IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 20 June 2023, accepted 4 July 2023, date of publication 10 July 2023, date of current version 14 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3293644

RESEARCH ARTICLE

Millimeter Wave SAR Imaging Denoising and Classification by Combining Image-to-Image Translation With ResNet

PHAM THE HIEN AND IC-PYO HONG, (Member, IEEE)
Department of Smart Information and Technology Engineering, Kongju National University, Cheonan 31080, South Korea

Corresponding author: Ic-Pyo Hong (iphong@kongju.ac.kr)

This work was supported in part by the Basic Science Research Program under Grant 2020R1I1A3057142, in part by the Priority Research Center Program through the National Research Foundation under Grant 2019R1A6A1A03032988, and in part by the Research Grant of Kongju National University in 2022.

ABSTRACT Synthetic aperture radar (SAR) imaging has recently attracted considerable attention due to its variety of applications in both military and civilian aspects. However, a SAR image scheme can be affected by various elements that can lead to poor image reconstruction performance, especially for the target recognition mission; for instance, the complex environment, irregular sampling intervals, sample scarcity, imaging parameters, etc. The rapid development of deep learning currently makes it a great solution to deal with the aforementioned problems. In this paper, we propose a SAR image model based on conditional generative adversarial networks (cGAN), which combines image-to-image translation (pix2pix) and residual networks (ResNet) in order to diminish the noise and artifacts on SAR images, increase their signal-to-clutter-noise ratio (SNCR) of the images, and improve the short-range target recognition rate. Unlike conventional cGAN, we employ a ResNet-based discriminator (RbD) to effectively improve the SAR image denoising ability of the model. On the other hand, another similar discriminator is simultaneously trained to classify 14 familiar metallic object types with high accuracy and avoid the over-fitting problem. This discriminator is built by replicating the RbD one, and then we replace the last layer with the standard softmax function to classify multiple objects based on class probability outputs. The experiment results in this paper illustrate that the proposed scheme achieves higher image denoising performance and SNCR enhancement than the other conventional approaches. Besides, the target recognition rate of the proposed scheme outperforms the other common classification models.

INDEX TERMS Automatic target recognition, frequency-modulated continuous wave, generative adversarial networks, millimeter wave radar, residual networks, synthetic aperture radar.

I. INTRODUCTION

Synthetic aperture radar (SAR) utilizes the capability of electromagnetic (EM) waves to generate images without the influence of visually opaque elements coming from the weather, light, clouds, fog, walls, etc. Compared to other techniques, for instance, optical sensors, infrared sensors, etc., SAR can perform in all types of weather conditions to accomplish continuous observation with its high penetration ability. It was first introduced as an effective method in the remote sensing field to reconstruct aerial images supporting ground or sea detection [1]. Furthermore, in recent years, SAR has also been widely adopted in short-range applications, including indoor detection [2], [3], non-destructive testing (NDT) [4], [5], structure health monitoring (SHM) [6], [7], and medical imaging [8], due to its valuable characteristics. Especially when working with higher frequencies, high-resolution image results can be achieved.

However, there are several challenges that SAR has to encounter to generate high-quality images. First, the high-frequency EM wave is sensitive to complex environments including occlusion, stacking, concealment, EM interference, etc., bringing about the low signal-to-clutter-noise ratio (SNCR) [9]. Secondly, SAR imaging has been supposed

The associate editor coordinating the review of this manuscript and approving it for publication was Gerardo Di Martino.

to work in the regular space domain with free-space propagation [10], [11]. Nevertheless, recent researchers have extended their SAR scheme to operate in harsh situations by using UAV and freehand imaging to detect landmines [12] or using smartphones as handheld scanners [13] where irregular sampling is unavoidable [14], [15]. Although these approaches improve the flexibility of SAR imaging in extreme circumstances, the irregular sampling combined with the existing limitations of conventional SAR imaging leads to clutter and artifacts in the image results [16], [17]. Next, the sample scarcity is caused by the variation of the targets or even the difference of the same target; for example, different observation angles, variations of structure, or their variants come from the target damage, which results in weak target recognition ability [18]. Last but not least, the influence of configuration parameters of a SAR imaging scheme, including the frequency, imaging mode, incident angle, target range, signal-to-noise ratio (SNR), resolution, etc., makes the object classification performance of one SAR scheme hard to apply to other objects quickly. Therefore, a method to improve denoising and recognition performances needs to be studied to enhance the effectiveness and adaptability of a SAR imaging scheme.

In terms of denoising, compressed sensing (CS) algorithms are widely utilized as an effective method to restore SAR data from sparse representation. As a result, this approach reduces the amount of required measurement data and achieves high resolution and image quality more efficiently than other traditional deterministic techniques, such as low-pass filters and singular value decomposition (SVD). In [19], the authors propose a peak search and CS (PS-CS) to alleviate the computational complexity of conventional CS as well as generate the super-resolution image results by extrapolating the parameters of scattering along with the observing angle and frequency dimensions. Another research employs scaled CS adopting the basis pursuit denoising (BPDN) algorithm to reconstruct low-frequency information from randomly sub-sampled raw signals [20] collected by using a Ku-band frequency modulated continuous wave (FMCW) SAR scheme mounted on an automobile platform. The authors in [21] reduce the speckle noise in the original image using the CS with BM3D as a non-local mean as well as improve performance parameters such as peak signal-to-noise ratio (PSNR) and SNR values. The main problem of the CS approach is that it requires accumulating measurement data following a pre-defined observation matrix (Gaussian random matrix, Bernoulli random matrix, discrete cosine transform matrix, etc.), which is slightly difficult in practice. Besides that, the CS also needs a solver algorithm to solve the l_1 or $l_2 - norm$ problems, which are usually computationally complex and consume a considerable amount of time for finding optional solutions.

In recent years, deep learning has demonstrated its great potential in image processing for both image quality improvement and image target detection, which has been well utilized not only in optical but also SAR images. For instance, in the study [22], the authors proposed a SAR sea-ice image classification method in which the ResNet and long short-term memory (LSTM) networks are combined to classify seven ice types from the spatial and temporal input features with an accuracy rate of 95.69%. Another approach adopting ResNet is introduced in [23] for the semantic segmentation of high-resolution polarimetric SAR (PolSAR) images. The method reported in this paper defines a multipath ResNet (MP-ResNet) to enlarge the valid receptive fields (VRFs) to learn the semantic contexts of the 500 pairs of 512×512 pixels PolSAR images in the open dataset from the Gaofen contest to archive high segmentation accuracy. The authors in [24] define a model comprising a denoise subnetwork and a classification one where the classifier employs the denoised images produced by the despeckling subnetwork as the training data to reach 82.19% recognition rate of ten ground target classes at the highest noise level. After first being introduced by Ian Goodfellow [25], conditional generative adversarial networks (cGAN) have become an effective method to enrich data for a classification model due to their excellent performance in translating images to other potentially similar features. For example, researchers in [26] and [27] adopted cGAN to generate high-quality SAR aerial images before using them as a source for ground target recognition. The authors in [28] and [29] used cGAN as an effective tool to convert SAR images from their optical versions or generate realistic optical images from SAR images [30], [31], [32].

Furthermore, the study in [33] shows another excellent performance of cGAN in mitigating speckle noises and artifacts in near-field SAR images. The denoised SAR image of an object is generated from its noisy SAR image via the image-to-image translation (pix2pix) model using its real-shape image as the reference. The reference image is one of the main elements of the pix2pix model to generate as realistic images as possible. However, using the real-shape image as the reference without considering SAR imaging parameters including the target range, its dimension on the actual SAR image, and its reflectivity may lead to inaccurately generated output images. Motivated by it, in our study, we propose a combined network comprising the pix2pix model along with ResNet to improve the denoising performance on SAR images. In our model, we utilize noisy SAR images, which contain a lot of ghost effects, speckle noise, and low SNR due to the interference and side lobe effect caused by complex measurement environments and irregular sampling of SAR imaging schemes [34], [35]. While a high-quality SAR image of an object used as the reference image is reconstructed from 100% radar data, the corresponding noisy SAR image is formed by randomly taking 10% of that radar data, which represents the irregularly sampled data. Additionally, additive Gaussian noise is added to the noisy images as interference. Besides that, we combined three loss functions, namely GAN adversarial loss, mean absolute error (MAE), and structural

similarity index (SSIM), to update the model over training iterations for generating high-quality SAR images. A ResNet-based discriminator (RbD) is set up to improve the denoising ability of the model. Moreover, our model also simultaneously trains the ResNet-based object classifier with the involvement of 14 metallic object classes, which supports improving the accuracy rate in the concealed object detection mission. So in this paper, we can accomplish the following:

- Generate high-quality denoised SAR images;
- Enhance the SNCR of the generated SAR images by the model;
- Concurrently attain the ResNet-based object recognition model;

The rest of the paper is organized as follows: In Section II, the FMCW SAR imaging and some of the main sources of noise on SAR images are presented. The concept of cGAN in SAR image denoising is illustrated in section III. After that, section IV demonstrates the performance results of the proposed approach, followed by section V as the conclusion.

