IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 15 May 2023, accepted 30 May 2023, date of publication 5 June 2023, date of current version 9 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3282688

APPLIED RESEARCH

Resource-Efficient Range-Doppler Map Generation Using Deep Learning Network for Automotive Radar Systems

TAEWON JEONG[1], (Graduate Student Member, IEEE), AND SEONGWOOK LEE[2], (Member, IEEE)

[1]School of Electronics and Information Engineering, College of Engineering, Korea Aerospace University, Deogyang-gu, Goyang-si, Gyeonggi-do 10540, Republic of Korea
[2]School of Electrical and Electronics Engineering, College of ICT Engineering, Chung-Ang University, Dongjak-gu, Seoul 06974, Republic of Korea

Corresponding author: Seongwook Lee (seongwooklee@cau.ac.kr)

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) funded by the Korean Government [Ministry of Science and ICT (MSIT)] under Grant 2021-0-00237.

ABSTRACT In this paper, we present a deep neural network aimed at enhancing the resolution of range-Doppler (RD) maps in frequency-modulated continuous wave radar systems. The proposed deep neural network consists of an U-net-based generator and a discriminator. The low-resolution (LR) RD map is processed through the generator, resulting in a super-resolution (SR) RD map. Then, the discriminator compares the SR RD map obtained from the generator with ground truth high-resolution (HR) RD map. Finally, the generator continuously trains until the loss between the two RD maps is minimized. The efficacy of the proposed method has been verified through simulations and real-world measurements. When compared with the ground truth HR RD map, the generated SR RD map by proposed method showed only 5.24% increase in pixel-wise mean squared error and a 0.477% decrease in peak signal-to-noise ratio. Through the proposed method, target detection and tracking performance can be improved by efficiently operating radar resources.

INDEX TERMS Frequency-modulated continuous wave (FMCW), generative adversarial network (GAN), range-Doppler (RD) map, super-resolution (SR).

I. INTRODUCTION

Recently, the frequency-modulated continuous wave (FMCW) [1] has become the most commonly used waveform in automotive radar systems. The FMCW radar system determines the maximum detectable range, velocity, range resolution, and velocity resolution based on the time and frequency resources used. In other words, the radar system's target detection performance depends on its bandwidth or the number of chirps used, which are referred to as radar resources. For example, the range resolution depends on the bandwidth used by the waveform, and the velocity resolution depends on the frame time [2]. Because the chirp duration is constant, using more chirps can lead to a longer frame time. Therefore when more chirps are used in one frame, the velocity resolution of the target increases, but the number of frames that can be obtained is reduced. Conversely, if the frame time is shortened by reducing the number of chirps, more frames can be obtained, but the velocity resolution of the target decreases.

In this study, we propose a deep neural network to enhance the velocity resolution of the targets in range-Doppler (RD) maps obtained with the FMCW radar systems. The low-resolution (LR) RD map is defined as the RD map generated using a small number of chirps. The deep learning network aims to increase the velocity resolution of these RD maps as if more chirps were used. To accomplish this, we first create a database of LR RD maps that use fewer radar resources. Meanwhile, we also create a database of ground truth high-resolution (HR) RD maps that use more radar resources.

The associate editor coordinating the review of this manuscript and approving it for publication was Chengpeng Hao.

Then, we design a generative adversarial network (GAN) [3]-based network to transform the LR RD map into a super-resolution (SR) RD map. In general, the structure of the GAN consists of two main parts: a generator and a discriminator. When the generator receives input data of noise vector or arbitrary form, it converts the input into the data desired by the user. On the other hand, the discriminator numerically compares the generated data with the ground truth data to determine how similar the generated data is. After all, the training objective of the GAN is to make the generated data very similar to ground truth data, making the two data indistinguishable.

Some studies have been conducted to improve the resolution of the radar data by applying the deep neural network. For example, the Dense U-Net, which consists of convolution layers and skip connections, is used to enhance the resolution of weather radar data in [4]. In [5], authors proposed the modified residual deep neural network to enhance the direction-of-arrival resolution. In [6], authors proposed a noise-free GAN to suppress the noise and enhance the overall resolution in synthetic aperture radar images. Also, the GAN-based network for medical image translation, which consists of a CasNet generator and a patch discriminator, was used to enhance the resolution in the time-velocity plane (i.e., micro-Doppler signature) [7] and range-angle plane in [8]. In [9], a radar-SRGAN using a radar coordinate transfer module and a digital beam-forming method was proposed to improve the resolution in the range-angle plane. Moreover, a deep mutual GAN consisting of two generators and one discriminator was used to enhance the angular resolution in the radar system [10].

In our study, we employed the pix2pix [11]-based SR image generation algorithm to enhance the image resolution. Previous studies in the generation of LR radar data have utilized 20-50% of the available radar resources. Our method utilized only 12.5% of the chirps in one frame to produce the LR data. This highlights that our method can achieve a comparable enhancement of the resolution by utilizing a smaller amount of radar resources. In addition, our method does not require additional processing steps when generating low-resolution images or using them as input for deep neural network. Furthermore, the U-Net [12] structure, which can show comparable results with a relatively small dataset, was used as a pix2pix-based generator. Thus the training time can be shortened.

Finally, the proposed network's performance is evaluated through simulations and actual radar signal measurements. The SR RD map generated by the proposed network is compared with the ground truth HR RD map generated using more radar resources. Additional experiments are conducted to verify the effectiveness of the proposed method in terms of radar resource operation. Through the proposed method, the number of frames that can be measured during the same time interval increases so that the trajectory of the target can be tracked more effectively.

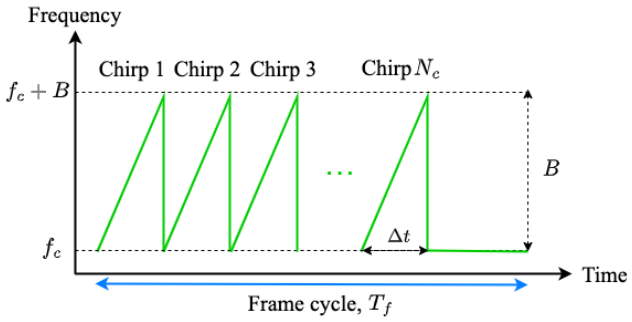

FIGURE 1. Signal transmitted from the FMCW radar system.

In summary, the major contributions of our work can be summarized as follows:

- Unlike the methods proposed in [8] and [9], our method uses less radar resources and does not require additional processing steps when generating low-resolution or high-resolution radar data.
- In contrast to the approach proposed in [10], which utilizes a deep learning network to increase resolution in the range-angle plane, our method employs a simple structure to enhance resolution in the range-Doppler plane through the use of deep learning.
- Because the same effect as using many chirps can be obtained even with smaller number of chirps, the frame time in radar can be shortened, and thus, the trajectory of a target can be identified efficiently during the same time period.

The remainder of the paper is organized as follows. In Section II, we introduce a conventional RD map generation method in the FMCW radar system. Then, the GAN-based deep learning network for generating SR RD maps is presented in Section III. Next, in Section IV, the performance of the proposed method is verified through simulations and actual measurements. Finally, we conclude this paper in Section V.

II. RD MAP GENERATION IN FMCW RADAR SYSTEM

In this section, we describe the basic principles of generating RD maps in FMCW radar systems. In addition, we introduce a conventional HR RD map generation method for building a ground truth HR RD map database.

A. BASIC RD MAP GENERATION IN FMCW RADAR SYSTEM

Because the FMCW can simultaneously obtain range and velocity information of targets, it has been widely used in automotive radar systems in recent years [1]. As shown in Fig. 1, a total of N_c chirps are transmitted sequentially. In each chirp, the frequency linearly increases over a constant time interval called the chirp duration. In Fig. 1, f_c and B represent the carrier frequency and operating bandwidth of the waveform, respectively. In addition, the entire transmission period, which is expressed as T_f in the figure, is defined as one frame.

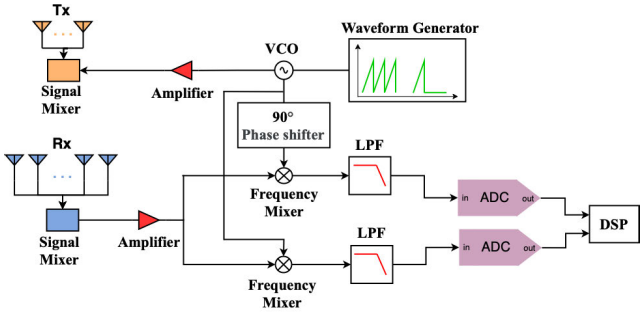

FIGURE 2. Block diagram of the FMCW radar system.

