Artificial Intelligence Review
https://doi.org/10.1007/s10462-023-10469-5

Application of deep generative networks for SAR/ISAR: a review

Jiawei Zhang[1] · Zhen Liu[1] · Weidong Jiang[1] · Yongxiang Liu[1] · Xiaolin Zhou[1] · Xiang Li[1]

© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract
Military, agricultural, and urban planning have all made extensive use of SAR/ISAR in the realm of remote sensing. SAR/ISAR images are more capable of identifying the details of the targets than optical images and can be taken in any condition. Due to the challenges associated with SAR/ISAR imaging, the lack of data causes many jobs relying on data-driven deep learning algorithms to perform less than satisfactorily. Cropping, rotation, and other procedures are examples of classic data augmentation techniques now in use, although they do not fundamentally differ from basic replication and cannot increase the model's stability and robustness. Deep generative models are used to generate SAR/ISAR images, which is a more efficient way than the conventional ones. The generation techniques are outlined and organized depending on the application fields in this review, including SAR/ISAR data augmentation (26 papers), SAR/ISAR image translation (29 papers), SAR/ISAR image enhancement (22 papers), azimuth interpolation (9 papers), and deceptive jamming (1 paper). The connected works are then summarized based on several deep generative models. 87 linked studies and 5 associated survey papers from 2017 to 2022 are compiled in this review. Finally, the summarized works are systematically analyzed. There are 27 papers using MSTAR for image generation, which is the mostly applied dataset. For evaluation, the combination of SSIM and PSNR is applied most widely (32.19%). In conclusion, this review offers fresh perspectives on the direction in which deep generative models for SAR/ISAR image generation are headed. The cutting-edge methods outlined in this paper are also available to researchers in other domains.

Keywords Synthetic aperture radar image · Deep learning · Generative adversarial network · Image generation · Artificial intelligence

✉ Zhen Liu
zhen_liu@nudt.edu.cn

[1] College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

Published online: 28 March 2023

1 Introduction

1.1 Basic knowledge of (inverse) synthetic aperture radar

Radar, first used to identify enemy ships and planes in bad weather and at night, has also advanced significantly with the rapid growth of remote sensing (RS) and digital technology (Perry et al. 1999). By processing the Doppler frequency shift, in 1951, Carl Wiley increased the angular resolution of radar, allowing for the capturing of a two-dimensional image of the earth's surface (Wiley 1985). In order to mimic the imaging performance of a longer antenna, synthetic aperture radar (SAR), which uses signal processing technologies, is developed. Multiple electromagnetic waves can be delivered to the same target area by SAR-equipped satellites during flight, and the echoes of each signal can be analyzed to create a composite signal (Solimene et al. 2014). As a result, SAR can be thought of as a device that can combine several radar antennas into one long radar antenna and then record the resulting image. Inverse synthetic aperture radar (ISAR), which is based on SAR imaging technology and has been developed with the development of RS technology, can produce two-dimensional and three-dimensional high-resolution radar images that reflect the size, shape, structure, azimuth, and other details of the target (Prickett and Chen 1980). ISAR is typically used to image uncooperative moving targets on a fixed platform, while SAR is typically used to image fixed targets and scenery on a moving platform, and its own motion rules are known. The comparison between a SAR image and a conventional optical image is presented in Fig. 1.

SAR/ISAR has been widely used in the field of remote sensing and offers three key advantages over traditional optical imaging technology:

- Using traditional optical imaging technology, the target surface's reflected light must be captured. However, SAR/ISAR has a built-in lighting source and can operate regularly in low-light conditions (Li et al. 2006).
- When confronted with severe weather, such as an overcast day and thunderstorms, optical imaging systems are unable to get ground target images beneath the cloud, however, SAR/ISAR can image through the cloud with minimal distortion (Yates et al. 2006).



(a) (b)

Fig. 1 An example of an optical image and corresponding SAR image. **a** The optical image; **b** the SAR image. The image pair is selected from WHU-OPT-SAR dataset (Li et al. 2022c), which contains 100 optical images of pixels and SAR images in the same area. The optical images are collected by GF-1 with a ground resolution of 2 ms, and the SAR images are collected by GF-3 fine strip II (FS II) with a sampling resolution of 5 ms

– The optical scattering energy of objects is different from the electromagnetic wave scattering energy of SAR/ISAR because the wavebands of the two systems are different. SAR/ISAR can therefore compensate for the shortcomings of the optical imaging system in specific jobs, and vice versa. As a result, they work best together, and occasionally SAR/ISAR may differentiate the characteristics of ground targets better than optical images (Green 1998).

As a microwave active observation system, SAR/ISAR has the ability to observe the target in all conditions and at all times, which is one of the main reasons why it may be used broadly (Sharifi 2020). SAR/ISAR, which is a high-resolution observation system, can provide correct image data of the ground surface by not only obtaining the distance delay information of ground objects but also achieving the high resolution of scattering spots in the distance and azimuth dimensions (Bi et al. 2019). SAR/ISAR data contains a wealth of information on the characteristics of the ground objects because the scattering properties of the ground objects are intimately tied to the particular wave band and polarization mode. SAR/ISAR can provide moving target recognition and detection based on the differences in Doppler characteristics between moving targets (such as ships, vehicles, icebergs, ocean currents, etc.) and stationary objects (Tomiyasu 1978).