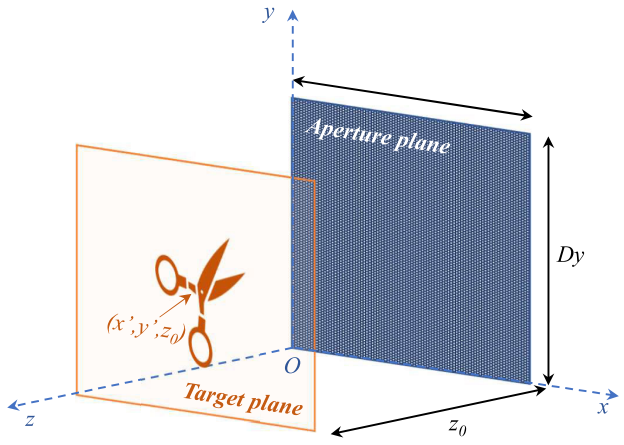


FIGURE 1. Mono-static measurement of a target by radar on aperture plane.

II. FMCW SAR IMAGING AND NOISE SOURCES

A. FMCW SAR IMAGING

In this paper, we adopt the range migration algorithm (RMA) to reconstruct the SAR image from the acquired signal data, as it works with motion compensation and requires a one-dimensional (1-D) interpolation to completely compensate for the curvature of the wavefront [36], [37]. To achieve high-resolution images, a low-cost FMCW radar working in the millimeter wave band is utilized. Assuming that the signal measurement is continuously acquired at all points on a synthetic aperture plane as shown in Fig. 1. The emitted signal from the radar can be expressed as:

$$s_{tx}(t) = e^{\left(j2\pi(f_0 t + \frac{1}{2}\alpha t^2)\right)}, \tag{1}$$

where f_0 is the carrier frequency at $t = 0$ and α is the slope frequency over time, which is indicated as $\alpha = \frac{BW}{T_c}$, with the bandwidth BW over the chirp duration T_c. The reflected signal from a target received by the radar can be formulated as the delayed version of the transmitted one:

$$s_{rx}(t-\tau) = \frac{\rho}{R^2} e^{\left(j2\pi\left(f_0(t-\tau)+\frac{1}{2}\alpha(t-\tau)^2\right)\right)}, \tag{2}$$

where τ is the delay between the transmitted signal and its echo, which is proportional to the range from the radar to the scatter point (R) as $\tau = \frac{2R}{c}$ with c being the speed of light, and ρ is the target reflectivity of the reflector over range R. Then, by multiplying the transmitted and received signals through the frequency mixer (IQ mixer) and passing the output through the low-pass filter, the intermediate frequency (IF) signal can be obtained:

$$\begin{aligned} s_{IF}(t) &= s_{tx}(t) s_{rx}(t-\tau) \\ &= \frac{\rho}{R^2} e^{\left(j2\pi\left(\alpha\tau t + f_0\tau - \frac{1}{2}\alpha\tau^2\right)\right)}, \end{aligned} \tag{3}$$

The IF signal can be recognized here as the sinusoidal signal with the beat frequency $f_b = \alpha\tau$ consisting of the time delay so that it can provide the range information of the scatterer. For near-field imaging, the residual video phase (RVP) which is the last term of (3) can be approximated as $f_0\tau - \frac{1}{2}\alpha\tau^2 \approx f_0\tau$. Hence, the IF signal in (3) becomes:

$$s_{IF}(t) = \frac{\rho}{R^2} e^{(j2\pi(f_b t + f_0\tau))}, \tag{4}$$

the IF signal peaks at the beat frequency f_b which is projected to carry the target distance from the radar (z_0). In order to reconstruct the two-dimensional (2-D) target image, the reflected signal acquisition for all scatter points in the volume under test is required. Assuming that the measurement system adopts the single-input-single-output (SISO) technique and continuously captures the signal over the synthetic aperture plane Oxy (as shown in Fig. 1), the IF signal of the target plane located at (x', y', z_0) can be inferred by taking the integral of (4) as:

$$s(x,y,t) = \iint \rho(x',y') \frac{e^{(j2\pi(f_b t + f_0\tau))}}{R_0^2} dx'dy', \tag{5}$$

where $R_0 = \sqrt{(x-x')^2 + (y-y')^2 + z_0^2}$. By getting the received signal at all positions on the target plane, the reflectivity of the target $\rho(x', y')$ on that plane can be characterized. The time-domain IF signal in (5) can be written in the wavenumber domain as:

$$s(x,y,k) = \iint \rho(x',y') \frac{e^{(j2kR_0)}}{R_0^2} dx'dy', \tag{6}$$

with $\frac{2\pi f_0}{c} \le k \le \frac{2\pi f_b}{c}$. For near-field SAR, a spherical wave can be represented as a superposition of plane waves by using the method of stationary phase [38]:

$$\frac{e^{j2kR_0}}{R_0} \approx \frac{j}{2\pi} \iint \frac{e^{j(k_x(x-x')+k_y(y-y')+k_z z_0)}}{k_z} dk_x dk_y, \tag{7}$$

where k_x, k_y and k_z are three components of k in the Cartesian coordinate system and obey the following dispersion relation:

$$k_z = \sqrt{4k^2 - k_x^2 - k_y^2}, \quad k_x^2 + k_y^2 \leq 4k^2. \tag{8}$$

The signal data becomes as below by submitting (7) to (6):

$$s(x, y, k) = \frac{j}{2\pi} \iint \left[\iint \rho(x', y') e^{-j(k_x x' + k_y y')} dx' dy' \right] \times \frac{e^{jk_z z_0}}{k_z} e^{j(k_x x + k_y y)} dk_x dk_y. \tag{9}$$

By applying discrete Fourier transform, $s(x, y, k)$ can be expressed as:

$$s(x, y, k) = \frac{j}{2\pi} \iint \mathcal{F}_{2D}[\rho(x', y')] \times \frac{e^{jk_z z_0}}{k_z} e^{j(k_x x + k_y y)} dk_x dk_y, \tag{10}$$

where $\mathcal{F}_{2D}[\cdot]$ denotes the 2-D Fourier transform operation. The equation (10) gives:

$$s(x, y, k) = \mathcal{F}_{2D}^{-1} \left[\mathcal{F}_{2D} \left[\rho(x', y') \frac{e^{jk_z z_0}}{k_z} \right] \right], \tag{11}$$

where $\mathcal{F}_{2D}^{-1}[\cdot]$ denotes the inverse 2-D Fourier transform operation. Hence, the target complex reflectivity can be inferred as:

$$\rho(x', y') = \mathcal{F}_{2D}^{-1} \left[\mathcal{F}_{2D}[s(x, y, k)] \times k_z e^{-jk_z z_0} \right]. \tag{12}$$

The detailed process of SAR image reconstruction using RMA is revealed in the Algorithm 1

Algorithm 1 SAR Image Reconstruction Using RMA

Input: The 3-D data cube $\boldsymbol{S}$ of the FMCW echo in the time domain is accumulated by measuring at all measurement points on the aperture plane.

Output: Image reconstruction results

1: Applying the 2-D Fourier transform to attain the wave number domain data, denoted as:

$$\boldsymbol{M} = \mathcal{F}_{2D}[\boldsymbol{S}] \tag{13}$$

2: Define k_z as in equation (8), and phase factor (PF) at distance z_0 as:

$$\text{PF} = e^{-jz_0 k_z} \tag{14}$$

3: Using the approach in (12) to reconstruct the image by taking the inverse 2-D Fourier transform on the Hadamard product of the wave number domain data, k_z and phase factor:

$$\hat{\boldsymbol{M}} = \mathcal{F}_{2D}^{-1}[\boldsymbol{M} \otimes k_z \otimes \text{PF}], \tag{15}$$

where the $\otimes$ represents the Hadamard (or element-wise) product. The image of the target area can be achieved by taking the modulus display.

B. NOISES ON SAR IMAGES

SAR imaging creates a radar image with coherent radiation from a small area of the image scene (cell resolution). The speckle noise mainly comes from the random interference of several scatters within a resolution cell. The high differences in the phases of back-scattered signals from those scatters produce light and dark pixels (or speckle noise) [39], [40]. In contrast, there might only be one scatter in the resolution cell if the resolution is high. Hence, the image is considered to not suffer from speckle noise [41].

In the previous section II-A, the imaging scheme works under the assumption that the radar continuously captures echo signals over the synthetic plane. Nevertheless, in practice, the radar cannot perform all measurements at all points on the synthetic aperture plane. Instead, the signals from the grid of the target plane are discretely sampled. In the first step of the image reconstruction algorithm, to improve the efficiency of the 2-D Fourier transform, uniform spatial sampling is required. However, discrete data capturing requires equal sampling, which is hard to accomplish, especially when using some imaging schemes with flexibility sampling techniques such as freehand imaging [42] or unmanned aerial vehicle (UAV) SAR imaging schemes [43]. Hence, the spectrum data obtained by the Fourier transform overlapped due to undersampling, resulting in the side lobe effect. Moreover, the side lobes of scatters interference with other nearby scatters can lead to an increment of the speckle noise and the loss of radiometric information [34], [35]. Additionally, the SAR imaging process pursues extracting the reflectivity using approximations behind each step that yield some artifacts in the image results.