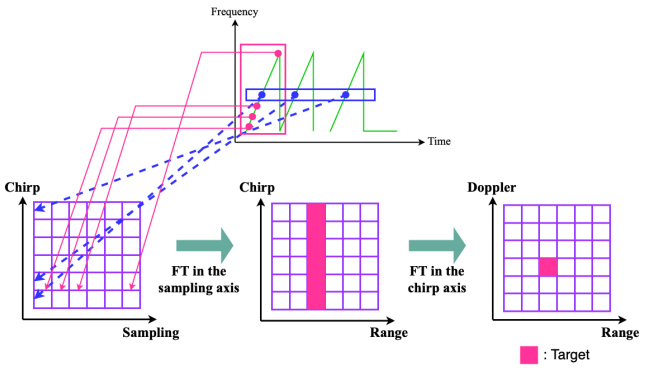

FIGURE 3. The basic process of generating the RD map.

Fig. 2 shows the overall block diagram of the FMCW radar system. The FMCW radar system is composed of a waveform generator, a voltage-controlled oscillator (VCO), amplifiers, signal mixers, transmitting and receiving antennas (Tx and Rx), a 90° phase shifter, frequency mixers, low-pass filters (LPFs), an analog-to-digital converter (ADC), and a digital signal processor. Let us assume that a total of L targets are located in the field of view of the radar system. Also, let R_l and v_l represent the relative distance to the l-th target and the relative velocity of the l-th target, respectively. When the signal transmitted from the radar is reflected by the l-th target and then returned to the radar, the time delay component is added to the received signal due to the relative distance R_l. In addition, the Doppler shift, which can be expressed as $f_d = 2v_l f_c/c$, is added to the received signal due to the relative velocity v_l, where c denotes the speed of light.

As shown in Fig. 2, the received signal passes through the amplifier to compensate for the path loss in the receiving process, and is converted into a baseband signal by frequency mixer and the LPF. Finally, the signals passed through the LPF are sampled at the ADC. The signal after passing through the ADC can be expressed as

$$T[s,\ c] = \sum_{l=1}^{L} A_l \exp\left(j2\pi\left(\frac{2R_l B}{c\Delta t}s + f_d T_f c + \frac{2f_c R_l}{c}\right)\right), \tag{1}$$

where A_l denotes the amplitude of the baseband signal corresponding to the l-th target. In addition, $s\,(s = 1,\ 2,\ \ldots,\ N_s)$ and $c\,(c = 1,\ 2,\ \ldots,\ N_c)$ in (1) denote the index of time samples in each chirp and the index of each chirp.

Then, the time-sampled signal in (1) can be expressed as a two-dimensional (2D) matrix, as shown in Fig. 3. The ranges and the Doppler frequencies for multiple targets can be obtained by applying the Fourier transform (FT) to the time-sampled baseband signal. For example, the range information of the target can be extracted by applying the FT to the sampling axis (i.e., s-axis). In addition, the Doppler shift by the target can be estimated by applying the FT to the chirp axis (i.e., c-axis). To summarize, applying the 2D FT to (1), the relative distance and velocity information of multiple detected targets can be obtained simultaneously [13].

The 2D FT for (1) can be expressed as

$$U[d,\ v] = \frac{1}{N_s N_c}\sum_{s=0}^{N_s-1}\sum_{c=0}^{N_c-1} T[s,\ c] \times \exp\left(-j2\pi\left(\frac{s}{N_s}d + \frac{c}{N_c}v\right)\right). \tag{2}$$

In this work, we define the absolute value of $U[d,\ v]$ (i.e., $|U[d,\ v]|$) as a RD map. In general, the range resolution of the FMCW is inversely proportional to the bandwidth ($R_{res} \propto \frac{1}{B}$) [14] and the velocity resolution is inversely proportional to the number of chirps ($V_{res} \propto \frac{1}{T_f}$).

B. CONVENTIONAL HR RD MAP GENERATION

In the radar system, HR frequency estimation algorithms can be used to obtain HR RD maps. In this study, we use spectrum-based frequency estimation algorithms, such as the conventional beam-forming algorithm (i.e., Bartlett) [15] and the multiple signal classification (MUSIC) algorithm [16]. The Bartlett algorithm finds a weight vector of received signal that maximizes the signal strength while keeping the noise component constant in terms of the signal-to-noise ratio (SNR). Meanwhile, the MUSIC, which is one of the subspace-based algorithms, uses the orthogonality between the signal subspace and the noise subspace.

Both of these methods use the correlation matrix of the received signal, and the correlation matrix for generating a HR RD map can be expressed as

$$R_n^C = (\mathfrak{F}^C\,(T[s,\ c]))_n^H\,(\mathfrak{F}^C\,(T[s,\ c]))_n \tag{3}$$

or

$$R_n^S = (\mathfrak{F}^S\,(T[s,\ c]))_n^H\,(\mathfrak{F}^S\,(T[s,\ c]))_n. \tag{4}$$

In (3) and (4), $\mathfrak{F}(\cdot)$ and $(\cdot)^H$ represent one-dimensional (1D) FT and hermitian operator, respectively. In addition, the superscript of $\mathfrak{F}(\cdot)$ represent the axis to where the FT is applied, as shown in Fig. 4 (a). If the correlation matrix of (3) is used, HR target detection on the range axis is possible. On the other hand, HR target detection on the Doppler axis is possible by using the correlation matrix of (4).

1) HR RD MAP GENERATION USING BARTLETT METHOD

First, we describe how to generate a HR RD map with respect to the chirp axis. Before calculating the correlation matrix as written in (4), 1D FT is performed on the sampling axis to the data obtained in (1). Then, the pseudospectrum of the Bartlett algorithm with n-th vector can be expressed as

$$P_n(v) = \frac{\mathbf{a}_A^H(v)\, R_n\, \mathbf{a}_A(v)}{\mathbf{a}_A^H(v)\, \mathbf{a}_A(v)}, \tag{5}$$

where R_n is correlation matrix obtained by (4) and $\mathbf{a}_A(v)^T = [1,\ e^{-j\pi \frac{2v}{\lambda} T_0},\ \ldots,\ e^{-j\pi \frac{2v}{\lambda}(N_c-1)T_0}]$ is the steering vector considering the chirp duration. When generating a RD map using pseudospectrum of the Bartlett, the range of the relative velocity are required. Let us assume that we divide the range of relative velocity into k number of intervals. By using k number of velocity values, k number of steering vectors can be generated as shown in Fig. 4 (b). Then we create a correlation matrix using the first vector based on the sampling axis and calculate the pseudospectrum with k number of steering vectors. The $k \times 1$ size vector will be generated as a result. If this process is repeated with all of the vectors based on the sampling axis, the RD map can be generated consequently.

2) HR RD MAP GENERATION USING MUSIC ALGORITHM

Because we assumed L number of targets and N_c number of chirps in II-A, the correlation matrix obtained through (4) becomes a size of $N_c \times N_c$ matrix. In the MUSIC algorithm, an eigenvalue decomposition is applied to the correlation matrix and separates the entire space into the signal subspace and the noise subspace. When the eigenvectors obtained through eigenvalue decomposition are arranged in the order of magnitude of the eigenvalues, the first L eigenvectors correspond to targets, and the remaining $N_c - L$ eigenvectors correspond to noise components:

$$N = \underbrace{\begin{pmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,L} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,L} \\ \vdots & \vdots & \vdots & \vdots \\ v_{N_c,1} & v_{N_c,2} & \cdots & v_{N_c,L} \end{pmatrix}}_{\text{Signal subspace}} \underbrace{\begin{pmatrix} \cdots & v_{1,N_c} \\ \cdots & v_{2,N_c} \\ \ddots & \vdots \\ \cdots & v_{N_c,N_c} \end{pmatrix}}_{\text{Noise subspace}}. \tag{6}$$

To calculate the pseudospectrum of the MUSIC algorithm, a matrix consisting of eigenvectors corresponding to the noise components is used. The pseudospectrum of the MUSIC algorithm can be expressed as

$$P_M(v) = \frac{\mathbf{a}_A(v)^H\, \mathbf{a}_A(v)}{\mathbf{a}_A(v)^H\, E_N\, E_N^H\, \mathbf{a}_A(v)}, \tag{7}$$

where E_N represents a matrix composed of eigenvectors constituting the noise subspace. Finally, the velocity of the target is determined by v that maximizes the value of the normalized pseudospectrum. In order to generate the HR RD map with the sampling axis, it is necessary to use the correlation matrix of (4) and the steering vector considering the sampling interval T_s.