Due to its all-weather and all-time imaging capabilities, SAR/ISAR is widely used in both the military and civic sectors. Military applications include strike effect analysis, military reconnaissance, and military surveying and mapping. Civil applications include disaster monitoring (Yamaguchi 2012), topographic mapping (Heygster et al. 2009), vegetation growth observation (Liu et al. 2019), land and sea traffic monitoring (Renga et al. 2018), offshore oil spill detection (Topouzelis 2008), ocean current (Kudryavtsev et al. 2014), glacier movement monitoring (Fan et al. 2019), etc. Due to the low resolution and the adversary's countermeasures when conducting reconnaissance and surveillance on the target, it is difficult to obtain enough clear and complete target images for the military target reconnaissance mission.

There are currently primarily two classic ways for SAR/ISAR imaging: the experimental measurement method, which involves measuring solid or scaled targets in an anechoic microwave chamber or outdoors (Wu 1976); The other is computational electromagnetic computer simulation calculation and analysis, in which computer-aided design software is used to precisely represent the geometric shape of the target (Franceschetti et al. 1995). The electromagnetic calculation approach is then achieved by computer code programming. Following that, simulation calculations can be used to determine the target's electromagnetic scattering echo characteristics. The aforementioned approaches, however, have flaws: Experimental measurement calls for a significant investment in time, materials, and money. It is challenging to obtain the SAR/ISAR imaging results by actual experimental measurement, particularly for uncooperative targets. To obtain a sufficiently accurate SAR/ISAR image, computer geometric modeling of the detailed target surface is necessary. However, because the real target surface can change the scattering characteristics of the model, this process is challenging (Chang et al. 2011). Additionally, although computer modeling and simulation can accurately restore target details in the simulation of small size targets, it is challenging to do so for large size targets (Chen et al. 2012).

1.2 Motivation of this review

Computer image processing has been widely used in a variety of SAR/ISAR image generation thanks to the quick growth of computer technology, particularly deep

learning technology. Such as SAR/ISAR target detection, super-resolution reconstruction, automatic target recognition (ATR), image classification, image segmentation, and other fields (Zhang et al. 2022d, e).

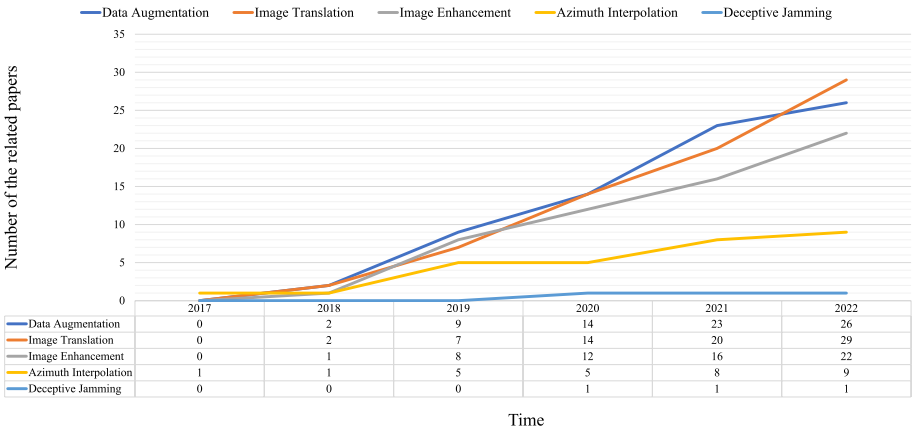
Because creating a probability distribution model using data-driven techniques is the core process of deep learning (LeCun et al. 2015; Zhang et al. 2022a). The quantity and quality of training data therefore affect the model's performance more and more as the deep learning network developing and the network layer continue to deepen (Sung et al. 2018; Zhang et al. 2022c).

However, the amount of satisfactory data is relatively small during the SAR/ISAR imaging process due to some objective reasons, such as the data acquisition of uncooperative targets, limited resolution, and other unfavorable factors, which results in inadequate deep learning network training and poor performance in a variety of tasks (Zhang et al. 2018, 2022b).

The data should be supplemented when the training samples fall short of the deep learning model's training needs. To generate images that are consistent with but not identical to current training examples so that the training dataset has enough samples, conventional data augmentation methods include randomly flipping, clipping, rotating, and adding artificial noise (Yang et al. 2020). However, these conventional techniques and straightforward replication are not fundamentally different, and the trained model is unable to reach the necessary stability and robustness.

The artificial intelligence approach outperforms conventional approaches in the task of data augmentation. A revolutionary method to learn the representation of the input information was developed in 2013 by the invention of the variational auto-encoder (VAE), which may be used for image generation, denoising, and style transformation (Kingma and Welling 2013). Then, in 2015, the generative adversarial network (GAN) was proposed, which was brilliant in the tasks of image generation (Goodfellow et al. 2020). GAN can be trained by optimizing the adversarial loss between a generator and a discriminator. After training, the optimal generator can generate images that the discriminator cannot judge whether they are real or fake. With the continuous development of network structure, GAN has been widely applied in image augmentation, super-resolution, image conversion, image style transformation, and other tasks. At the same time, GAN is also widely applied in SAR/ISAR image augmentation tasks (Wang et al. 2017). The existing methods confirm the application potential of GAN in the field of SAR/ISAR image augmentation. They can also prove that GAN can restore image details well and perform well in image augmentation tasks. The development trend of deep generative models for SAR/ISAR image generation tasks is shown in Fig. 2.

As shown in Fig. 2, the application of deep generative networks in SAR/ISAR image generation tasks has grown steadily since 2017. Before 2021, the number of papers applied for data augmentation is the largest and the fastest growing. However, after 2021, the applications for image translation and image enhancement develop more rapidly, while the growth rate of papers only used for data augmentation is slowing. Compared with the three applications mentioned above, the number of papers applied for azimuth interpolation is relatively slow, only consists of six papers till 2022. The papers used for deceptive jamming are even fewer, which proves that there is still some vacancy in this field, indicating that it has a certain potential for development in the future.