III. DENOISING SAR IMAGE BY CGANS WITH RESNET

GAN is a generative framework that can generate new outputs through adversarial processes according to a certain distribution learned from the system [25]. In GAN, two models are trained at the same time to compete with each other in a min-max game. The first model is known as the generator (G), which tries to generate the output from a latent space input based on a data distribution. The other model is referred to as the discriminator (D), whose binary output is the probability estimation of the generator output following the desired distribution.

In conventional GANs, due to the excessive freedom of the output from the G, unreasonable samples are generated to deceive the D. To deal with this problem, cGAN, which is a special version of GAN, comes into play. In cGAN, the outputs are conditioned by certain inputs, which are aimed at minimizing the distance between the generated outputs and their corresponding inputs. In this paper, we utilize the noisy SAR images as the input condition of G, which can reduce the degree of freedom of the output images. This approach is known as pix2pix, which is one of the most successful

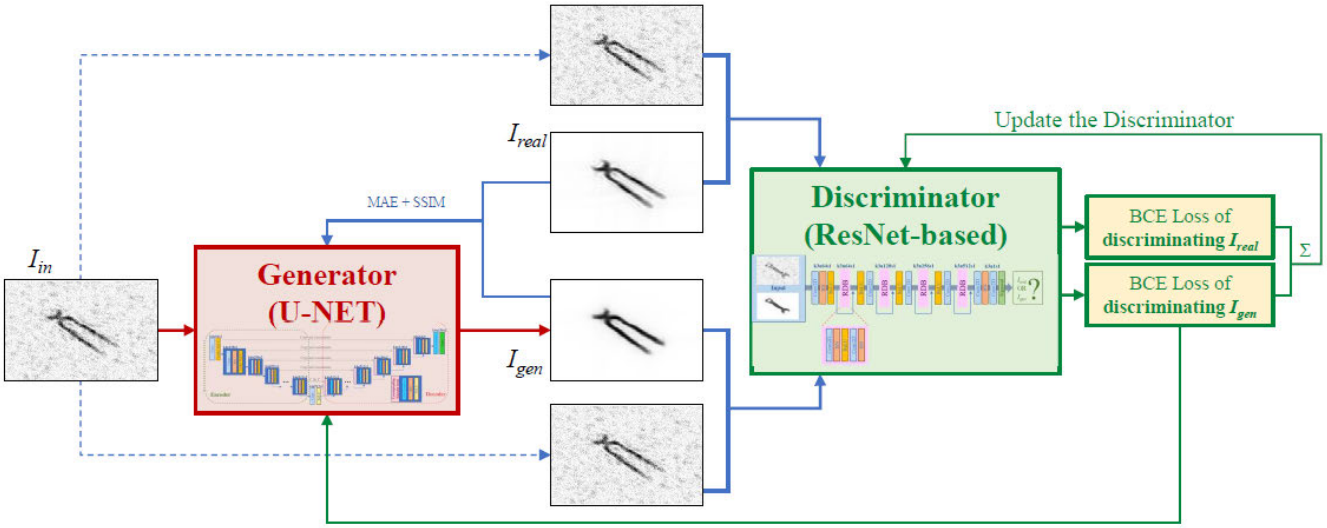


FIGURE 2. Architecture diagram of cGAN in denoising SAR image.

applications of cGAN [44], [45]. The architecture of the model is described in Fig. 2.

As shown in Fig. 2, the model employs three types of different images. The input images are referred to as I_{in} which are the SAR images containing artifacts and alias areas coming from the conventional imaging scheme without applying any further processing. These images are utilized by the G during the training phase to generate the corresponding clean SAR images I_{gen}. In general, the G can be trained using one or two channels of the input images, including the real and imaginary parts of the reflectivity. However, in this paper, we only take the magnitude information as the input channel to the model, as the SAR images are simply reconstructed from the magnitude of the reflectivity. It is important to note that the model requires a large amount of input image data during the training phase, but it may not be possible to accumulate a complete dataset from measurements. Therefore, we need to generate SAR images (e.g., via simulation) taking into account the imaging configuration parameters to reasonably achieve the SAR image outputs I_{gen}. The last term of the images used in the model is real SAR images, which are related to I_{real}. These SAR images are the clean version of the corresponding input images, which are also simulated with careful consideration of imaging parameters. The model employs these images as a reference for idea representations that need to be reconstructed.

A. IMAGE TRANSLATION MODEL DESIGN

First, the structure of the discriminator is presented. Unlike conventional GAN models, in which the discriminator is formed by convolutional neural networks (CNN) [30], [33], we employ a ResNet-based structure to build it up. The residual blocks in the ResNet architecture have been shown to effectively mitigate the problem of vanishing gradients during training, which can occur in traditional deep neural networks (DNN) with many layers, leading to slow or stalled training progress [46]. In addition, ResNet allows for the construction of very deep neural networks to capture fine-grained features and patterns in data that can help the discriminator in GANs make more accurate and confident decisions. In general, the discriminator works during the training phase only to distinguish whether the input image is the real image or the artificially generated image, but in this research, we also construct another similar discriminator structure for the multi-classification objective, which will be introduced in the next subsection.

The structure of the RbD in our model is illustrated in Fig. 3a. The input layer is followed by a 2-D convolutional (Conv2D) layer, a batch normalization (BN) layer, and a rectified linear unit (ReLU) activation layer. The kernel size of 3×3 and the kernel stride of 1×1 are applied to all Conv2D layers in the RbD model. Furthermore, this RbD comprises four similar-structure residual blocks, as shown in Fig. 3b. Each block contains a standard feed-forward CNN stacked on five layers, including two Conv2D layers, two BN layers, and one ReLU activation layer. A shortcut connection identically maps the input of the residual block x_{in}, then its output is added to the outcomes of the stacked layers $f(x_{in})$ before going through the ReLU layer. This identity shortcut link does not add calculation complexity or parameters, which can enable the deeper network to be effectively and implicitly tuned over the course of training. The dimensions of the input and output x_{out} of each block are the same, and the output of each residual block takes the contributions of the residual function and the shortcut connection:

$$x_{out} = \text{ReLU}\left(f(x_{in}) + x_{in}\right) \text{ with } \text{ReLU}(x) = \max(x, 0). \tag{16}$$

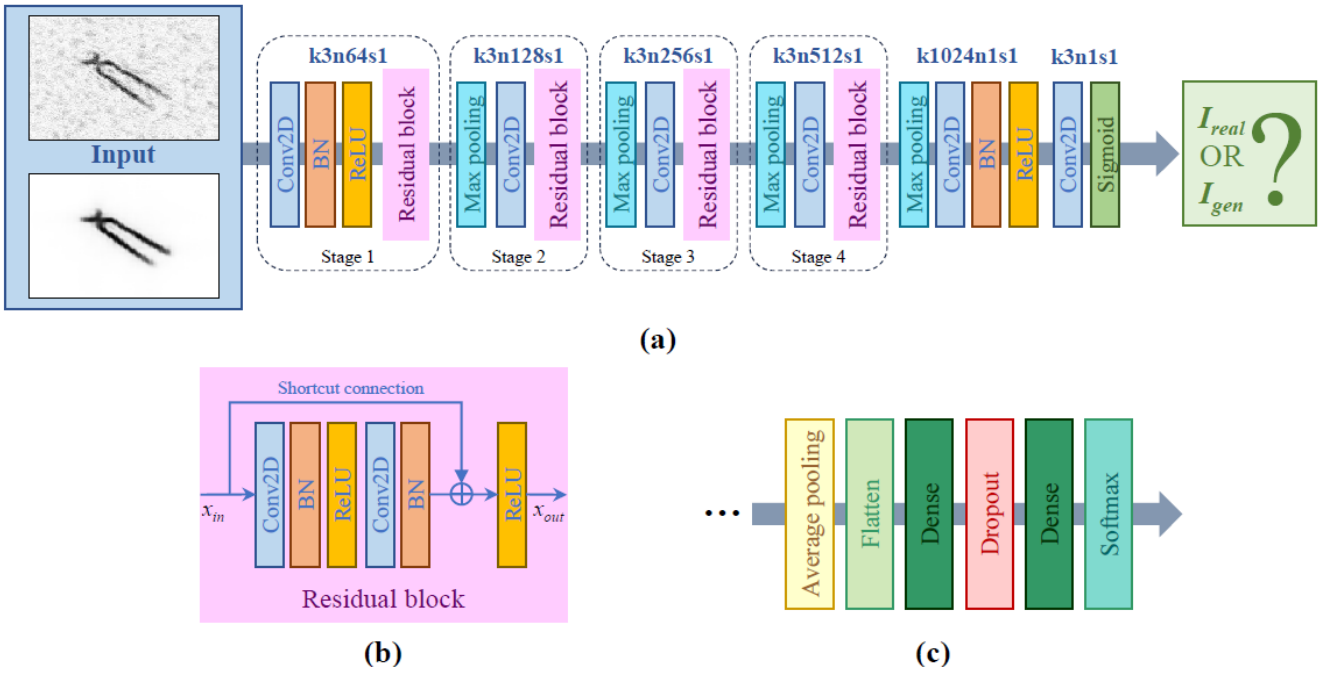


FIGURE 3. The architectures of (a) ResNet-based discriminator of denoising model, (b) the residual blocks, and (c) the classification layers of the object recognition model.