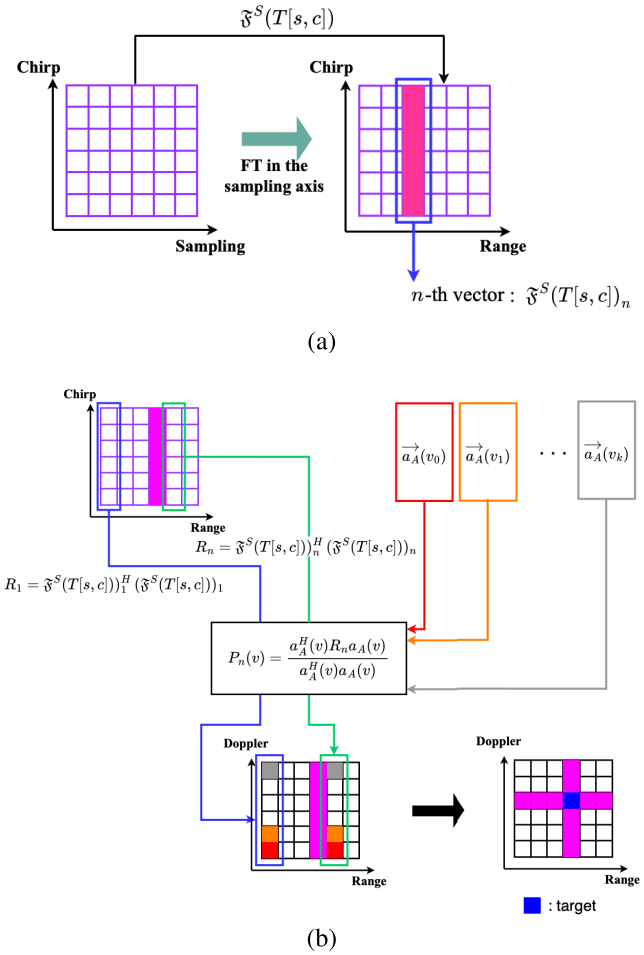

FIGURE 4. (a) Visual representation of $\mathfrak{F}(\cdot)$ operator. (b) The overall scheme of velocity estimation with the Bartlett algorithm.

III. PROPOSED NETWORK FOR SUPER-RESOLUTION IMAGE GENERATION

In this section, we explain the structure of the generator and discriminator used in the proposed deep learning network. We also explain which loss functions are used to train the generator and discriminator.

A. STRUCTURE OF THE GENERATOR

We designed a generator based on the U-net [12] that consists of contracting paths and expansion paths, which is different from the structure of the GAN proposed in [3]. The generator proposed in [3] starts from an empty noise vector and makes it look like the ground truth image, but the U-net-based generator does not. The U-net is an end-to-end (E2E) structured deep neural network [17] proposed for image segmentation in biomedical fields (e.g., finding boundaries between cells). When deep neural network-based SR imaging techniques did not exist, an algorithm that finds a target in an image must be applied first. Then, the SR imaging was completed by applying an algorithm that increases the resolution of the target image. In other words, two different algorithms had to

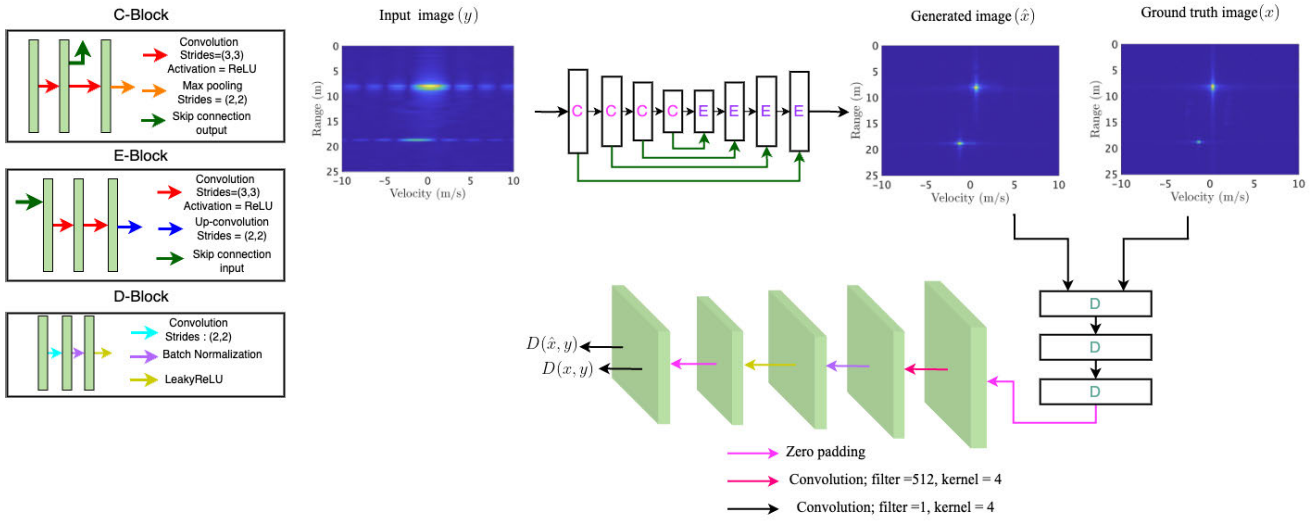

FIGURE 5. The pipeline of proposed network for enhancing the resolution of the RD map.

be used for the SR imaging tasks. However, the SR imaging tasks can be completed with a single algorithm if a deep neural network with an E2E structure is used. Also, most deep neural networks require large amounts of data because they use a data-driven approach rather than a rules-based one. The most significant advantage of the U-net is that it shows relatively accurate performance even with a small amount of data.

Fig. 5 shows the structure of the proposed network consisting of the generator and the discriminator. As shown in the figure, in the contracting path, a total of 8 convolution layers were used, and the kernel size and the number of strides in each convolution layer are 4 and 2, respectively. Also, all 8 convolution layers use leaky rectified linear unit (Leaky ReLU) functions as an activation function. Furthermore, the number of filters used in each of the eight layers is set to 64, 128, 256, 512, 512, 512, 512, and 512, respectively.

Next, the expansion path is a process opposite to the contracting path and consists of 8 layers identically. To reconstruct the image size reduced after passing through 8 consecutive convolution layers, up-convolution layers (i.e., convolution transpose layers) are used. The kernel size and the number of strides of the up-convolution layers are the same as those in the contracting path, and the number of filters used in each layer is 512, 512, 512, 512, 256, 128, 64, and 3, respectively. In addition, the rectified linear unit (ReLU) function is used as activation function for up-convolution layers in the expansion path.

B. STRUCTURE OF THE DISCRIMINATOR

As shown in Fig 5, the patch discriminator [18] is used. Unlike the pixel discriminator in [19], which compares every corresponding pixel from the generated image and the ground truth image, the patch discriminator determines the authenticity of a generated image in a specific size of patch unit from the entire image. By using a patch discriminator, determining whether an image is real or fake with fewer parameters is possible.

When training is performed by calculating the difference of each pixel, such as L1 loss, low frequency components in the image are well generated, while high frequency components are not. To accurately generate high frequency components in the image, it is necessary to focus on a local part of the image rather than focusing on the entire image. Therefore, the L1 loss and the patch discriminator are used to restore the low and high frequency components in the image, respectively.

A $256 \times 256 \times 3$ image generated through the U-net is combined with an exact same-sized ground truth image to create $256 \times 256 \times 6$ size image. The created $256 \times 256 \times 6$ size image is used as input to the patch discriminator. After the input layer, it passes through three convolution layers. In each convolution layer, the number of filters increases in the order of 64, 128, and 256. In addition, the kernel size and the number of strides are 4 and 2, and the Leaky ReLU function is used as the activation function, respectively. Finally, a convolution layer and a zero padding layer with kernel size and number of strides of 4 and 1 are used twice as a pair. As a result, we get a feature map of size $30 \times 30 \times 1$.

C. LOSS FUNCTIONS IN THE GENERATOR AND THE DISCRIMINATOR

Let $I_{LR}^{m,n}$ be the input for the proposed U-net-based generator and $G_{\text{unet}}^{m,n}$ be the generated output. Additionally, let $GT^{m,n}$ be the ground truth image of corresponding LR image (i.e., $I_{LR}^{m,n}$), where m and n represent the width and height of image. In the generator, two loss functions are defined, which can be expressed as

$$L_{\text{G1}} = -\frac{1}{mn} \sum_{m \in M} \sum_{n \in N} (O^{m,n} \log(D(G_{\text{unet}}^{m,n}))) \tag{8}$$

TABLE 1. Parameters of the radar system.

Parameter	Value
Carrier frequency, f_c	79 GHz
Bandwidth, B	3.1 GHz
Range resolution, R_{res}	4.8 cm
The number of chirps, N_c	8 or 64
The number of time samples in each chirp, N_s	256
Frame time, T_f	40 ms

and

$$L_{l1} = \sum_{m\in M}\sum_{n\in N}\left\|GT^{m,n} - G^{m,n}_{\text{unet}}\right\|_1, \tag{9}$$

where $D(\cdot)$ and $O^{m,n}$ denote the output of the discriminator and a matrix of size $m \times n$ in which all elements are 1, respectively. In other words, the first loss function in (8) is the binary cross-entropy (BCE) loss between $D(\cdot)$ and $O^{m,n}$. The second loss function in (9) is the L1 loss function. The loss of the entire generator is obtained through the weighted sum of the two loss functions. The total loss function of the generator is defined as

$$L_G = \lambda_1 L_{G1} + \lambda_2 L_{l1}, \tag{10}$$

where λ_1 and λ_2 the weights for each loss function. To determine the values of the two weights, the value of λ_1 was fixed at 1 and the value of λ_2 was gradually increased. Finally, λ_1 and λ_2 were set to 1 and 40, respectively.