	2017	2018	2019	2020	2021	2022
Data Augmentation	0	2	9	14	23	26
Image Translation	0	2	7	14	20	29
Image Enhancement	0	1	8	12	16	22
Azimuth Interpolation	1	1	5	5	8	9
Deceptive Jamming	0	0	0	1	1	1

Fig. 2 The development trend of different SAR/ISAR image generation tasks based on the deep generative models

1.3 Related reviews

In the research of radar image analysis, SAR/ISAR is one of the most essential topics. Hence, there exist some surveys about the deep learning techniques applied for SAR/ISAR image generation, which are summarized as follows: In Denis et al. (2021), the techniques based on deep learning for SAR image restoration are summarized and analyzed. The authors also indicate that the speckle reduction based on deep learning is one of the most active research topics in the research of SAR image, which performs the optimal by comparing with the classical approaches, such as wavelet transform. Meanwhile, the authors claim that the deep learning methods for speckle reduction can be classified as supervised, self-supervised, and single-image self-supervised methods. Besides, the L2 loss is widely applied for regression, but for image generation, L1 loss is a better choice to reduce the impact of poor modeled samples. For supervised training, the perceptual loss functions can be applied to allocate more weight for the meaningful areas of the image, and the adaptive loss functions are most applied for self-supervised training. For GAN-based models. the adversarial loss is necessary to establish an adversarial system between the generator and discriminator. There are 16 papers in total and 10 papers are related to SAR image restoration.

In the review Lalitha and Latha (2022), the CNN-based remote sensing image augmentation approaches are summarized. Remote sensing images are hard to acquire sometimes, hence it is easily caused the insufficient training in the tasks of ATR and target detection. Firstly, the author proposes a project for image augmentation, namely Augmentor, which can be applied to easily customize the augmentation of data. In addition, the detection accuracy of several improved CycleGAN are compared, and the multi-scale CycleGAN with attention mechanism can achieve the optimal performance. Besides, the authors summarize the works about image augmentation for style transferring, image super-resolution, and the classification of the hyperspectral image. There are 19 papers in total and 11 papers are related.

In Jozdani et al. (2022), the author provides a comprehensive review of the applications of GAN in remote sensing image processing. The summarized remote sensing papers

include LiDAR, SAR, hyperspectral, multispectral, and visible images. And the review classifies the related papers based on different applications, including image classification, image reconstruction, data translation, domain adaptation, and miscellaneous. The authors indicate that the methods based on the conjunction of adversarial loss and pixel-based loss, such as L1 and L2 loss, are most widely applied for image generation. And the combination of the pixel-space loss and perceptual loss is very popular while training a GAN-based model. Moreover, the authors also summarize the most used evaluation metrics in the tasks of remote sensing image generation. However, though the review is excellent enough and consists of 231 papers in total, there are only 35 papers are about the application for SAR images.

In the review Qiongnan et al. (2021), the authors indicate that the applications of GAN in the task of SAR image processing consist of two parts, including data augmentation and super-resolution reconstruction. By using the proper data augmentation strategy, the generalization and robustness of the model can be improved without over-fitting. However, SAR augmentation using the traditional methods only changes the shape of the image at the geometric level, without considering the imaging mechanism of SAR images. The authors summarize the application of DCGAN, WGAN, and Spectral Normalization GAN (SNGAN) in the task of data augmentation, and the application of Super-resolution GAN (SRGAN) for super-resolution reconstruction is summarized. In addition, the loss function of SRGAN is designed based on the difference between features, including the adversarial loss, the total variance loss, and the feature map loss. In addition, several most used evaluation indices are listed. There are 46 papers in total and 16 of them are related.

In the review Qingling and Wei (2020), the authors summarize the applications of SAR in the task of image generation and style transformation. In the topic of SAR image generation, the application of DCGAN, cGAN, and InfoGAN are summarized. And the application of cGAN in the task of style transformation is summarized. There are 29 papers in total and 15 papers are related.

Briefly, there exists many researchers have published comprehensive and detailed review papers on SAR image generation based on deep learning technology, proving that in the field of SAR image analysis, image processing using deep generative networks is one of the primary research branches. However, despite several papers being excellent enough (Jozdani et al. 2022), these papers did not specifically focus on the application of deep generative networks in the field of SAR image. In addition, although GAN is an important branch of the generation network, there exist some other networks, such as VAE-based and probabilistic graphical-based generation networks, which have been applied in SAR image processing tasks. Therefore, it is necessary to supplement this vacant research field. For a clearer illustration, a histogram (Fig. 3) is applied to represent the contribution of each related paper to this research field. This paper mainly summarizes the application of deep generative networks in SAR (including SAR and ISAR) image generation tasks, and the development and potential of these approaches are summarized. This review can provide a considerable reference value for radar image and computer vision researchers. In this paper, there are 87 papers are about the application of deep generative networks for SAR image processing.