The input layer of our RbD model takes two input images corresponding to the number of channels: the first one is the noisy image I_{in}, and the second one is the image that needs to be classified as a real image I_{real} or a generated image I_{gen}. It is expected to correctly determine whether I_{real} and I_{gen} are real or generated images. If the input images are I_{in} and I_{real}, then the discriminator output should be $D(I_{in}, I_{real}) = 1$. On the other hand, its output should be $D(I_{in}, I_{gen}) = 0$ for the input images of I_{in} and I_{gen}. Therefore, the output of this network is the sigmoid activation layer, which only brings about the probability of real or fake (artificially generated) images. Consequently, the following objective function is utilized for training the D to teach the G to generate images I_{gen} as close to the scenes of I_{real} as possible [44]:

$$\mathcal{L}_{cGAN}(G, D) = \max_{D} \left(\mathbb{E}_{I_{real}} \left[\log D(I_{in}, I_{real}) \right] + \mathbb{E}_{I_{gen}} \left[\log \left(1 - D(I_{in}, G(I_{in})) \right) \right] \right), \quad (17)$$

where $G(I_{in})$ represents the image (I_{gen}) generated by G, which takes only one latent space input (I_{in}), and $\mathbb{E}$ denotes the mathematical expectation output from the discriminator. While D tries to maximize the loss function, generator G strives to minimize it by generating images that can fool the discriminator. During the training process, the generator should pursue the minimization of the loss function (17) i.e. $G^* = \arg\min_G \max_D \mathcal{L}_{cGAN}(G, D)$. Eventually, the model distribution can be learned through this adversarial min-max optimization.

In the regular pix2pix model, MAE loss is used as the regression function constraint to not only reduce the degree of freedom of the output of G but also minimize the distance between the I_{gen} and I_{real}, and thus enhance the SNCR of the output SAR images. It can be expressed as:

$$\mathcal{L}_{MAE}(G) = \mathbb{E}_{I_{real}, I_{gen}} \left[\| I_{real} - I_{gen} \|_1 \right], \quad (18)$$

where $\| \cdot \|_1$ is the ℓ_1-norm. Using the MAE loss can decrease the distance between the output image and the real one, but it tends to blur the output. Therefore, in our pix2pix model, we combine the cGAN loss, MAE loss, and another SSIM loss to train the generator. The use of SSIM loss aims to adjust the G for generating the SAR image with better structural information. The SSIM loss between the I_{real} and I_{gen} can be defined as:

$$\mathcal{L}_{SSIM}(G) = \frac{1}{N} \sum_{\substack{p \in I_{real}, \\ p \in I_{gen}}} 1 - \frac{2\mu_m \mu_n + C_1}{\mu_m^2 + \mu_n^2 + C_1} \times \frac{2\sigma_{mn} + C_2}{\sigma_m^2 + \sigma_n^2 + C_2}, \quad (19)$$

where N is the number of pixels in each image; m and n are the pixels at the position p in I_{real} and I_{gen}, respectively; μ_m and μ_n represent the mean values of pixels in I_{real} and I_{gen}; σ_{mn} is the covariance of the pixels in two images; σ_m and σ_n denote the standard deviation values of pixels in I_{real} and I_{gen}. C_1 and C_2 are the constants to stabilize the division with a weak denominator.

The final loss function to update the weights of the discriminator and the generator is defined as:

$$\arg\min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda_1 \mathcal{L}_{MAE}(G) + \lambda_2 \mathcal{L}_{SSIM}(G), \quad (20)$$

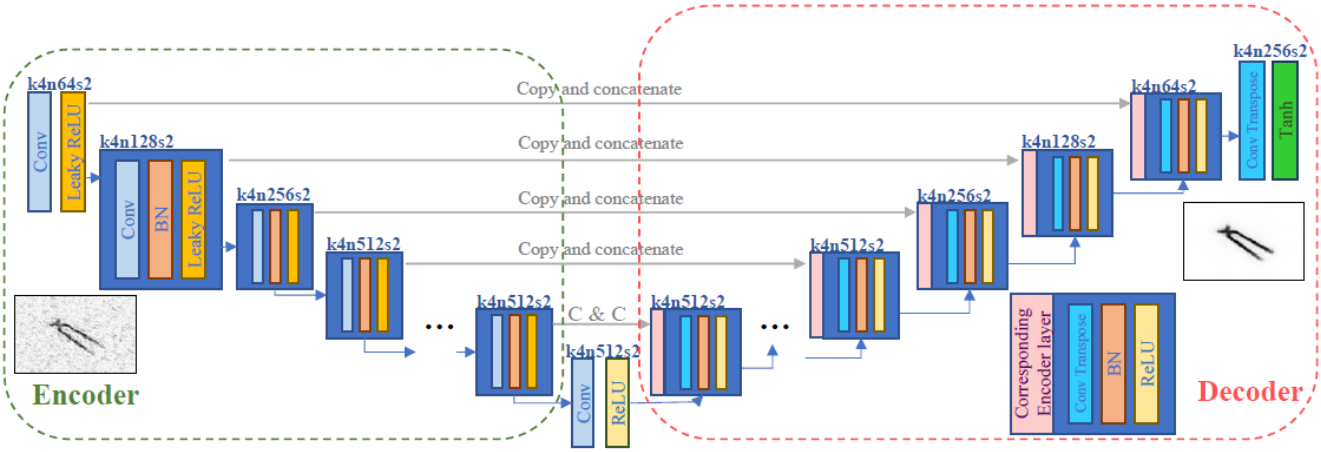


FIGURE 4. Generator with U-Net structure.

where the λ_1 and λ_2 denote the weights of the MAE and SSIM losses. This min-max game allows the discriminator to evolve during training and, thus, improves the generator's ability to produce more natural outputs over iterations. Nevertheless, the discriminator is not used after the training phase, as the generated SAR images from the generator now are expected to resemble the real ones (clean SAR images), which are not able to be distinguished by the discriminator.

The architecture of the generator is designed by utilizing the U-Net structure [44] as shown in Fig. 4 in which the skip connections are added between the i^{th} layer and the $n - i^{\text{th}}$ layer of the generator so that it enables the input feature information at each downsampling layer to be directly transferred to its corresponding upscaling layer. This approach is projected to provide clean SAR images of the targets as lots of low-level information, such as structure and detail information, is shared between output and input. Specifically, the first layers of encoder blocks can detect low-level features (e.g., corners, edges, etc.), while the later layers can detect the complex shapes of the targets. On the decoder blocks, the transposed convolutions are combined with the transferred information from the skip connection to upscale the image to the final output I_{gen}.

B. OBJECT CLASSIFICATION MODEL

In this paper, we deploy another discriminator for metallic object classification by replicating the RbD in the previous section III-A. However, unlike the discriminator in our cGAN model, which only has two outputs, 0 and 1, for the real and generated images, respectively, this discriminator performs its classification ability on multiple object types. Therefore, we replace the last output sigmoid activation layer of the original one with the softmax activation function, whose output is a vector of class probabilities. Specifically, the structure of the classification layers of the object classifier is depicted in Fig. 3c. After the final residual block of the RbD, we added an average pooling layer, a flattening layer, and dense layers. In addition, to avoid the overfitting problem, we add a dropout layer with a rate of 0.5 before the output layer. Besides that, the L2 regularizer is also applied to all 2-D convolutional layers in all the residual blocks.

During the training phase, this classification model only takes pairs of I_{in} and the I_{gen} as the input features. It is different from the RbD of the denoising model, which also takes the I_{real} into discrimination. Therefore, although two discriminators simultaneously learn similar input features of I_{gen} for each iteration, their network parameters are always dynamically different. Furthermore, the local receptive domain of this classifier uses features without taking into account the translation, scaling, and rotation of the image; hence, it can directly work as SAR automatic target recognition (ATR).