For the loss function of discriminator, the sum of the two BCE losses is used. The two BCE loss functions can be expressed as

$$L_{D1} = -\frac{1}{MN}\sum_{m\in M}\sum_{n\in N} O^{m,n}\log(D(GT^{m,n})) \tag{11}$$

and

$$L_{D2} = -\frac{1}{MN}\sum_{m\in M}\sum_{n\in N}\left(1 - Z^{m,n}\right) \times \left(1 - \log\left(D(G^{m,n}_{\text{unet}})\right)\right), \tag{12}$$

where $Z^{m,n}$ denotes and a matrix of size $m \times n$ in which all elements are 0. Finally, the total loss function consisting of the two BCE loss functions is expressed as

$$L_D = L_{D1} + L_{D2}. \tag{13}$$

These two loss functions, the BCE of the ground truth image and the BCE of the generated image, are added with each other without multiplying weights.

IV. PERFORMANCE EVALUATION

A. SR RD MAP GENERATION RESULTS FROM SIMULATIONS

1) SIMULATION CONDITIONS

First, we verify the performance of the proposed network through the simulations. The simulation dataset was generated using the radar parameters in Table 1 and the signal model in (1). To increase the similarity with the radar sensor data obtained in real-world measurements, the radar cross section (RCS) of the target and signal attenuation according to the distance between the target and the radar was considered in (1). The RCS value of the target was set based on the shape of a trihedral with a side length of 20 cm, which is used in the real-world measurement. In addition, white Gaussian noise is added to (1) in consideration of the noise component generated in the experimental environment. Because the simulations were designed based on actual radar system parameters [20], the results obtained in the simulation and the actual environment show a high degree of similarity except for slight differences due to noise components. Therefore, the weights obtained through simulation can be saved (i.e., pre-trained weights) and used in the training process with the actual dataset. Using the pre-trained weights, large amounts of training data are not required, and the training time is also reduced.

In addition, the number of targets appearing in the RD map is set from 1 to 3 in the simulation. 250 RD maps were generated for each case, resulting in a total of 750 different RD maps. When generating the RD map, each target's relative distance and velocity information is set randomly between $0 \sim 25$ m and $-10 \sim 10$ m/s. By applying the data augmentation technique that flips the image horizontally, vertically, and diagonally based on the image's origin, a total of 3000 RD maps were defined as the training dataset. For the test dataset, 30 RD maps were generated for each case.

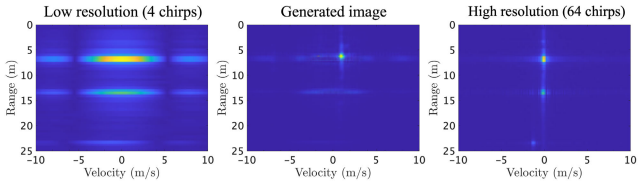

FIGURE 6. Generated SR RD map with the LR RD map ($N_c = 4$) and the HR RD map ($N_c = 64$).

2) SIMULATION RESULTS

In general, the frame time determines the resolution of the Doppler axis (i.e., the velocity axis). As mentioned in I, the chirp duration in the FMCW radar system is constant. Therefore, we varied the number of chirps to adjust the frame time. In the simulation, we verified the performance of the proposed network by changing the number of chirps for the LR and the HR RD maps. First, when the LR and the HR RD maps are generated using 4 chirps and 64 chirps, respectively, the SR RD map generated by the proposed network is shown in Fig. 6. As shown in the figure, even if the resolution of the RD map was increased through the proposed method, the target's location in the SR RD map cannot be accurately found. In addition, the target located farthest from the radar and detected with weak signal strength, disappears from the generated SR RD map.

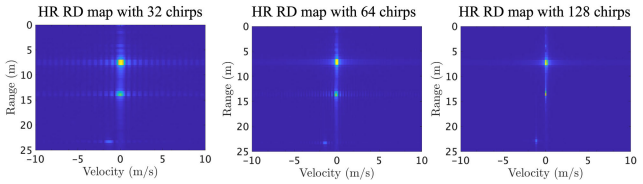

FIGURE 7. Generated RD maps with the different numbers of chirps used.

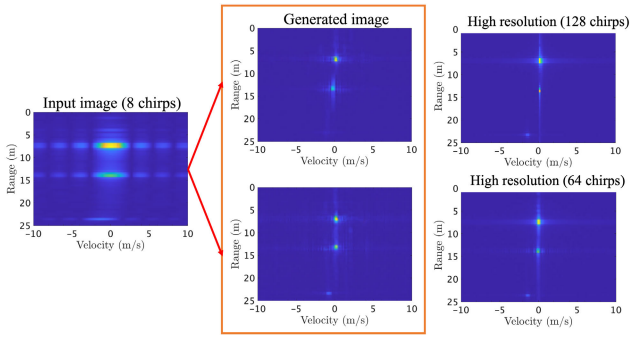

FIGURE 8. Comparison between the generated SR RD maps and the HR RD maps.

In addition, in generating the HR RD map, we performed simulations by changing the number of chirps to 32, 64, and 128 as shown in Fig. 7. As the number of chirps increases from 32 to 64, the increase in resolution is noticeable. However, there is no significant improvement in resolution when the number of chirps is increased from 64 to 128. Therefore, when generating the ground truth HR RD maps, it is appropriate to use 64 or 128 chirps instead of 32 chirps.

Moreover, we compared the outputs (i.e., generated SR RD maps) of the network trained using the HR RD maps of 64 chirps and the network trained using the HR RD maps of 128 chirps, where the LR RD map was generated from 8 chirps. As shown in Fig. 8, if the resolution of the HR RD map for training the proposed network is too high, a target with relatively weak signal strength disappears from the SR RD map generated by the proposed network. In summary, the upper and lower bounds of the number of chirps used for generating the LR and the ground truth HR RD maps are 8 and 64, respectively. Finally, Figs. 9 (a) and (b) show examples of the LR RD maps and ground truth HR RD maps when the number of targets is 1, 2, and 3, respectively.

As mentioned in the previous section, we generated a training dataset consisting of 3000 LR RD maps and ground truth HR RD maps. In addition, we also generated a test dataset consisting of 90 RD maps to validate the performance of the proposed network. When the LR RD map shown in Fig. 9 (a) is used as input and the ground truth HR RD map shown in Fig. 9 (b) is used as the ground truth image, the newly generated RD map through the proposed method is shown in Fig. 10. As shown in Fig. 10, when the LR map is used as an input, the new RD map obtained through the GAN-based network is very similar to the ground truth HR RD map.

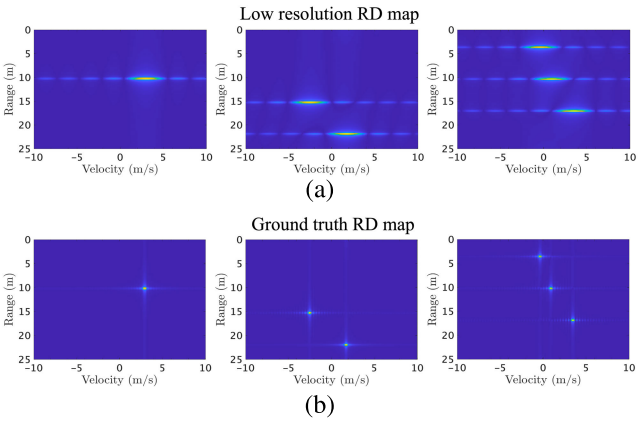

FIGURE 9. Generated RD map with simulation: (a) LR ($N_c = 8$) RD map. (b) Ground truth ($N_c = 64$) HR RD map.

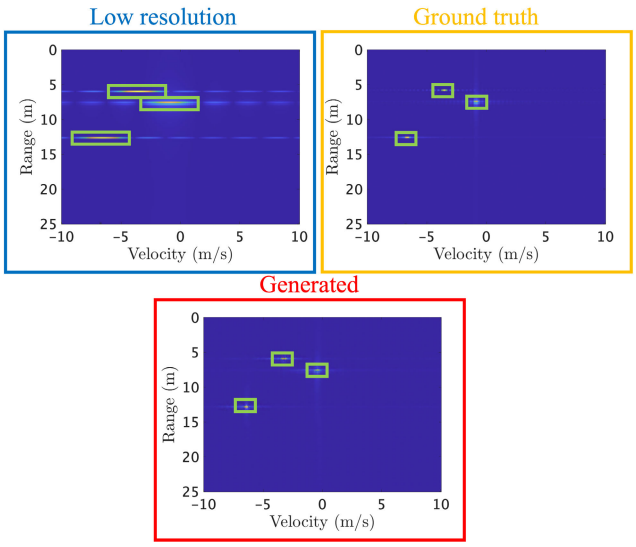

FIGURE 10. Generated RD map images with proposed network.

