECTION PERFORMANCE OF DIFFERENT FUSION ALGORITHMS

Table 4 lists the comparison results of different fusion algorithms, where \( t (s) \) represents the time required to fuse an image. The level of the PSDWT algorithm is 3 and the sub-band is set to 10. The level of LP, DTCWT and NSCT fusion methods is set to 4. The level of CT fusion method is set to 5. The detection performance of all fusion methods is based on the fusion of reconnaissance images segmented by the UPSNet, which belongs to the category of indirect fusion.

As can be seen from Table 4, the detection performance of the proposed fusion algorithm is significantly higher than other fusion algorithms in quantitative evaluation metrics. Looking at the \( P_d \), compared with single-scale fusion algorithms, the accuracy of the SSWF algorithm is improved by 27% (Avg) and 30% (Max), respectively. Compared with the multi-scale fusion algorithms, it improves by 22.5% (PSDWT), 39% (DTCWT), 44% (LP), 26.5% (NSCT) and 32% (CT), respectively. Looking at the \( F o M \), the SSWF algorithm achieved 0.7595. In other fusion algorithms, the best algorithm is NSCT, which reached 0.6106. Looking at the computational efficiency (\( t \)), the SSWF algorithm has the fastest speed in the multi-scale fusion algorithms. Although the fusion speed of the SSWF algorithm is slower than the single-scale fusion methods (Avg and Max), the \( P_d \) were improved by 7% (Avg) and 20% (Max), respectively. Compared with the multi-scale fusion algorithms, the SSWF algorithm not only has good detection probability, but also has the advantage of low complexity. Compared with the single-scale fusion algorithms, the improved detection performance of the SSWF algorithm can make up for the disadvantage caused by computational complexity.

5) ABLATION EXPERIMENTS

Table 5 lists target detection performance of different numbers of the MA images, where the 2 MA images are \( \text{Sat}_4 \) and \( \text{Sat}_7 \). The 4 MA images are \( \text{Sat}_1, \text{Sat}_4, \text{Sat}_5 \) and \( \text{Sat}_7 \). The 6 MA images are \( \text{Sat}_1, \text{Sat}_2, \text{Sat}_4, \text{Sat}_5, \text{Sat}_7 \) and \( \text{Sat}_8 \). The principle of the selecting MA images is that the look angle of SAR are followed as evenly as possible in the range of \( 0^{\circ} \sim 360^{\circ} \). As the number of MA images increases, the object detection performance is gradually improved. Looking at the \( P_d \), the accuracy of 2 MA images is 37.5%, while the accuracy of 4 MA images is 65%. But the result of \( P_d \) of 6 MA images is 9% better than that of 4 MA images. Compared with the results of 6 MA images, the \( P_d \) of 8 MA images is increased by 16%. It shows the same trend as the accuracy rate in \( F o M \). In the above ablation experiments, the results of the 2 MA images are the worst. The reason is that the selected images are \( \text{Sat}_4 \) and \( \text{Sat}_7 \). Their \( P_d \) are 35.5% and 3.5%. But the \( P_d \) of the fused image reaches 37.5%, which is improved by 2% (\( \text{Sat}_4 \)) and 34% (\( \text{Sat}_7 \)), respectively.

Fig. 13 lists the ROC curves of different numbers of the MA images. The fused image of 2 MA images have the worst performance of detection. As the number of MA images increases, the detection performance of fused images improves significantly.

Fig. 14 shows the fusion results of different numbers of MA images, and the local details of the fused images are given to the right. As the fusion angle increases, the visual effect of...
is proposed for reducing the negative effects of shadows and overlays. The method mainly includes a spaceborne SAR collaborative observation model, an unsupervised learning progressive segmentation network for shadow and overlapping regions, and a single-scale-based weighted fusion algorithm for segmented images. Firstly, to obtain the MA SAR images of the same target, a spaceborne SAR-CO model is designed. The spaceborne SAR-CO model uses distributed SAR satellites to imaging the target. In SAR imaging processing, discrete frequency encoded signals and FBP algorithm are chosen, where the discrete frequency encode signal is used to reduce signal interference between different SAR satellites. The FBP algorithm can project multiple SAR images in the same coordinate system, which can avoid image registration. Secondly, according to the pixel difference of shadow and layover regions, an UPSNet is proposed. The UPSNet can sequentially complete the detection and segmentation of shadow and layover regions. In addition, the UPSNet reduces the reliance on ground-truth image labels. Thirdly, to improve the brightness and contrast of the fused image, a SSWF algorithm is proposed to fuse segmented MA SAR images. Finally, the proposed method is validated by employing visual subjective results and the performance of target object detection. The experimental results show that the proposed methods significantly improve the accuracy of target recognition.

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, Apr. 2013.
[2] Y. Deng, W. Yu, H. Zhang, W. Wang, D. Liu, and Y. Wang, “Forthcoming spaceborne SAR development,” J. Radars, vol. 9, no. 1, pp. 1–33, 2020.
[3] W. G. Kropatsch and D. Strohl, “The generation of SAR layover and shadow maps from digital elevation models,” IEEE Trans. Geosci. Remote Sens., vol. 28, no. 1, pp. 98–107, Jan. 1990.
[4] Z. Chen, Z. Zeng, Y. Huang, J. Wan, and X. Tan, “SAR raw data simulation for fluctuant terrain: A new shadow judgment method and simulation result evaluation framework,” IEEE Trans. Geosci. Remote Sens., vol. 60, pp. 1–18, 2022.
[5] U. Stilla, U. Soergel, and U. Thoennessen, “Potential and limits of InSAR data for building reconstruction in built-up areas,” ISPRS J. Photogramm. Remote Sens., vol. 58, nos. 1–2, pp. 113–123, Jun. 2003.
[6] M. Eineder and N. Adam, “A maximum-likelihood estimator to simultaneously unwrap, geocode, and fuse SAR interferograms from different viewing geometries into one digital elevation model,” IEEE Trans. Geosci. Remote Sens., vol. 43, no. 1, pp. 24–36, Jan. 2005.
[7] U. Soergel, K. Schulz, and U. Thoennessen, “Enhancement of interferometric SAR data using segmented intensity information in urban areas,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., Take Pulse Planet, Role Remote Sens. Manag. Environ. (IGARSS), Jul. 2000, pp. 3216–3218.
[8] X.-Y. Du, Q. Yang, B. Cai, and D.-N. Liang, “A new method on shadow and layover detection of InSAR,” in Proc. 6th Asia-Pacific Conf. Antennas Propag. (APCAP), Oct. 2017, pp. 1–3.
[9] L. Wu, H. Wang, Y. Li, Z. Guo, and N. Li, “A novel method for layover detection in mountainous areas with SAR images,” Remote Sens., vol. 13, no. 23, p. 4882, Dec. 2021.
[10] A. Lapini, T. Bianchi, F. Argenti, and L. Alparone, “Blind speckle decorrelation for SAR image despeckling,” IEEE Trans. Geosci. Remote Sens., vol. 52, no. 2, pp. 1044–1058, Feb. 2014.
[11] X. Cai, L. Chen, J. Cai, J. Xing, X. Xing, R. Luo, S. Tan, and J. Wang, “Automatic extraction of layover from InSAR imagery based on multilayer feature fusion attention mechanism,” IEEE Geosci. Remote Sens. Lett., vol. 19, pp. 1–5, 2022.