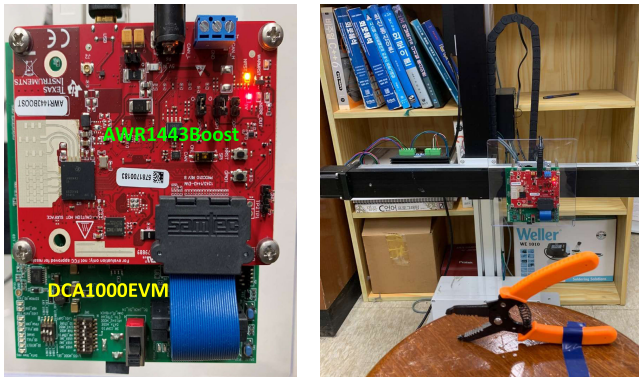
IV. RESULTS AND DISCUSSION

To validate the performance of the proposed method in SAR image denoising and ATR, we take the flexible measurement SAR approach into consideration, whose main noise source comes from the irregular spatial sampling of data acquisition and sparse measurement observations. As stated in the early aforementioned section III, only the magnitude of the reflectivity images are utilized as inputs to the networks, whose size is 128×128. A similar output size is achieved for the output images after denoising. To handle that size of input and output data, the generator is designed with six encoder blocks and six corresponding decoder blocks of the U-Net structure, along with a 2-D convolution layer at the bottom. The numbers of filters in the encoder block are, respectively, 64, 128, 256, 512, 512, and 512 from the top to the bottom. In the first encoder block, the batch normalization layer is not required as the input data have already been normalized to the range of $[-1, 1]$. On the other hand, the same number of filters are applied to the corresponding decoder blocks. Whereas the RbD is built up with four residual blocks whose filter sizes are 64, 128, 256, and 512 to bring about deeper learning but avoid

the vanishing gradient problem. The network parameters of the model are summarized in Table 1.

TABLE 1. The system parameters of the denoising model.

Common parameters	
Optimizer	Adam
Learning rate	0.0002
β	0.5
Generator parameters	
Input/output size	128×128
Leaky ReLU scale	0.2
Filter size	4×4
Stride	2×2
Discriminator parameters	
Input size	$2 \times 128 \times 128$
Output size	1×1 (0 or 1)
Filter size	3×3
Stride	1×1



(a) (b)

FIGURE 5. Measurement system: a) the TI AWR1443Boost and b) the motion platform.

In this paper, since we aim to detect the shape of the target at short range, the FMCW SAR imaging topology is utilized [47]. This topology exploits the SISO technique since the distance between transmitted and received antennas is close enough to be represented as a single full-duplex antenna. Thus, our SAR imaging scheme employs the Texas Instrument AWR1443Boost, which is the evaluation module using a single-chip AWR1443 millimeter wave (mmWave) sensor [48]. The device supports 4 GHz bandwidth sweeping from 77 GHz to 81 GHz, and it is configured to interface with the data card DCA1000EVM to stream the captured data from the radar to the host PC over an Ethernet connection. Two of these devices are mounted on a 2-D motion platform to perform the measurements on the aperture plane. Fig. 5 shows our measurement system containing the radar, data card device, and motion platform with two orthogonal linear actuators. The measurement setup of the imaging scheme is depicted in Fig. 5b. In this setup, the radar is moved on the 200 mm $\times$ 200 mm aperture plane to cover the target dimension. The object is placed at a distance $z_0 = 200$ mm from the radar. To reconstruct the real SAR image of an object I_{real}, we take the involvement of all captured signal data within the aperture plane with a movement step of 2 mm in both horizontal and vertical directions, resulting in a total of 10,000 measurement points. The imaging parameters are shown in Table 2.

TABLE 2. FMCW SAR imaging parameters.

FMCW radar parameters	
Center frequency	77 GHz
Bandwidth (B)	4 GHz
Slope frequency (α)	63.343 MHz/second
Sampling rate	9.121×10^6 samples/second
Imaging parameters	
Horizontal length (Dx)	200 mm
Vertical length (Dy)	200 mm
Movement step ($\Delta x, \Delta y$)	2 mm
Target range (z_0)	200 mm

To generate the dataset for training the model, we simulate the SAR images of 14 common metallic objects by using electromagnetic simulation using the above radar and SAR configuration parameters, which take into consideration the target distance, the electromagnetic wave propagation, and the target reflectivity. For each type of target, there are several different target samples taken into the simulation, and the shapes of all target samples in one specific target class are different from the others in that class. In addition, there are eight I_{real} and eight corresponding I_{in} images for each target sample according to eight different rotation angles. Additionally, to form the corresponding noisy SAR image I_{in} of a target, we randomly take 10% of the data from the simulated measurement, corresponding to 1000 random measurement points, which represent the irregular spatial sampling. Besides that, we also add different levels of additive Gaussian noise to those images (with SNRs of 5, 10, 15, and 20 dB), which would properly model the noise source described in the section II-B. Consequently, combining the rotation angles and different SNR levels produces 16,400 simulated images. 70% of the data is taken into the training phase, and the remaining 30% is used to test the model. The problem is simulated using MATLAB [47] adopting the simulation parameters as shown in Table 2. The detailed information on 14 targets for SAR image simulation is depicted in Table 3.

During the training phase, to quantify the error between the real image (I_{real}) and the denoised image (I_{gen}), the mean square error (MSE) is derived as:

$$\mathrm{MSE}(I_{gen}, I_{real}) = \frac{\sum_{\substack{i=1,\\ j=1}}^{M,N} \left[I_{gen}(i,j) - I_{real}(i,j)\right]^2}{M \times N}, \quad (21)$$

where M and N are the number of rows and columns in the input images. Fig. 6 illustrates the MSE between the denoised images and the ground truths during the training phase. It is important to note that the image quality is steadily improving as the MSE of generated images decreases over iterations with a fast convergence rate. On the other hand, Fig 7 shows

TABLE 3. Object information for simulating SAR images.

#	Object type	No. of object	No. of simulated images
1	cogwheel	15	1200
2	fork	10	800
3	gun	20	1600
4	hammer	20	1600
5	key	10	800
6	knife	30	2400
7	nut	5	400
8	pliers	15	1200
9	scissors	10	800
10	screwdriver	30	2400
11	spanner	20	1600
12	spatula	5	400
13	spoon	10	800
14	whisk	5	400

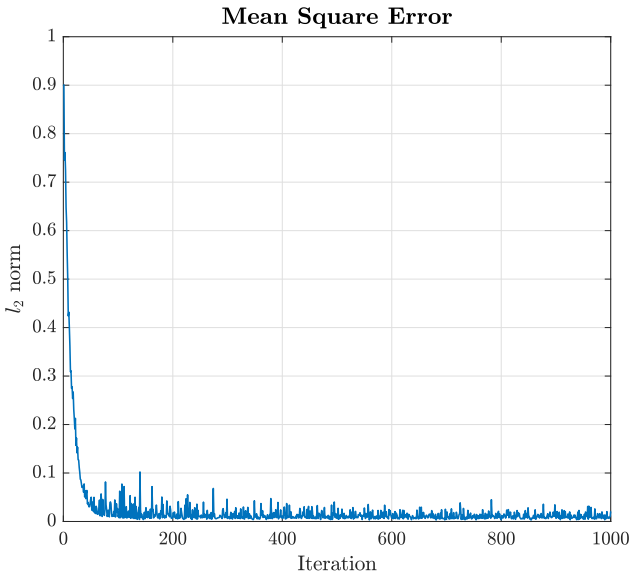


FIGURE 6. Mean square error of generated images during training.

the simulated SAR images at the SNR of 5 dB used for training and testing the model, including the noisy images I_{real} (Fig 7a-b), the ground truth I_{in} (Fig 7c-d) and the denoised images through the proposed method I_{gen} (Fig. 7e-f). Additionally, we also adopted the PSNR as a metric to quantify the reconstructed image quality, whose higher value indicates a better-quality image with less distortion or noise. It is expressed as:

$$\mathrm{PSNR}(I) = 20\log_{10}\left(\frac{\max_I}{\sqrt{\mathrm{MSE}}}\right), \tag{22}$$

where $\max_I$ is the maximum valid value of image I. After finishing the training phase, the PSNRs of the noisy images (Fig. 7a-b) are significantly improved by the proposed denoising method from 11.22 dB and 9.80 dB to 20.28 dB and 21.82 dB, respectively.