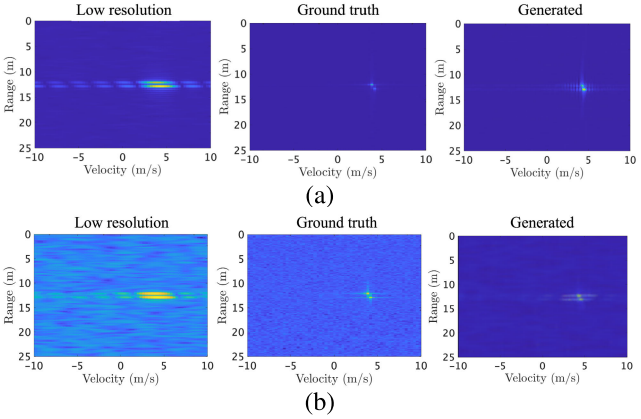

FIGURE 11. Generated RD map with proposed network: (a) When the targets are closely located. (b) When targets are in noisy environment (i.e., low SNR).

In addition to a simple simulation scenario where the targets are all separated, we have also obtained data for simulations that more closely resemble real-world measurements.

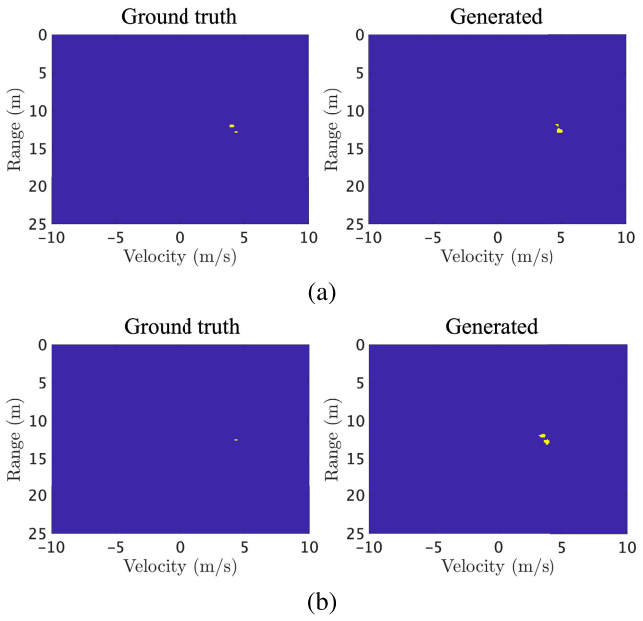

FIGURE 12. The results after applying the CFAR algorithm: (a) When the targets are closely located. (b) When targets are closely located in the low SNR scenario.

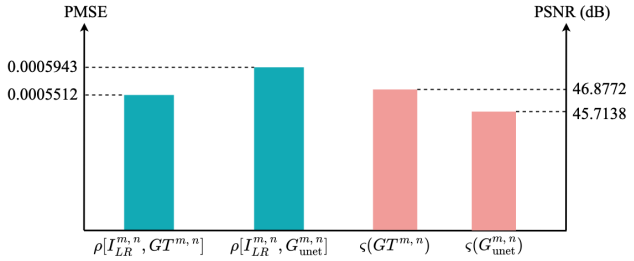

FIGURE 13. Comparison between the ground truth and the generated SR RD map.

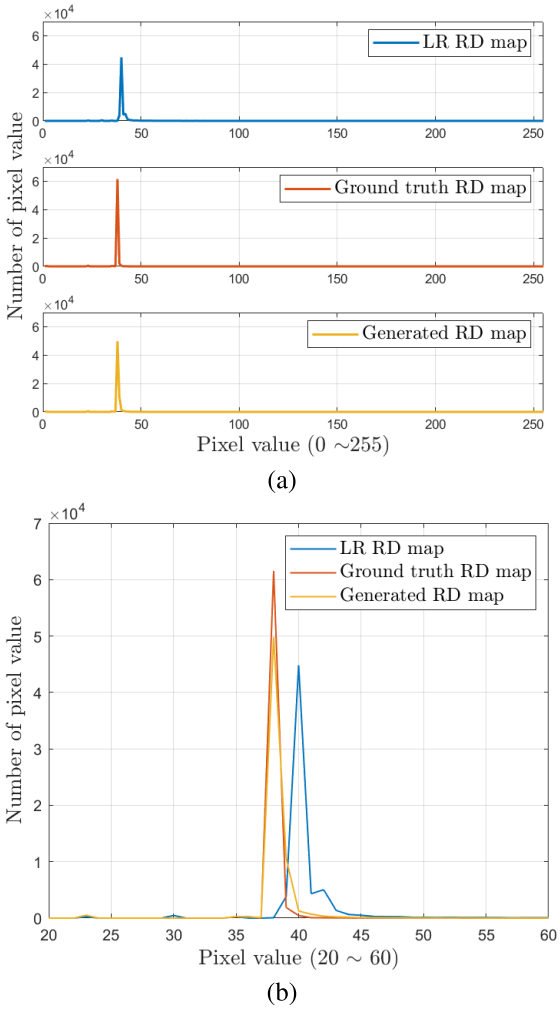

FIGURE 14. Distribution of average pixel values in the test dataset: (a) For all pixel values. (b) Pixel values between 20 to 60.

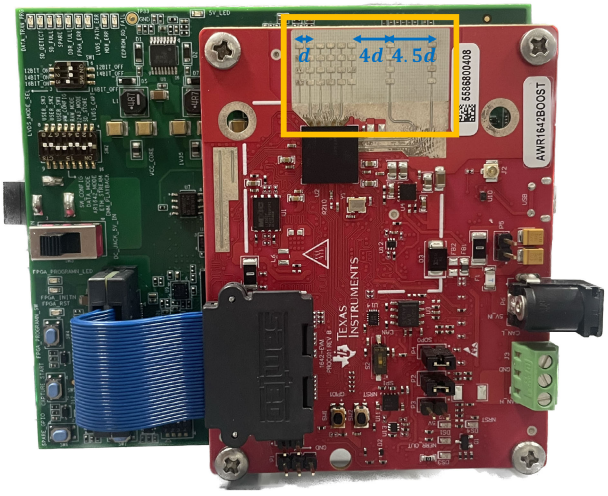

FIGURE 15. AWR1642BOOST and DCA1000EVM manufactured by Texas Instruments.

These include the cases where targets are closely located in the RD map, making it difficult to distinguish their individual areas, as well as the cases where the SNR is low. Fig. 11 (a) shows the result via the proposed method when there is partially overlapping region in the RD map due to closely located targets. Also, Fig. 11 (b) shows the result of applying the proposed method to the RD map obtained with a lower SNR value than ideal simulation scenarios.

Moreover, we verified whether the targets can be successfully detected in the RD maps generated through the proposed method in Fig. 11 by applying the constant false alarm rate (CFAR) algorithm. Figs. 12 (a) and (b) show the results of applying the CFAR to the ground truth RD maps and generated RD maps in Figs. 11 (a) and (b), respectively. In the case where the targets are partially overlapped, the CFAR algorithm can detect both targets in both the ground truth RD map and the generated RD map. On the other hand, in a noisy environment with partially overlapping targets, only one target was detected in the ground truth RD map, while both targets were detected in the generated RD map. Therefore, the proposed method can enhance the resolution of closely located targets in the RD map and effectively detect targets even when low SNR. However, when the targets are

FIGURE 16. Experimental environment for radar signal measurement.

perfectly overlapped in the RD map and cannot be distinguished, the resolution of the targets could not be enhanced even if the proposed method was used.

We quantitatively evaluate how similar the newly generated RD map is to the ground truth HR RD map. Because the GAN is not a deep neural network for classification or prediction tasks, the accuracy score cannot be used as an evaluation measure. Therefore, several image quality assessment (IQA) methods have been proposed to evaluate the similarity between images quantitatively. For example, a pixel-wise mean squared error (PMSE), peak SNR (PSNR) [21], structural similarity index measure (SSIM) [22], and visual information fidelity (VIF) [23] can be used for the IQA. Among these IQA methods, the PMSE and PSNR were used to calculate the similarity between the generated RD map and the ground truth RD map. Those two measures are defined as

$$\mathrm{PMSE} = \sum_{i \in R,\, G,\, B} \left[\frac{1}{MN} \sum_{m \in M} \sum_{n \in N} (G_{\mathrm{unet}}^{m,\,n} - GT^{m,\,n})^2 \right]_i \tag{14}$$

and

$$\mathrm{PSNR} = 10 \log_{10} \left(\frac{(255.0)^2}{\mathrm{PMSE}} \right). \tag{15}$$

In addition, we evaluated the similarity through the distribution of the pixels in various RD maps. In particular, we used average and standard deviation of the pixel values.