1.4 SAR/ISAR image generation workflow

To illustrate the general steps in the tasks of SAR/ISAR image generation based on the deep generative networks, the organization chart is shown in Fig. 4, which consists of the

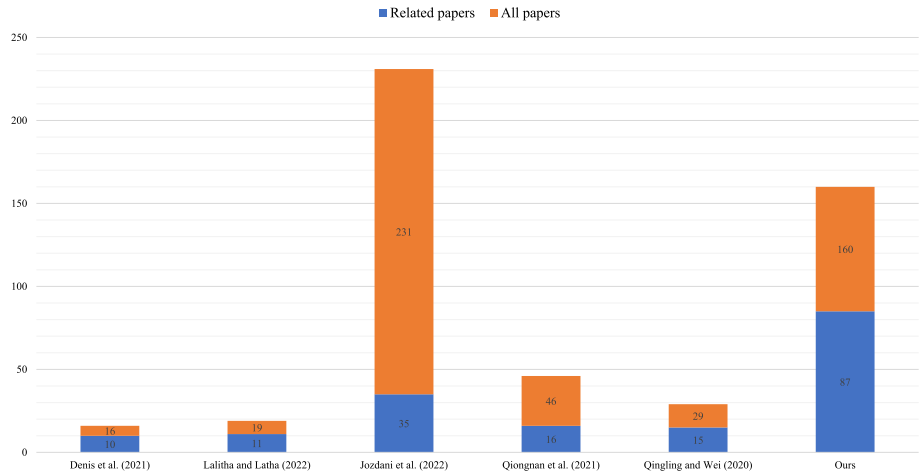


Fig. 3 The comparison of the number of papers summarized in related review papers. The numbers in each column represent the number of the related papers and all papers, respectively

categories of the application domain, SAR/ISAR images, and processing methods, including image pre-processing, target task, deep generative networks, and evaluation indices.

As shown in Fig. 4, the SAR/ISAR image can be acquired for the tasks of the military, agriculture (Sharifi et al. 2022; Sharifi 2021; Farmanov et al. 2023), city planning, aerial survey, aerial remote sensing, satellite ocean observation (Mohammadi et al. 2021), space reconnaissance, and so on. Then, according to the different imaging mechanisms, SAR images are composed of SAR and ISAR images.

The SAR is a coherent system, and speckle noise is the inherent characteristic. To suppress the noise and speckles generated while imaging, and enhance the contrast

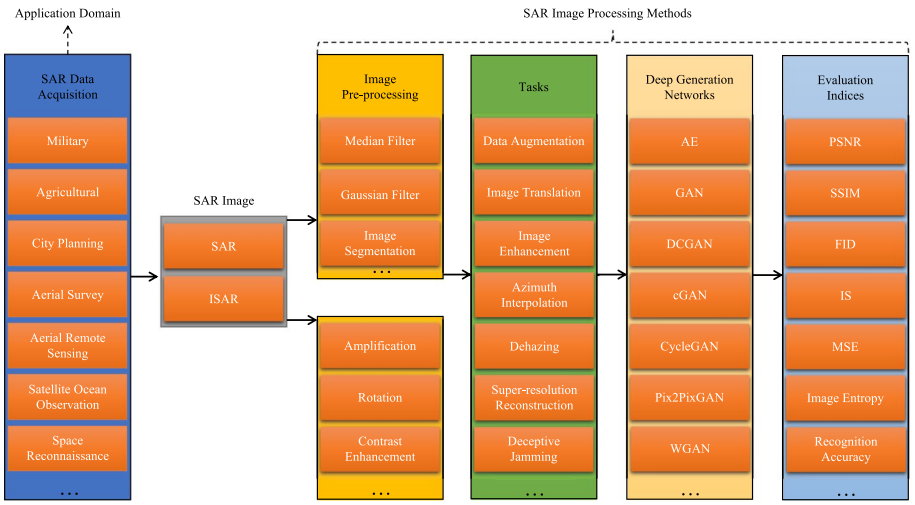


Fig. 4 The organization chart of SAR/ISAR image generation based on the deep generative networks

between the target and the background, image pre-processing technologies can be applied to enhance the quality of the SAR images, such as the application of wire filters for denoising, and image segmentation can be applied for preparation of the further target generation (Singh et al. 2021). The common filtering methods consist of low-pass filtering and structural filtering (Ma et al. 2022). Low pass filtering can effectively suppress speckle noise and blur the image, including mean filtering, Gaussian filtering, etc. The structural filtering can be designed for convolution operation, which can not only suppress speckle noise but also keep the structure information of the image without loss. Besides, ISAR is usually applied for spatial target imaging, hence the background of ISAR image is relatively clearer, and speckle noise is relatively fewer than in SAR image (Vehmas and Neuberger 2021). So, denoising filtering methods are rarely applied for ISAR image pre-processing.

After that, the deep generative networks can be applied for different SAR/ISAR image generation tasks. Due to the difficulties of capturing the SAR/ISAR images, deep generative networks are always applied for data augmentation to solve the problem of the limited dataset (Ding et al. 2016). Besides, in several cases when the optical imaging system cannot capture satisfactory images of the target, such as the cloudy day and thunderstorms, the SAR/ISAR imaging system can complement the damaged images due to its capability of all-weather and all-time imaging. Hence, the translation between the SAR/ISAR and optical images is one of the most significant application fields in the tasks of SAR/ISAR image generation (Fuentes Reyes et al. 2019). In addition, SAR/ISAR images with low resolution cannot meet the requirement of accurate recognition and detection, or if there exist hazes in the images, super-resolution reconstruction and dehazing technologies can be applied for image enhancement (Shen et al. 2020). Moreover, it is difficult to capture the SAR/ISAR images of a target with all azimuths, so the deep generative networks can be applied to generate the SAR/ISAR image with a specific azimuth to solve the problem of insufficient training in some cases (Zhang et al. 2018). In addition, each task proposed above is part of image generation. For example, image translation can be regarded as one kind of image generation, which generates the color features but not the texture features. Analogously, image enhancement can generate detailed texture or color features, and azimuth interpolation is to generate the image of one single target with another azimuth in the same image domain. Hence, in this review, the description of 'Image generation' refers to all tasks summarized above.