The denoising performance of the proposed approach is validated in the testing phase, which first employs only the simulated synthetic data at a SNR level of 5 dB for 14 target types with different rotation angles; none of the data were

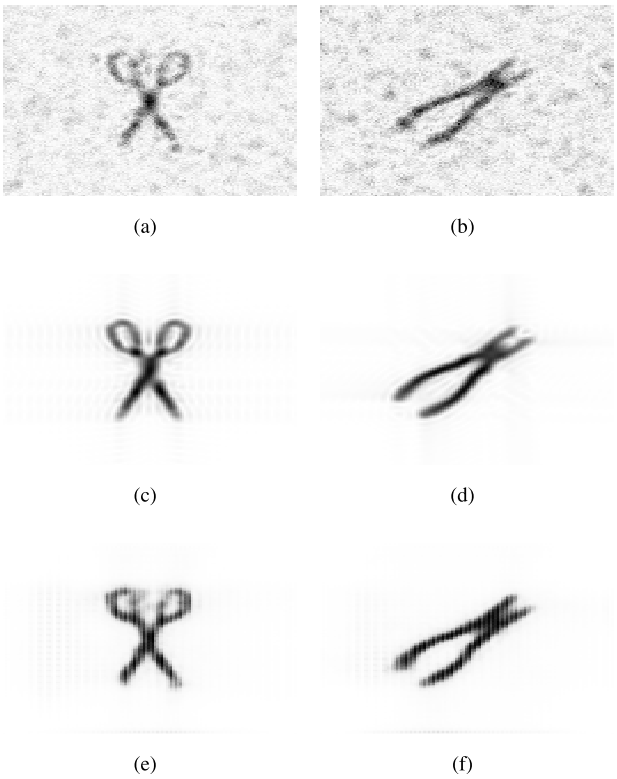
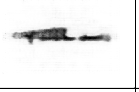
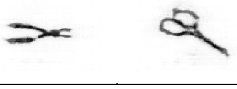
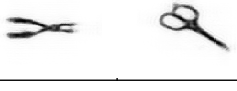
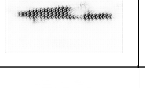
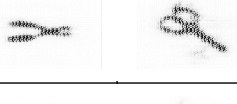
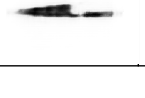
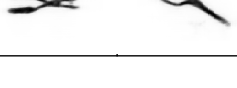


FIGURE 7. Different target simulation results of (a) scissors and (b) pliers noisy images, (c)-(d) the ground truths, and (e)-(f) the denoised images.

TABLE 4. Results of different denoising methods on simulated SAR images.

Ground truth			
Compressed sensing			
DCGAN			
DCGAN with MAE loss			
Proposed method without losses			
Proposed method with losses			

used during the training phase. The ground truth SAR images (I_{real}) of the targets are illustrated on the first row of Table 4 while the corresponding denoised images (I_{gen}) through the

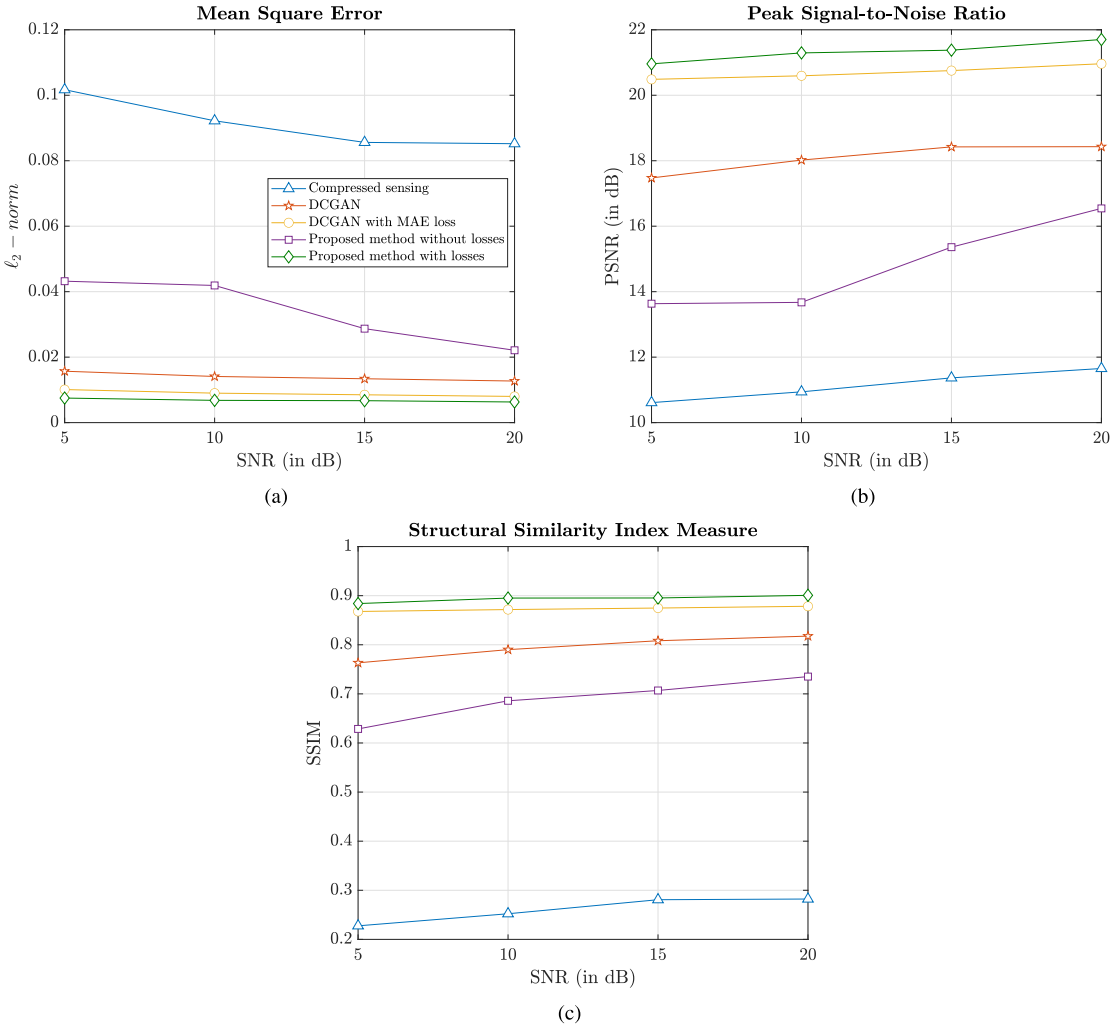


FIGURE 8. Denoising performances of different methods.

proposed method are in the last two rows with and without the involvement of different loss functions of the generator. It can be seen that the proposed approach can remove most of the noise in the SAR images as well as improve their contrast. Besides that, the MAE loss significantly contributes to the performance of the proposed model. It is shown in the 5^{th} row of Table 4 that without MAE loss, the reconstructed images still contain the artifacts, while the PSNR is also very low, resulting in the low contrast images. On the other hand, with the involvement of the MAE and SSIM losses, the model produces clean images, less noise, and higher PSNR, as in the 6^{th} row of Table 4. Furthermore, a comparison between the proposed approach and other additional denoising methods is also conducted for the sake of completeness. The compressed sensing algorithm is first taken into consideration. It also employs 10% of the original data for reconstructing an image. The random Bernoulli matrix is adopted as the observation matrix, and the optimization problem to recover data from sparse representation is solved by using the basis pursuit denoising (BPDN) solver. It can be seen in the second row of Table 4 that although the CS-BPDN approach can slightly improve the image quality, it is less effective than the proposed method as a lot of noise remains in the image results. In addition, the shapes of the targets are blurred, and they are not well reconstructed since there is lost information in the small details. Another method that is taken into comparison is the deep convolutional GAN (DCGAN), whose discriminator is a CNN network [33]. The model adopts the same dataset for both the training and testing phases as the proposed one. Although this approach achieves better denoising ability than the CS method, in general, it is less efficient than the proposed one as lower contrast images are achieved, as well as little distortion in the shapes of the targets as depicted in the third and fourth rows.

In order to achieve a detailed assessment of the performance among the employed methods, we also evaluate three image parameters, namely MSE, PSNR, and SSIM, of the denoised images, as shown in Fig. 8. These parameters are

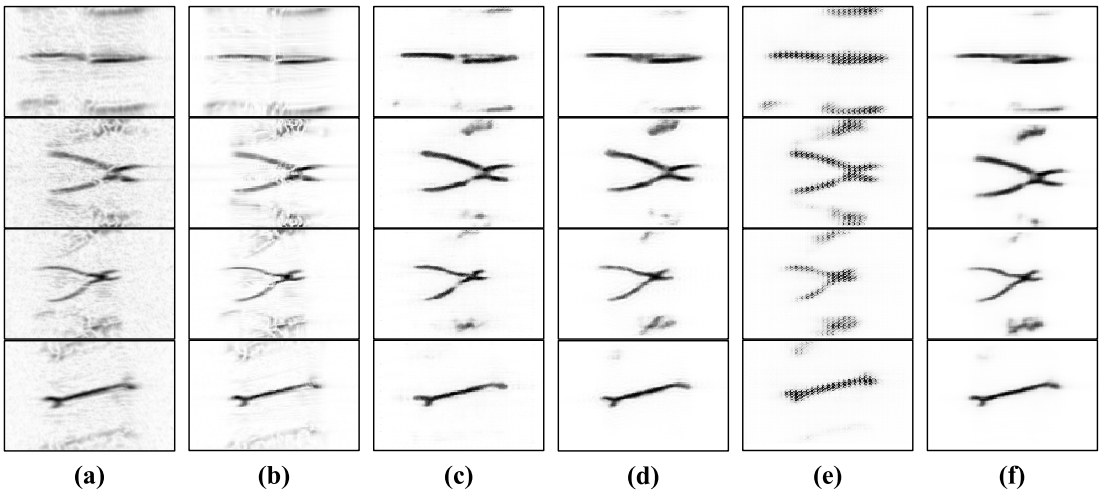


FIGURE 9. SAR image results from (a) the experiment and the corresponding denoised images by (b) compressed sensing, (c) DCGAN, (d) DCGAN with MAE loss, (e) proposed method without losses, and (f) proposed method with losses.

derived using the denoised images from different SNR levels and their corresponding ground truths. In general, denoising performance is proportional to the SNR. It can be seen that, by adopting the RbD, our method outperforms its counterparts as the images generated by the model attain less MSE (Fig. 8a) but higher PSNR (Fig. 8b) and SSIM (Fig. 8c) which results in higher quality images with less deformation and higher contrast. Again, we can also witness the significant contribution of MAE loss to denoising ability. Without the support of MAE loss, the model works worse than other DCGAN models.