The overall network training process can be seen through Alg. 1 below. Before training, the dataset generated by simulations and those obtained through the actual experiments are required. The network is then trained based on the dataset generated through simulations. When the two conditions are satisfied simultaneously in the process of training, training is set to be stopped and the weight vectors are saved. The two

Algorithm 1 Training Process of the Proposed Network

Require: simulation, actual dataset $S(x,\ \hat{y},\ y)$, $A(x,\ \hat{y},\ y)$
Require: $L_{\mathrm{G}}^{(s)}$, $L_{\mathrm{D}}^{(s)}$, L_{G}, L_{D}, $O^{m,\,n}$, $Z^{m,\,n}$
Require: Epoch = n
With $S(x,\ \hat{y},\ y)$:
  for epoch ∈ n :
    if (PSNR and PMSE) > E_1% :
      Obtain generated image $D(G_{\mathrm{unet}}^{m,\,n})$
      Calculate loss functions $L_{\mathrm{G}}^{(s)}, L_{\mathrm{D}}^{(s)}$
      $\theta_{\mathrm{G}}^{(s)} \leftarrow \theta_{\mathrm{G}}^{(s)} + \omega \frac{dL_{\mathrm{G}}^{(s)}}{d\,\theta_{\mathrm{G}}^{(s)}}$
      $\theta_{\mathrm{D}}^{(s)} \leftarrow \theta_{\mathrm{D}}^{(s)} + \omega \frac{dL_{\mathrm{D}}^{(s)}}{d\,\theta_{\mathrm{D}}^{(s)}}$ **apply gradients**
    else:
      Save θ^p; $\theta^p = (\theta_{\mathrm{G}}^{(s)}, \theta_{\mathrm{D}}^{(s)})$
  end
end
With $A(x,\ \hat{y},\ y)$:
  for epoch ∈ Epochs :
    if epoch = 1 :
      $\theta_{\mathrm{G}},\ \theta_{\mathrm{D}} \leftarrow \theta^p$ **apply pretrained gradients**
      Obtain generated image $D(G_{\mathrm{unet}}^{m,\,n})$
      Calculate loss functions $L_{\mathrm{G}}, L_{\mathrm{D}}$
    else:
      if (PSNR and PMSE) > E_2% :
        Obtain generated image $D(G_{\mathrm{unet}}^{m,\,n})$
        Calculate loss functions $L_{\mathrm{G}}, L_{\mathrm{D}}$
        $\theta_{\mathrm{G}} \leftarrow \theta_{\mathrm{G}} + \omega \frac{dL_{\mathrm{G}}}{d\,\theta_{\mathrm{G}}}$
        $\theta_{\mathrm{D}} \leftarrow \theta_{\mathrm{D}} + \omega \frac{dL_{\mathrm{D}}}{d\,\theta_{\mathrm{D}}}$ **apply gradients**
      else:
        Save (PSNR and PMSE)
  end
end

conditions can be expressed as follows:

$$\frac{\rho[I_{LR}^{m,\,n}, G_{\mathrm{unet}}^{m,\,n}] - \rho[I_{LR}^{m,\,n}, GT^{m,\,n}]}{\rho[I_{LR}^{m,\,n}, GT^{m,\,n}]} < \frac{E_1}{100} \tag{16}$$

and

$$\frac{\varsigma(GT^{m,\,n}) - \varsigma(G_{\mathrm{unet}}^{m,\,n})}{\varsigma(G_{\mathrm{unet}}^{m,\,n})} < \frac{E_1}{100}. \tag{17}$$

$\rho[\cdot\,,\ \cdot]$ represents the PMSE value between two variables in $\rho[\cdot\,,\ \cdot]$, and $\varsigma(\cdot)$ represents the PSNR value, respectively.

After the network is trained with simulation dataset, the network is retrained with actual dataset. When training the network with the actual dataset, the saved weight vectors are used in the first epoch. From the second epoch, the weight vectors that are updated during the training process are used. Finally, when the two conditions mentioned above are satisfied again, the training is set to be stopped. However, there is one thing that has changed. Because the network is retrained based on the pre-trained weights, E_1 value is substituted with smaller value, E_2. In this paper, we set E1 and E2 as 10 and 5. After the network was trained with

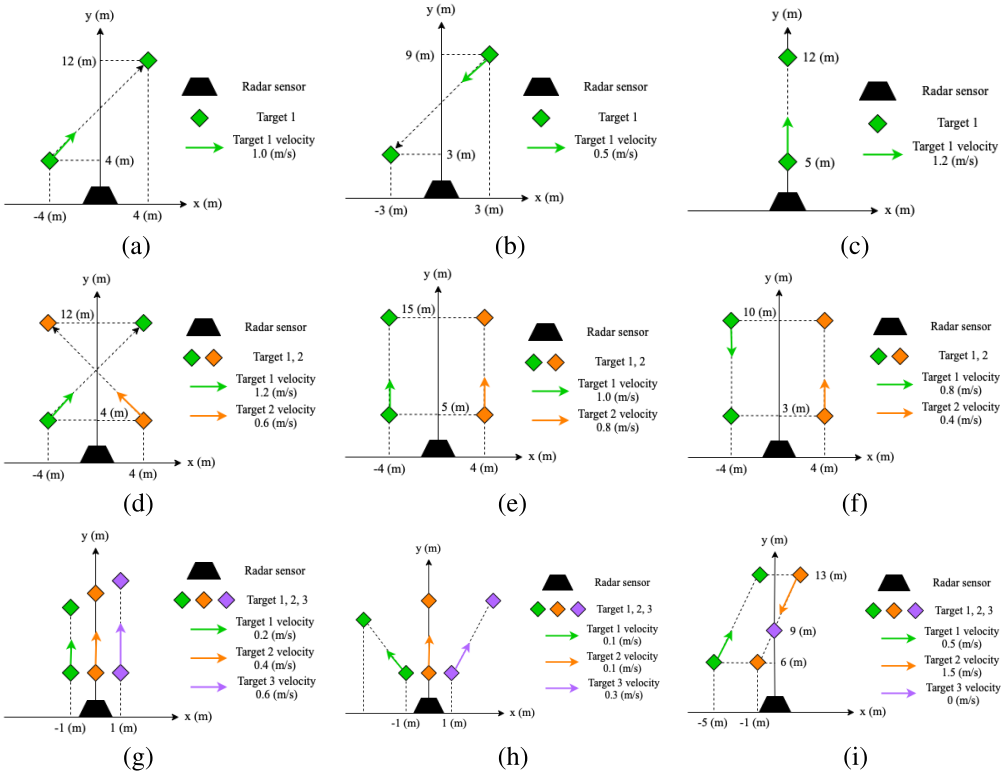

FIGURE 17. 9 experimental scenarios for the performance evaluation.

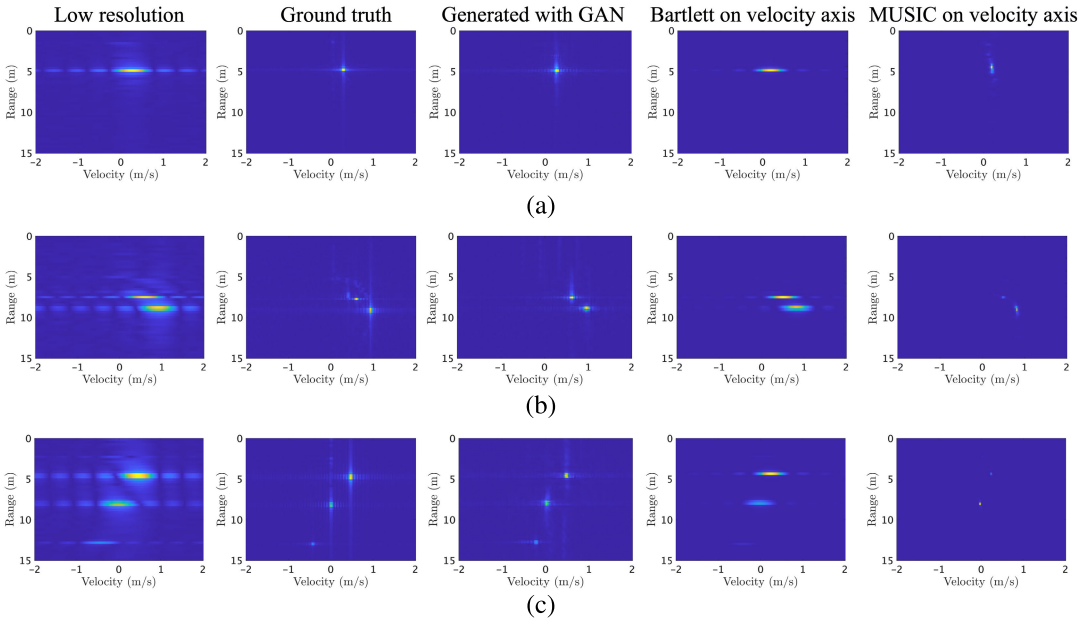

FIGURE 18. Comparison between generated RD map from the conventional HR imaging algorithms and the proposed deep learning-based method.

the training dataset generated through simulations, the network was verified with test dataset generated through simulations. As a measure of the verification, the PMSE value between the LR RD map and ground truth HR RD map (i.e., $\rho[I_{LR}^{m,n}, GT^{m,n}]$) and the PMSE value of the LR RD map and the deep learning-based SR RD map (i.e., $\rho[I_{LR}^{m,n}, G_{\text{unet}}^{m,n}]$) were used. In addition as another measure of the verification, the PSNR value of ground truth HR RD map (i.e., $\varsigma(GT^{m,n})$)

and the PSNR of the deep learning-based SR RD map (i.e., $\varsigma(G_{\text{unet}}^{m,n})$) were used. As shown in Fig. 13, when comparing the deep learning-based SR RD map with the ground truth, the PMSE value increased by 7.819% and the PSNR decreased by 2.481%.