Moreover, the choice of deep generative networks is the most significant part to complete different tasks. Auto Encoder (AE) is an unsupervised learning generative model based on the backpropagation algorithm and optimization methods (such as gradient descent method), which can use the input data itself as the supervision to guide the neural network to try to learn a mapping relationship and obtain a reconstructed output (Bank et al. 2020). The AE model consists of two main parts: encoder and decoder. The encoder can encode a high-dimensional input into a low-dimensional hidden variable, thus forcing the neural network to learn the most informative features. The decoder can restore the latent variables of the latent layer to the initial dimension, trying to perfectly or approximately restore the original input. The AE-based models have a strong generalization and can generate new data without data annotation (Chen et al. 2019).

Generative adversarial network (GAN) is one of the most significant research projects in the field of artificial intelligence (AI). GAN is developed based on the deep neural network architecture, which consists of a generator and a discriminator (Goodfellow et al. 2020). The generator can generate fake data and try to cheat the discriminator. The discriminator can be trained to distinguish the real and generated fake images. In the process of training

iteration, the generator and discriminator continue to evolve until they reach the equilibrium state (Nash equilibrium). The objective function of GAN is as follows:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}\left[\log D(x)\right] + \mathbb{E}_{z\sim p_z(z)}\left[\log(1-D(G(z)))\right] \tag{1}$$

where x represents the real image, z represents the input noise, $G(z)$ represents the generated image and $D(*)$ represents the discrimination result of the discriminator. The use of log can enlarge the loss and accelerate the convergence while training a GAN. However, the vanilla GAN uses full-connected layers to construct the neural network, which cannot obtain satisfactory generated images. Deep convolutional generative adversarial network (DCGAN) induces the convolution layers to GAN to improve the generation performance and accelerate the convergence (Radford et al. 2015). In DCGAN, the transposed convolution is applied for up-sampling and the convolution with a stride is applied to replace the pooling for down-sampling.

However, the generated images of GAN are random and unpredictable. Conditional generative adversarial network (cGAN) is proposed to introduce the conditional information to the generator and discriminator. In cGAN, the extra conditional information is added in both generator and discriminator, and the generated image can be discriminated as real only if it is true enough and meets the conditions (Mirza and Osindero 2014). The objective function of cGAN is shown as follows:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}\left[\log D(x|y)\right] + \mathbb{E}_{z\sim p_z(z)}\left[\log(1-D(G(z|y)))\right] \tag{2}$$

which is similar to Eq. 1, and the conditional information y is added in the network. However, in cGAN, conditional information is known and must be supplied before training. InfoGAN is proposed to obtain the decomposable feature representation through unsupervised learning (Chen et al. 2016). A hidden code is added to the generator, and mutual information is introduced to represent the correlation degree between the input noise and the hidden code. The value of the mutual information is zero when the two variables are independent. In addition, an extra network, Q-network can be trained to calculate the mutual information, which uses the same convolution layers of the discriminator and a full-connected layer is added to output the predicted conditional information.

Auxiliary classifier generative adversarial network (ACGAN) is another implementation of cGAN. The label information in ACGAN is not only applied for training, but also for reconstruction of label information (Odena et al. 2017). The difference between the cGAN and ACGAN is that the discriminator in cGAN can receive the generated images and the category information at the same time, while the discriminator in ACGAN only receives images, including the generated and real images. Therefore, the cross entropy of the real and predicted label of the generated image is added in the loss function of the discriminator and generator, expecting to get the correct classification results.

The deep generative networks proposed above can generate the image using a set of random noise. Pix2PixGAN is designed for the tasks of image translation or coloring, which is essentially an improved cGAN, and an image can be applied as the conditional information of the network. For the tasks like image translation, much information, such as the contour and texture features, can be shared between the input and output images. Hence, the U-Net is applied to construct the encoder of the Pix2PixGAN, which can share the features in different layers using the concatenation operation (Isola et al. 2017). Moreover, the patch discriminator is designed to accelerate the training process. And L1 loss is added in the generator loss function to improve the similarity between the real and generated images.

Besides, the paired images are required for training, and discrimination to guarantee the mapping relationship.

However, it's always difficult to capture the paired images, which are necessary while training a Pix2PixGAN, and CycleGAN is designed to realize the domain transformation without the paired images (Zhu et al. 2017). The main idea of CycleGAN is to learn the feature mapping between two different domains. There are two generators and two discriminators in CycleGAN. Furthermore, the cycle consistency loss is added in CycleGAN to guarantee the generated image retains the features of the original image. The architecture of the models summarized above is shown in Fig. 5.

Besides, the super-resolution generative adversarial network (SRGAN) is designed for super-resolution reconstruction of the low-resolution images (Wang et al. 2021). Traditional methods can restore high-resolution images by adding some prior information, such as interpolation, sparse learning, and random forests based on regression methods, and CNN-based models perform the best in the task of super-resolution reconstruction. The generator in SRGAN consists of a set of residual blocks, which can help to restore more detailed information, and VGG19 is applied to construct the discriminator. Moreover, the content loss is added as a part of perceptual loss, which is defined as the mean square error (MSE) between the generated high-resolution image and the feature map of the real image obtained through the first 16 layers of the VGG network.