The effectiveness of the model is then assessed by involving it in denoising the actual noisy SAR images reconstructed from the experiments using the same aforementioned radar and imaging parameters as in the beginning of this section. We also conducted measurements of common metallic objects, namely scissors, pliers, spanners, and knives. It is shown in Fig. 9a that the original SAR images contain a lot of artifacts due to the aliasing coming from the irregular spatial sampling intervals and the speckle noises from the environmental interference. Generally speaking, a similar phenomenon in the denoising ability of all considered methods applied to simulated data can also be witnessed in measured data. Although there are still artifacts remaining on the denoised images of all methods, the models employing RbD produce cleaner images with less noise and distortion than the others while improving contrast and image quality and keeping object information without distortion (as illustrated in Fig. 9f). On the other hand, a lot of speckle noise is still present in the images provided by the CS-BPDN algorithm (Fig. 9b) while the images produced by DCGAN models consist of object shape distortions due to losing detailed information (Fig. 9c-d).

The target recognition performance of the RbD model in the aforementioned section III-B is also trained by taking two input channels: the noisy image (I_{in}) and the ground truth (I_{real}). This classification model is simultaneously trained with the denoising model using the same input dataset of 14 target types at four SNR levels in the training phase. The test set for this discriminator is built up by taking the noisy SAR images (I_{in}) into the denoising model to get corresponding denoised images (I_{gen}), then all pairs of those images are input to the classifier to detect the recognition rate. By using both the noisy and the denoised images as two input channels, the classification rate can be improved compared to using only one input channel for the images without denoising.

TABLE 5. Recognition rate of different classification models.

Model	5 dB	10 dB	15 dB	20 dB
ResNet34	82.86%	83.65%	84.49%	86.74%
ResNet50	83.88%	85.92%	86.33%	88.16%
CNN	83.06%	85.31%	87.14%	87.55%
Proposed	84.08%	86.12%	87.14%	89.39%

To validate the efficiency of the proposed model, we take into account the involvement of three other deep learning image classification models, namely ResNet30, ResNet50, and CNN. These models are updated to adopt two input channels during the training and testing phases. The classification rates of these models are also accumulated by employing the same training and testing dataset. The recognition results of all models are depicted in Table 5 which shows that the RbD classification model achieves higher accuracy rates than other models at all SNR levels. The denoising model enhances the data as it reduces the noise and improves the SNR, achieving better effectiveness in target recognition of SAR images.

V. CONCLUSION

In this paper, we investigate a deep-learning model that combines the pix2pix model with the ResNet models to both

denoise the SAR images and improve ATR. The RbD model is built up to provide deeper learning on the two-channel input data that guides the generator to improve its performance in producing cleaner SAR images and getting over the shortcoming of the low SNR of the noisy images. The proposed model employs data, taking into account the influences of imaging parameters and noise sources, which significantly impact the quality of the SAR images. The noisy SAR images reconstructed from both simulated and measured data suffer considerable speckle noise and artifacts due to the sparse sampling and the approximation procedure of the imaging procedure. Nevertheless, the image results show that the proposed scheme has appreciably decreased the noise and artifacts on the images as well as considerably enhanced the PSNR of the input data. Additionally, the contributions of various loss functions to the model's capacity for image denoising are taken into account, and MAE loss effectively boosts the imaging scheme's performance. The ability of the proposed scheme is validated along with other well-known denoising methods, namely CS-BPDN and DCGAN. The better results of imaging metric evaluation prove that the proposed method can efficiently contribute to image-denoising missions.

A similar multi-classification RbD model is designed and concurrently trained for target recognition, which takes two input channels for training and testing, including the denoised images provided by the denoising model and the noisy image. The denoised images become a valuable source of data that significantly improves the classification accuracy rate of the ATR model. Although this pix2pix RbD scheme demonstrates its beneficial abilities, it also exhibits some drawbacks. The first one is inherent in traditional machine learning approaches, which require a large amount of data. Another disadvantage could emerge when the number of targets under measurement becomes large. In this situation, the model would not correctly generate the features as the object distribution could be blurred. Therefore, future denoising models integrating innovative despeckling methods that require less training data would be studied to overcome the limitations mentioned.

REFERENCES

[1] H. J. Kramer. (2019). *SeaSat Mission-the World's First Satellite Mission Dedicated to Oceanography*. eoPortal Directory–Satellite Missions. [Online]. Available: https://directory.eoportal.org/web/eoportal/satellite-missions/s/seasat

[2] A. Batra, V. T. Vu, Y. Zantah, M. Wiemeler, M. I. Pettersson, D. Goehringer, and T. Kaiser, "Sub-mm resolution indoor THz range and SAR imaging of concealed object," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Linz, Austria, Nov. 2020, pp. 1–4, doi: 10.1109/ICMIM48759.2020.9299034.

[3] A. Batra, T. Hark, J. Schorlemer, N. Pohl, I. Rolfes, M. Wiemeler, D. Göhringer, T. Kaiser, and J. Barowski, "Fusion of optical and millimeter wave SAR sensing for object recognition in indoor environment," in *Proc. 5th Int. Workshop Mobile Terahertz Syst. (IWMTS)*, Duisburg, Germany, Jul. 2022, pp. 1–5, doi: 10.1109/IWMTS54901.2022.9832438.

[4] M. ELsaadouny, J. Barowski, J. Jebramcik, and I. Rolfes, "Millimeter wave SAR imaging for the non-destructive testing of 3D-printed samples," in *Proc. Int. Conf. Electromagn. Adv. Appl. (ICEAA)*, Granada, Spain, Sep. 2019, pp. 1283–1285, doi: 10.1109/ICEAA.2019.8879272.

[5] B. Wu and L. He, "Multilayered circular dielectric structure SAR imaging based on compressed sensing for FOD detection in NDT," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 7588–7593, Oct. 2020, doi: 10.1109/TIM.2020.2980801.

[6] M. Fallahpour and R. Zoughi, "Fast 3-D qualitative method for through-wall imaging and structural health monitoring," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2463–2467, Dec. 2015, doi: 10.1109/LGRS.2015.2484260.

[7] P. Giordano, Z. Turksezer, M. Previtali, and M. Limongelli, "Damage detection on a historic iron bridge using satellite DInSAR data," *Struct. Health Monitor.*, vol. 21, no. 5, pp. 2291–2311, Sep. 2022, doi: 10.1177/14759217211054350.

[8] D. Oloumi, R. S. C. Winter, A. Kordzadeh, P. Boulanger, and K. Rambabu, "Microwave imaging of breast tumor using time-domain UWB circular-SAR technique," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 934–943, Apr. 2020, doi: 10.1109/TMI.2019.2937762.

[9] X. Chen, J. Guan, X. Li, and Y. He, "Effective coherent integration method for marine target with micromotion via phase differentiation and Radon-Lv's distribution," *IET Radar, Sonar Navigat.*, vol. 9, no. 9, pp. 1284–1295, Dec. 2015.

[10] Z. Chen, Z. Xiong, and D. Lewis, "Direct wave removal in anechoic chamber range imaging from planar scanned data," in *Proc. 14th Eur. Conf. Antennas Propag. (EuCAP)*, Copenhagen, Denmark, Mar. 2020, pp. 1–5, doi: 10.23919/EuCAP48036.2020.9135194.

[11] Z. Wu, L. Zhang, and H. Liu, "Generalized three-dimensional imaging algorithms for synthetic aperture radar with metamaterial apertures-based antenna," *IEEE Access*, vol. 7, pp. 59716–59727, 2019, doi: 10.1109/ACCESS.2019.2912169.

[12] M. Garcia-Fernandez, G. Alvarez-Narciandi, Y. Alvarez-Lopez, and F. L. Heras, "Comparison of subsampling strategies for UAV-mounted subsurface radar imaging systems," in *Proc. IEEE Int. Symp. Antennas Propag. USNC-URSI Radio Sci. Meeting (AP-S/URSI)*, Denver, CO, USA, Jul. 2022, pp. 1162–1163, doi: 10.1109/AP-S/USNC-URSI47032.2022.9887155.

[13] G. Álvarez-Narciandi, J. Laviada, and F. Las-Heras, "Towards turning smartphones into mmWave scanners," *IEEE Access*, vol. 9, pp. 45147–45154, 2021, doi: 10.1109/ACCESS.2021.3067458.