Figs. 14 (a) and (b) shows the distribution of average pixel values in the test dataset. The x and y-axis represent the range of pixel value in the RD map image and the number of each pixel value, respectively. For pixel values below 50, the distribution of the LR RD map is more dispersed than that of the ground truth RD map. Also, the pixel distribution of the generated RD map follows the distribution of the ground truth RD map. The average pixel values of the LR RD map, ground truth RD map, and the generated RD map are 42.386, 37.849, and 38.402, respectively. Also, the standard deviations are 13.972, 3.287, and 5.98. Therefore, we verified that the generated RD maps are highly similar to the ground truth RD maps.

B. SR RD MAP GENERATION RESULTS FROM ACTUAL MEASUREMENTS

1) EXPERIMENTAL ENVIRONMENTS

To show the effectiveness of our proposed deep neural network, we conducted actual measurements using the AWR1642BOOST board, which was produced by Texas Instrument (TI) [20]. We used the AWR1642BOOST board connected with a DCA1000EVM, as shown in Fig. 15. The AWR-1642 radar module has two transmit antenna elements and four receiving antenna elements. The physical size of the antenna mounted on the board is 30 × 19 mm. Also, the spacing between the transmitting antenna elements, the spacing between the transmitting and receiving antenna elements, and the spacing between the receiving antenna elements are d, $4d$, and $4.5d$, respectively. Moreover, the 3dB beam width in the azimuth direction is 70 degrees, and the 3dB beam width in the elevation direction is 30 degrees. Because our goal is to enhance the resolution for target detection in the RD map, it can be sufficient to acquire sensor data using only one transmit antenna and one receiving antenna. In summary, one transmit and one receiving antenna element were used among the multiple-input and multiple-output antenna system [24]. The data acquired through the AWR-1642 radar module can be saved as a binary file through the DCA1000EVM board [25]. Then, the data stored as binary files can be read through TI-provided code implemented in Matlab or Python.

Fig. 16 shows the experimental environment for the radar signal measurements. In the experiment environment, radar sensor data were obtained through the 9 different scenarios as shown in Fig. 17. As mentioned in Section IV-A, the trihedral corner reflectors with a side length of 20 cm were used as targets in the measurement. Various RD maps were obtained because the moving direction and velocity of the targets are different in each scenario. For each scenario, we obtained 128 frames of radar data and of which the first 110 frames were used as the training dataset, and the remaining 18 frames were used as the test dataset.

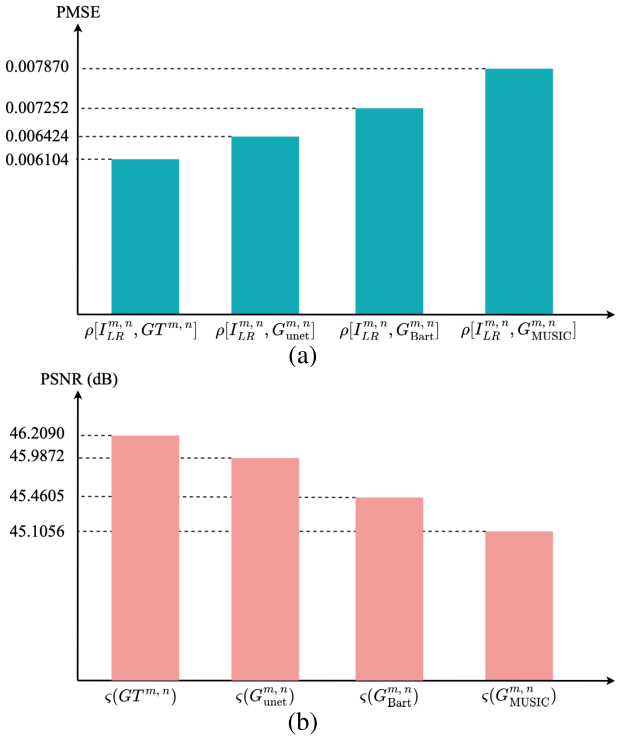

FIGURE 19. Comparison between the proposed and the conventional HR imaging methods: (a) PMSE and (b) PSNR values.

2) EXPERIMENTAL RESULTS

By changing the number of chirps used as mentioned in Section IV-A, a total of 990 pairs of LR and ground truth HR RD maps were used as the training dataset. Also, 162 pairs of LR and ground truth HR RD maps were used as the test dataset. Fig. 18 shows the LR RD map from the leftmost, the ground truth HR RD map, the generated SR RD map by the proposed deep learning network and HR RD maps generated by applying the Bartlett and MUSIC algorithms to the Doppler axis. When qualitatively evaluating the results, the RD maps generated through the MUSIC algorithm and the proposed deep neural network show the most similarity to the ground truth image. Applying Bartlett reduces sidelobes but increases the resolution very slightly. In addition, although the RD map from the MUSIC algorithm exhibits a high resolution comparable to the ground truth HR RD map, a target is often not detected in the RD map as shown in Fig. 18 (c). The target disappears because the MUSIC algorithm's performance is highly sensitive to the target's SNR. In addition, the significant disadvantage of the MUSIC algorithm is that the number of targets must be less than the number of chirps used in one frame [26]. Therefore, there is a limit to applying the MUSIC algorithm under the condition when the number of targets exceeds the number of radar resources.

Fig. 19 quantitatively compares the similarity between the RD maps generated by the existing methods and the RD maps generated based on the proposed deep neural network. In addition, $G_{\text{Bart}}^{m,n}$ and $G_{\text{MUSIC}}^{m,n}$ represent the HR image output with Bartlett and MUSIC algorithm. To compare the

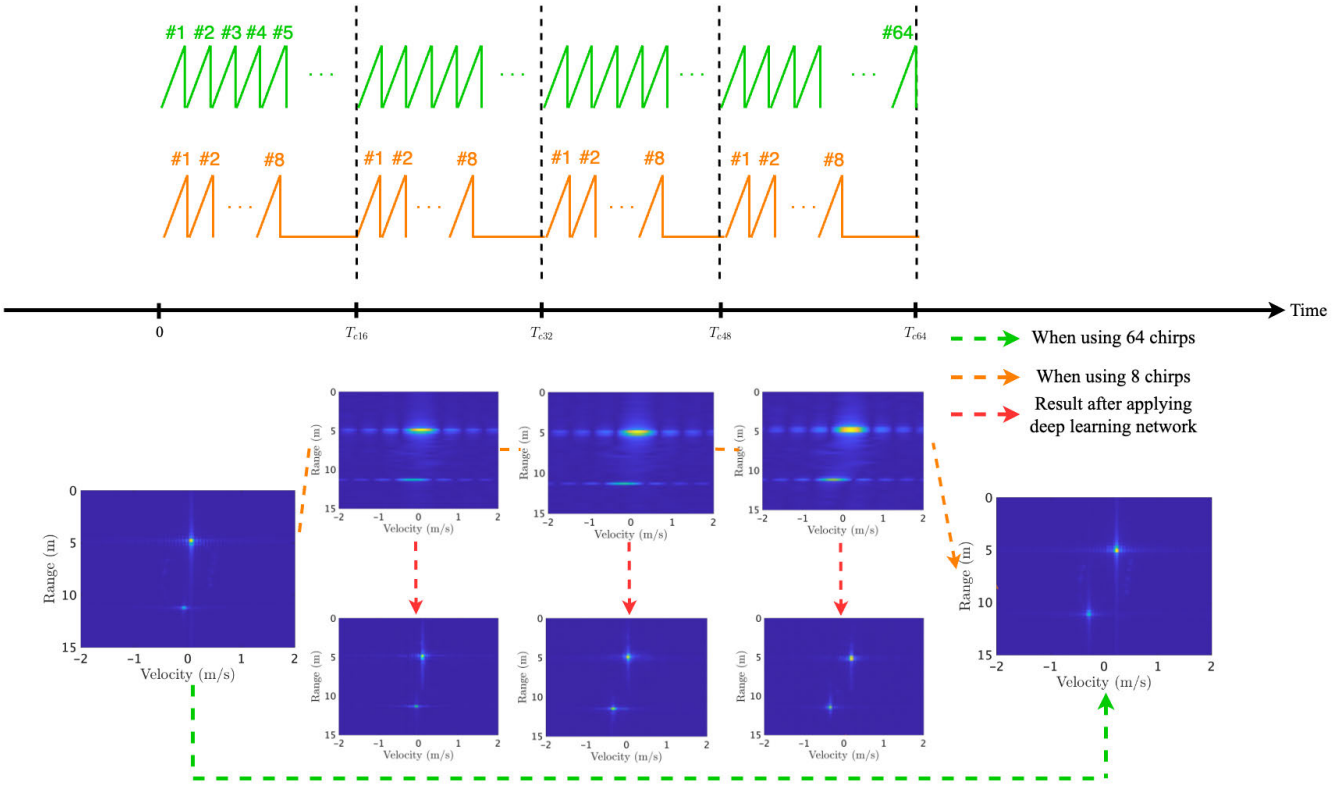

FIGURE 20. An example of efficient radar resource management.