It is well-known that the applicability of generative models to image generation tasks such as SAR image generation inevitably suffers from the so-called abnormal samples problem. Researchers can train a single classifier further to determine whether the input samples are outliers or not. However, it is very difficult to train a classifier to detect outliers facing unknown categories. AnoGAN (Schlegl et al. 2017) uses the generator to learn the distribution of the conventional data. The test sample can be compared with

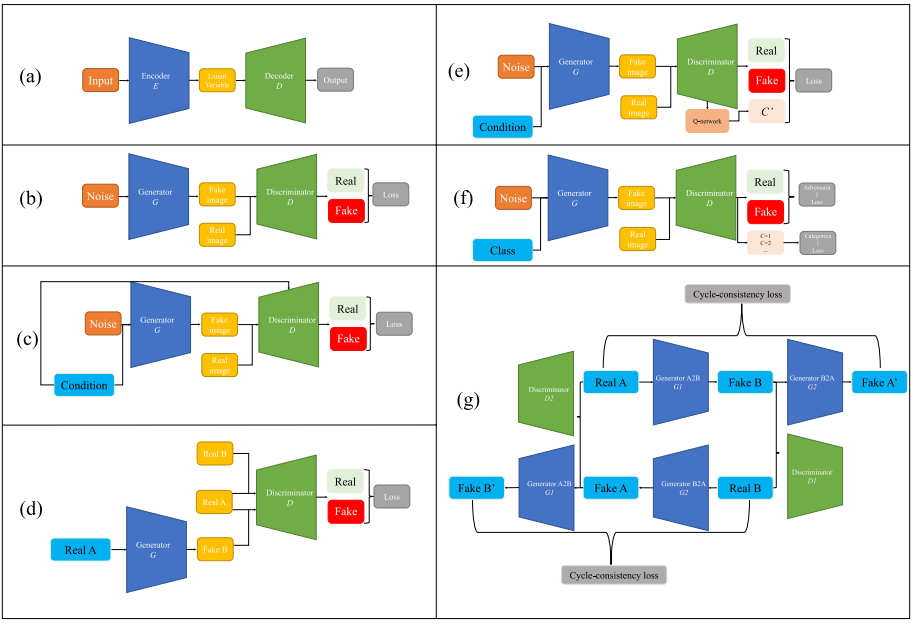


Fig. 5 The architecture of the summarized models. **a** AE; **b** GAN; **c** cGAN; **d** Pix2PixGAN; **e** InfoGAN; **f** ACGAN; **g** CycleGAN

the image generated by the trained generator and determined whether the sample is an outlier. DCGAN is applied as the backbone of AnoGAN, whose innovation lies in taking the initial noise z as an updatable parameter to calculate the $L1$ loss of the generated image and the real image. At the same time, the feature map difference of the middle layer of the trained discriminator is calculated to update the noise z and the abnormal area can be found. However, AnoGAN needs to update parameters during testing, causing the slow training speed. In 2018, a novel detection method based on BiGAN was proposed (Zenati et al. 2018), which can learn an encoder E that maps input samples x to a late representation z to avoid expensive calculation steps in the test process.

The traditional GANs use Jensen–Shannon (JS) divergence to measure the distance between two distributions, however, there exist two main limitations. The first one is the problem of pattern collapse, that is, the generator can only generate similar samples, which can only cover a single pattern in the data distribution (Arjovsky and Bottou 2017). The second one is that there is no interpretable index to reflect the convergence state of the network. Wasserstein generative adversarial network (WGAN) is proposed to address the problems above (Arjovsky et al. 2017). Wasserstein distance is used to measure the distance between two distributions, which has superior smoothing characteristics and can solve the problem of gradient disappearance compared with the traditional Kullback–Leibler (KL) divergence and JS divergence. The application of Wasserstein distance can continuously optimize the network even if the two distributions have no overlaps.

After image generation, the application of evaluation indices can quantitatively evaluate the performance of the deep generative networks. Generally, the most applied evaluation approach is to calculate the recognition accuracy (Bernardi et al. 2016). The generated images can be added as a part of the test dataset in different proportions to judge whether the trained model can distinguish the real and generated images. An alternative approach is applying the generated images as the training dataset and the real images as the test dataset, showing whether the generated images can be applied for data augmentation. The recognition accuracy is easy to use and can evaluate the generation performance, however, the value is determined by the recognition networks, which cannot be compared using different networks.

Similarly, inception score (IS) is proposed to evaluate the generation performance (Barratt and Sharma 2018). The Inception Net-V3 trained on ImageNet is applied as the classifier to output the probability that the input sample belongs to a certain category. IS considers certainty and diversity at the same time, that is, a single generated sample should be classified as a certain category, and all generated samples should distribute in a uniform distribution. The definition of IS is shown as follows:

$$IS(G) = \exp\left(E_{x\sim p_g} D_{KL}(p(y \mid x) \| p(y))\right) \tag{3}$$

where $p(y \mid x)$ means the classification for a single sample and $p(y)$ means the edge distribution of all samples in all categories. However, when the network is over-fitted, the generator can only generate samples of the training dataset, but its generalization ability is weak, and IS cannot find this problem. In addition, if the category of the generated image does not belong to the training sample categories of the pre-trained model, the IS will be relatively low.

Fréchet inception distance (FID) is another commonly applied evaluation indices for image generation (Soloveitchik et al. 2021). Firstly, the feature maps of generated samples are extracted using a CNN classifier. Then, the mean value and variance can be calculated

and compared with that of training samples. Finally, the Fréchet distance of the two distributions can be calculated, which is defined as follows:

$$FID = ||\mu_{data} - \mu_g|| + tr\left(\Sigma_{data} + \Sigma_g - 2\left(\Sigma_{data}\Sigma_g\right)^{\frac{1}{2}}\right) \tag{4}$$

Smaller FID means a closer distance between the two distributions and better generation performance. Besides, FID is robust to noise and can better evaluate the quality of the generated samples, which is consistent with human visual results.

In addition, there exist some evaluation indices for directly quantifying the quality of the generated images, but not relying on an extra classifier. Mean square error (MSE) can be applied to evaluate the difference between the real and generated images in the pixel level, which is defined as follows:

$$MSE = \frac{1}{M \cdot N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[I(i,j) - K(i,j)]^2 \tag{5}$$

where M and N represent the total pixel number of two images, respectively. The smaller MSE means the better generation performance (Chai and Draxler 2014).