[14] G. Álvarez-Narciandi, J. Laviada, and F. Las-Heras, "Freehand mm-wave imaging with a compact MIMO radar," *IEEE Trans. Antennas Propag.*, vol. 69, no. 2, pp. 1224–1229, Feb. 2021, doi: 10.1109/TAP.2020.3013745.

[15] C. Vasileiou, J. Smith, S. Thiagarajan, M. Nigh, Y. Makris, and M. Torlak, "Efficient CNN-based super resolution algorithms for mmWave mobile radar imaging," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2022, pp. 3803–3807, doi: 10.1109/ICIP46576.2022.9897190.

[16] X. Zeng, Y. Ma, Z. Li, J. Wu, and J. Yang, "A near-field fast time-frequency joint 3-D imaging algorithm based on aperture linearization," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Brussels, Belgium, Jul. 2021, pp. 5163–5166, doi: 10.1109/IGARSS47720.2021.9553669.

[17] J. Wang and A. Yarovoy, "NUFFT-based range migration for 3-D imaging with irregular planar array," in *Proc. Int. Conf. Electromagn. Adv. Appl. (ICEAA)*, Granada, Spain, Sep. 2019, pp. 0671–0675, doi: 10.1109/ICEAA.2019.8879356.

[18] Y. Tai, Y. Tan, S. Xiong, Z. Sun, and J. Tian, "Few-shot transfer learning for SAR image classification without extra SAR samples," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 2240–2253, 2022, doi: 10.1109/JSTARS.2022.3155406.

[19] K. Wu, W. Cui, and X. Xu, "Superresolution radar imaging via peak search and compressed sensing," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2022.3184067.

[20] D. Jung, H. Kang, C. Kim, J. Park, and S. Park, "Sparse scene recovery for high-resolution automobile FMCW SAR via scaled compressed sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10136–10146, Dec. 2019, doi: 10.1109/TGRS.2019.2931626.

[21] K. R. Mamatha, S. A. Hariprasad, P. Saranya, S. Shahi, S. Poddar, and V. Srivastava, "A non local approach to de-noise SAR images using compressive sensing method," in *Proc. IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT)*, Bangalore, India, May 2016, pp. 1603–1606, doi: 10.1109/RTEICT.2016.7808103.

[22] W. Song, M. Li, W. Gao, D. Huang, Z. Ma, A. Liotta, and C. Perra, "Automatic sea-ice classification of SAR images based on spatial and temporal features learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 9887–9901, Dec. 2021, doi: 10.1109/TGRS.2020.3049031.

[23] L. Ding, K. Zheng, D. Lin, Y. Chen, B. Liu, J. Li, and L. Bruzzone, "MP-ResNet: Multipath residual network for the semantic segmentation of high-resolution PolSAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 4014205, doi: 10.1109/LGRS.2021.3079925.

[24] J. Wang, T. Zheng, P. Lei, and X. Bai, "Ground target classification in noisy SAR images using convolutional neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 4180–4192, Nov. 2018, doi: 10.1109/JSTARS.2018.2871556.

[25] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 1–9.

[26] C. Mao, L. Huang, Y. Xiao, F. He, and Y. Liu, "Target recognition of SAR image based on CN-GAN and CNN in complex environment," *IEEE Access*, vol. 9, pp. 39608–39617, 2021, doi: 10.1109/ACCESS.2021.3064362.

[27] L. Li, C. Wang, H. Zhang, and B. Zhang, "SAR image ship object generation and classification with improved residual conditional generative adversarial network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2020.3016692.

[28] J. Guo, B. Lei, C. Ding, and Y. Zhang, "Synthetic aperture radar image synthesis by using generative adversarial nets," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1111–1115, Jul. 2017.

[29] K. Wang, G. Zhang, Y. Leng, and H. Leung, "Synthetic aperture radar image generation with deep generative models," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 912–916, Jun. 2019.

[30] Y. Li, R. Fu, X. Meng, W. Jin, and F. Shao, "A SAR-to-optical image translation method based on conditional generation adversarial network (cGAN)," *IEEE Access*, vol. 8, pp. 60338–60343, 2020, doi: 10.1109/ACCESS.2020.2977103.

[31] J. Noa Turnes, J. D. B. Castro, D. L. Torres, P. J. S. Vega, R. Q. Feitosa, and P. N. Happ, "Atrous cGAN for SAR to optical image translation," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2020.3031199.

[32] K. Doi, K. Sakurada, M. Onishi, and A. Iwasaki, "GAN-based SAR-to-optical image translation with region information," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Waikoloa, HI, USA, Sep. 2020, pp. 2069–2072, doi: 10.1109/IGARSS39084.2020.9323085.

[33] J. Laviada, G. Álvarez-Narciandi, and F. Las-Heras, "Artifact mitigation for high-resolution near-field SAR images by means of conditional generative adversarial networks," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022, doi: 10.1109/TIM.2022.3200107.

[34] H. Kim, S. You, B. J. Jeong, and W. Byun, "Azimuth angle resolution improvement technique with neural network," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2020, pp. 1384–1387.

[35] Z. Xu, H. Li, Q. Shi, H. Wang, M. Wei, J. Shi, and Y. Shao, "Effect analysis and spectral weighting optimization of sidelobe reduction on SAR image understanding," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3434–3444, Sep. 2019, doi: 10.1109/JSTARS.2019.2925420.

[36] J. M. Lopez-Sahcnez and J. Fortuny-Guasch, "3-D radar imaging using range migration techniques," *IEEE Trans. Antennas Propag.*, vol. 48, no. 5, pp. 728–737, May 2000, doi: 10.1109/8.855491.

[37] Z. Yang and Y. R. Zheng, "Near-field 3-D synthetic aperture radar imaging via compressed sensing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 2513–2516, doi: 10.1109/ICASSP.2012.6288427.

[38] H. Weyl, "Ausbreitung elektromagnetischer Wellen über einem ebenen Leiter," *Ann. Phys.*, vol. 365, no. 21, pp. 481–500, 1919.

[39] Z. Xu, Q. Shi, Y. Chen, W. Feng, Y. Shao, L. Sun, and X. Huang, "Non-stationary speckle reduction in high resolution SAR images," *Digit. Signal Process.*, vol. 73, pp. 72–82, Feb. 2018.

[40] Z. Xu, "Wavelength-resolution SAR speckle model," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2022.3145996.

[41] D. I. Alves, B. G. Palm, M. I. Pettersson, V. T. Vu, R. Machado, B. F. Uchôa-Filho, P. Dammert, and H. Hellsten, "A statistical analysis for wavelength-resolution SAR image stacks," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 227–231, Feb. 2020.

[42] G. Álvarez-Narciandi, M. López-Portugués, F. Las-Heras, and J. Laviada, "Freehand, agile, and high-resolution imaging with compact mm-wave radar," *IEEE Access*, vol. 7, pp. 95516–95526, 2019, doi: 10.1109/ACCESS.2019.2929522.

[43] M. Garcia-Fernandez, Y. Alvarez-Lopez, and F. L. Heras, "3D-SAR processing of UAV-mounted GPR measurements: Dealing with non-uniform sampling," in *Proc. 14th Eur. Conf. Antennas Propag. (EuCAP)*, Copenhagen, Denmark, Mar. 2020, pp. 1–5, doi: 10.23919/EuCAP48036.2020.9135650.

[44] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," 2016, *arXiv:1611.07004*.

[45] J. Zhang, M. Guo, and J. Fan, "A novel generative adversarial net for calligraphic tablet images denoising," *Multimedia Tools Appl.*, vol. 79, nos. 1–2, pp. 119–140, Jan. 2020.

[46] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[47] M. E. Yanik and M. Torlak, "Near-field MIMO-SAR millimeter-wave imaging with sparsely sampled aperture data," *IEEE Access*, vol. 7, pp. 31801–31819, 2019, doi: 10.1109/ACCESS.2019.2902859.

[48] (2023). *mmWave Radar Sensors*. [Online]. Available: http://www.ti.com/sensors/mmwave/overview.html

PHAM THE HIEN received the B.E. degree in communication and networking from the Ho Chi Minh City University of Transport, Vietnam, in 2012, and the M.E. degree in information and communication engineering from Kongju National University, South Korea, in 2022, where he is currently pursuing the Ph.D. degree in information and communication engineering. His major research interests include non-destructive testing, communication and networking, wireless communication systems, synthetic aperture radar imaging, machine learning, and optimization methods.

IC-PYO HONG (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electronics engineering from Yonsei University, Seoul, South Korea, in 1994, 1996, and 2000, respectively. From 2000 to 2003, he was with the Information and Communication Division, Samsung Electronics Company, Suwon, South Korea, where he was a Senior Engineer with CDMA Mobile Research. He was a Visiting Scholar with Texas A&M University, College Station, TX, USA, in 2006; and Syracuse University, Syracuse, NY, USA, in 2012. Since 2003, he has been with the Department of Information and Communication Engineering, Kongju National University, Cheonan, South Korea, where he is currently a Professor. His research interests include numerical techniques in electromagnetics, periodic electromagnetic structures, and their applications in wireless communications.