proposed deep neural network-based method with the existing SR imaging method, the PMSE value between the LR RD map and the ground truth HR RD map and the PSNR value of the ground truth HR RD map was set as reference values. In terms of the PMSE, the deep learning-based method, the Bartlett-based method, and the MUSIC-based method increased by 5.24%, 18.8%, and 28.9% compared to the reference value, respectively. In terms of the PSNR, the deep learning-based method, the Bartlett-based method, and the MUSIC-based method decreased by 0.477%, 1.619%, and 2.387% compared to the reference value, respectively. Furthermore, the average pixel values of the LR RD map, Ground turth RD map, and generated RD map were 45.384, 38.912, and 39.715, respectively. The standard deviation of pixel values were 15.857, 4.291, and 6.118. Consequently, the proposed deep learning-based method enhanced the resolution closer to the ground truth image than the existing HR imaging algorithms.

C. EFFICIENT MANAGEMENT OF RADAR RESOURCES

The following experiments were conducted to emphasize the resource-efficient aspect of our proposed method. The number of chirps required to generate the range-Doppler map was reduced to 12.5%, and the period of each frame was also decreased to 25%, as shown above in Fig. 20. Because the period of one frame is shortened to 25%, 4 frames can be measured within the same time when 8 chirps are used. In other words, because the measurement period for the target is reduced to 25%, it is possible to know the trajectory of the moving target more precisely. As shown in Fig. 20, when the detection result for a moving target is obtained using 64 chirps, the transition from the first frame to the second frame follows the green dotted line. At the same time, if we use 8 chirps to acquire detection results, 3 more frames can be obtained along the orange dotted line. If the reduced velocity resolution using only 8 chirps is regenerated into the SR RD map using the proposed deep neural network, the same effect as obtaining 4 frames with 64 chirps can be achieved. Finally, the red lines in the figure show the result of converting the LR RD map into the SR RD map.

V. CONCLUSION

In this paper, we proposed a deep learning-based network for enhancing the resolution of the LR RD maps in the FMCW radar system. The proposed network consists of the U-net-based generator and the discriminator. When the U-net receives an image as input, it divides the image into feature maps. Then it increases the resolution of the image in the process of reconstructing the image back to its original size. In addition, the discriminator evaluates the performance of the generator by comparing the resolution enhanced image with the ground truth image. The performance of the proposed network was verified through simulations and actual measurements. To evaluate the similarity between the

RD map generated by the proposed network and the ground truth HR RD map, the PMSE, and PSNR were calculated. Compared with the conventional HR imaging algorithms (i.e., Bartlett and MUSIC algorithms), the PMSE value decreased by 12.9% and 22.5%, respectively, and the PSNR value increased by 1.1% and 1.9%, respectively, in our proposed method. Based on these measures, we confirmed that the RD map generated by the proposed method showed a higher resemblance with the ground truth HR RD map than the RD maps generated from the conventional HR imaging algorithms. In addition, additional experiments were conducted to verify the performance of the proposed method in terms of radar resource operation, and the target tracking performance could be improved through the proposed method. Although the proposed deep neural network-based technique was trained with the data obtained through the automotive radar system, this does not represent that the proposed technique is limited to the automotive radar system. The proposed method can be applied to all radar systems capable of obtaining RD map data to enhance the resolution of targets.

REFERENCES

[1] V. Winkler, "Range Doppler detection for automotive FMCW radars," in *Proc. Eur. Radar Conf.*, Oct. 2007, pp. 166–169.

[2] S. Rao, *Introduction to mmWave Sensing: FMCW Radars*. Dallas, TX, USA: Texas Instruments, 2017. [Online]. Available: https://www.ti.com/video/series/mmwave-training-series.html

[3] I. Goodfellow, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2014, pp. 2672–2680.

[4] A. Geiss and J. C. Hardin, "Radar super resolution using a deep convolutional neural network," *J. Atmos. Ocean. Technol.*, vol. 37, no. 12, pp. 2197–2207, Dec. 2020.

[5] M. Alizadeh, M. Chavoshi, A. Samir, A. M. Hegazy, A. Bahri, M. Basha, and S. Safavi-Naeini, "Experimental deep learning assisted super-resolution radar imaging," in *Proc. 18th Eur. Radar Conf. (EuRAD)*, Apr. 2022, pp. 153–156.

[6] F. Gu, H. Zhang, C. Wang, and F. Wu, "SAR image super-resolution based on noise-free generative adversarial network," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 2575–2578.

[7] K. Armanious, C. Jiang, M. Fischer, T. Küstner, T. Hepp, K. Nikolaou, S. Gatidis, and B. Yang, "MedGAN: Medical image translation using GANs," *Computerized Med. Imag. Graph.*, vol. 79, Jan. 2020, Art. no. 101684.

[8] K. Armanious, S. Abdulatif, F. Aziz, U. Schneider, and B. Yang, "An adversarial super-resolution remedy for radar design trade-offs," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2019, pp. 1–5.

[9] H. Cho, W. Kim, S. Choi, M. Eo, S. Khang, and J. Kim, "Guided generative adversarial network for super resolution of imaging radar," in *Proc. 17th Eur. Radar Conf. (EuRAD)*, Jan. 2021, pp. 144–147.

[10] H. Xing, M. Bao, Y. Li, L. Shi, and M. Xing, "Deep mutual GAN for life-detection radar super resolution," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[11] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976.

[12] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Oct. 2015, pp. 234–241.

[13] S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: A review of signal processing techniques," *IEEE Signal Process. Mag.*, vol. 34, no. 2, pp. 22–35, Mar. 2017.

[14] M. N. Cohen, "An overview of high range resolution radar techniques," in *Proc. NTC Nat. Telesyst. Conf.*, Mar. 1991, pp. 107–115.

[15] M. S. Bartlett, "Smoothing periodograms from time-series with continuous spectra," *Nature*, vol. 161, no. 4096, pp. 686–687, May 1948.

[16] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[17] J. H. Saltzer, D. P. Reed, and D. D. Clark, "End-to-end arguments in system design," in *Proc. Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Paris, France, Apr. 1981, pp. 509–512.

[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[19] D. Pathak, P. Krähenbühl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2536–2544.

[20] *AWR1642 Single-Chip 77-79-GHz FMCW Radar Sensor*, Texas Instrum., Dallas, TX, USA, 2021. [Online]. Available: https://www.ti.com/lit/ds/symlink/awr1642.pdf

[21] C.-Y. Yang, C. Ma, and M.-H. Yang, "Single-image super-resolution: A benchmark," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2014, pp. 372–386.

[22] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[23] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[24] B. Donnet and I. Longstaff, "MIMO radar, techniques and opportunities," in *Proc. Eur. Radar Conf.*, Sep. 2006, pp. 112–115.

[25] *DCA1000EVM Data Capture Card*, Texas Instrum., Dallas, TX, USA, 2019. [Online]. Available: https://www.ti.com/tool/DCA1000EVM

[26] L. Osman, I. Sfar, and A. Gharsallah, "Comparative study of high-resolution direction-of-arrival estimation algorithms for array antenna system," *J. Res. Rev. Wireless Commun.*, vol. 2, no. 1, pp. 72–77, 2012.

TAEWON JEONG (Graduate Student Member, IEEE) is currently pursuing the integrated B.S. and M.S. degree in electronics and information engineering from Korea Aerospace University (KAU), Goyang-si, Gyeonggi-do, Republic of Korea. His research interests include radar signal processing, such as radar clutter suppression, deep learning-based target detection and tracking, and improved angle estimation.

SEONGWOOK LEE (Member, IEEE) received the B.S. and Ph.D. degrees in electrical and computer engineering from Seoul National University (SNU), Seoul, Republic of Korea, in February 2013 and August 2018, respectively. From September 2018 to February 2020, he was a Staff Researcher with the Machine Learning Laboratory, AI & SW Research Center, Samsung Advanced Institute of Technology (SAIT), Gyeonggi-do, Republic of Korea. He was an Assistant Professor with the School of Electronics and Information Engineering, College of Engineering, Korea Aerospace University (KAU), Gyeonggi-do, from March 2020 to February 2023. Since March 2023, he has been an Assistant Professor with the School of Electrical and Electronics Engineering, College of ICT Engineering, Chung-Ang University (CAU), Seoul. He has published more than 100 papers on signal processing for radar systems. His research interests include radar signal processing techniques, such as enhanced target detection and tracking, target recognition and classification, clutter suppression and mutual interference mitigation, and artificial intelligence algorithms for radar systems.