Based on MSE, the peak signal to noise ratio (PSNR) is proposed, which is defined as follows:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{6}$$

where MAX means the possible maximum value in the image. PSNR is usually applied to measure the similarity between the real high-resolution and generated super-resolution images (Korhonen and You 2012). However, the PSNR cannot be completely consistent with the intuitive feelings of the human eye, because the human eyes have no absolute visual sensitivity to the difference between two images.

The structure similarity index measure (SSIM) is designed to measure the similarity between two images, which is a perceptual model and more in line with the intuitive feelings of the human eye (Hassan and Bhagvati 2012). SSIM considers the luminance, contrast, and structure of the images, which is defined as follows:

$$SSIM(x,y) = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_{xy} + c_2\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{7}$$

where μ is the mean value of the image, σ^2 is the variance of the image, and c is the constant to improve the stability. SSIM is always applied to measure the similarity between the real and generated images, and higher SSIM means better generation performance.

Besides, the information entropy (IE) can be applied in this field to measure the information in the image, which is defined as follows:

$$H(A) = -\sum_a P_A(a)\log p_A(a) \tag{8}$$

where a means the grey-scale value and P represents the probability distribution (Núñez et al. 1996). The higher IE means the image has more abundant information and better quality.

Furthermore, there are more evaluation indices for specific tasks, such as precision, recall, and F1-score for image classification, intersection over union (IoU) and pixel accuracy for image segmentation, and so on. The analysis of these evaluation indices is summarized in Sect. 7.

1.5 Structure of this review

The flow chart of paper selection in this review is shown in Fig. 6. The initial papers are searched from Google Scholar, IEEE, ACM, Nature, Science, Elsevier, Wiley, Hindawi, IOP, PloS, BMC, and Springer, and the keywords contain "SAR/ISAR image generation", "Deep generative network", "GAN" and "Generative Adversarial Network". Then the duplicate and irrelevant papers are deleted. 21 papers do not conform to this review. Finally, 92 papers are selected about the deep generative network for SAR image processing, containing 5 review papers, 13 papers for ISAR images, and 74 papers for SAR images.

The review is structured as follows: From Sects. 2 to 6, the applications of deep generative networks for different tasks are summarized, including the task of SAR/ISAR data augmentation, image translation, image enhancement, azimuth interpolation, and deceptive jamming. Each part is classified and organized based on deep generative networks. Then, in Sect. 7, the works summarized in this review are analyzed, including their application fields, deep generative models, and evaluation indices. Furthermore, the potential methods for SAR/ISAR image generation are analyzed. Finally, the conclusion and future work of this review is analyzed and listed in Sect. 8.

Briefly, due to the difficulties while capturing SAR/ISAR images, the deep learning models cannot be trained sufficiently, and cannot perform well in different tasks. Besides, the classic methods (such as flipping and cropping) cannot meet the requirement of efficient data augmentation. To this end, the purpose of this survey is to summarize the application of deep generative networks for SAR/ISAR image generation, and the applications of deep generative models, datasets, and evaluation indices are systematically analyzed. In addition, few existing works concentrate on deep generative networks for SAR/ISAR image generation. For example, in the review (Jozdani et al. 2022), the authors present a comprehensive survey of GAN-based models for radar image generation, but the content of SAR/ISAR images is relatively less. Moreover, this paper is also a review of the evolutionary

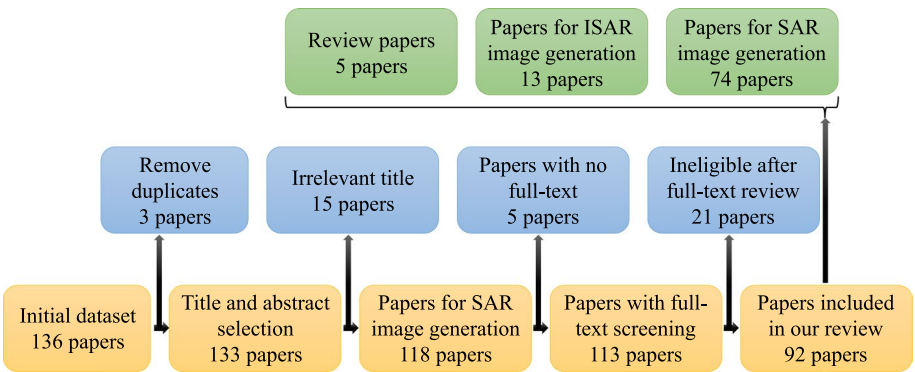


Fig. 6 The systematic flow chart of paper selection for our work

history of the deep generative network. With the development of deep learning technologies, the images generated by deep generative models are much more excellent and satisfactory. Hence this review can be a sort of reference for radar and artificial intelligence workers.

2 SAR/ISAR image generation for data augmentation

Deep learning is to build a probability distribution model based on data-driven, hence, the performance of a deep model is heavily dependent on data quality and data quantity. However, there exist many objective and minus factors in the process of SAR imaging, which leads to few captured data meeting the requirements, thus resulting in poor performance in various tasks. Because the imaging principles of SAR images and optical images are different. For SAR images, the scattering characteristics of targets captured by different wavebands are also different, and SAR images contain more information, such as scattering characteristics and azimuth of targets, so the traditional approaches commonly applied for optical image data augmentation, such as inversion, translation, and cropping, are no longer suitable for SAR image data augmentation. Compared with the traditional methods, deep generative models can generate SAR images with better quality with abundant categories and can maintain the modal information. Therefore, the generation models based on deep learning such as variational auto-encoder (VAE), restricted Boltzmann machine (RBM), and generative adversarial network (GAN) are widely applied in the tasks of image generation. Especially, in the tasks of SAR image generation, GAN-based models are the most widely applied, the fastest growing approaches, relying on their ability to learn abundant image information from a large number of unlabeled data in an unsupervised way. This chapter summarizes the applications of SAR image generation for data augmentation.

2.1 DCGAN-based methods

The existing approaches for SAR image recognition are mostly based on supervised learning, which strongly dependent on image labels. However, the labeling process of SAR images is time and manpower-consuming, which cannot meet the requirement of a sufficient dataset for deep neural network training. In Bhamidipati et al. (2020), a traditional DCGAN is trained using only 100 images of MSTAR. Then, DenseNet is applied to measure the mean accuracy of generation images. DenseNet can be applied to accelerate convergence and improve the generalization ability. The accuracy score achieves 95.93% and Fischer's inception distance (FID) score achieves 0.00143, showing the satisfactory generation performance of the proposed method.

Moreover, the architecture of DCGAN can be modified to improve the generation performance with a specific purpose. In Gao et al. (2018), an improved DCGAN with an extra discriminator is designed to augment the SAR image dataset, whose loss can be calculated using the average value between the two discriminators. Softmax is applied in the last layer of the discriminator for multi-categories recognition. Firstly, two discriminators are trained using the labeled image, and the trained discriminator can predict the unlabeled image. Then, the newly labeled images with high confidence will be added to the labeled dataset for the next round. For the generator, cross-entropy is applied as a new loss function to represent the difference between the generated image and the real image. The proposed method achieves the accuracy of 98.14%, 97.97%, 97.22%, and 95.72% by using the rate

of 0.2, 0.4, 0.6, and 0.8 unlabeled data, respectively. In Yan et al. (2022), the authors consider applying the multi-scale GAN for SAR image aircraft target data augmentation. A single image can be applied to train the generator and avoid the tedious process of making a large number of data pairs and can obtain a large number of high-quality generated samples in a short time. There are three different sizes of the convolutional kernel are applied to extract the multi-scale features of the images, including the size of 3×3, 5×5, and 7×7. The idea of multi-scale for remote sensing image target detection is first carried out by Khoshboresh-Masouleh et al. (2020), which provides the basis and hypothesis for further research. After that, the Markov discriminator is applied to distinguish the real SAR images and the generated images. For the design of the loss function, BCE loss is applied as the adversarial loss. Finally, the proposed improved SinGAN can achieve the ENL of 2.0093, SSIM of 0.9081, FSIM of 0.6676, and CSIM of 0.2661. In the task of target detection, Faster R-CNN can achieve an accuracy of 74.2%. In Hwang and Shin (2021), a traditional DCGAN with an extra classifier is designed, namely TripleGAN. Firstly, the classifier can generate a fake label for a real image. Then, the generator can generate a fake image based on a real label. After that, the discriminator is applied to judge whether the image and label pairs are real or fake. In addition, cross-entropy loss and weight normalization are applied for training and optimization. The dataset of MSTAR is applied in this experiment. Finally, VGG16 is applied to evaluate the generation performance of the proposed model, whose accuracy achieves 0.9141, showing the optimal generation capability. In Li et al. (2019b), DCGAN and Online Hard Examples Mining (OHEM) are applied for SAR image generation in the task of ship detection. To improve the adaptability to spatial transformation, some existing approaches will do some data expansion preprocessing such as segmentation, amplification, and rotation. The spatial transformer network (STN) allows the operation of images in the network. The network can generate a feature map with spatial transformation characteristics so that the trained model has strong spatial transformation invariance without collecting many samples. The architecture of the proposed model is shown in Fig. 7. The feature maps are set as input, and the localization network will estimate transformation parameters (such as rotation angle, transformation distance, and scale factor). These parameters will be used as the input of the generator and sampler, and the output is the feature map after deformation. After image generation, the OHEM algorithm is applied to replace the SGD and optimize the Fast RCNN for target detection. An existing

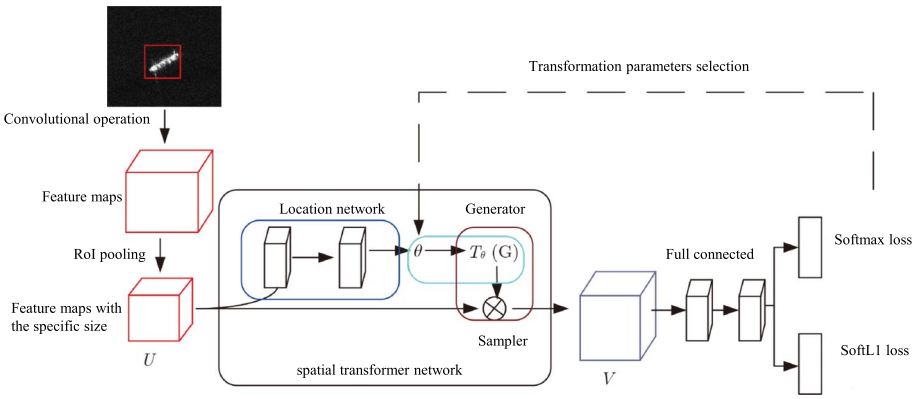


Fig. 7 The architecture of the proposed method [in Li et al. (2019b, Fig. 2)]

SAR Ship Detection Dataset (SSDD) is applied for training and validation. The proposed combined approach can achieve an improvement of 2.1% in target detection.

Besides, the abundant features of SAR images can be extracted to improve the generation performance. In Zhou (2020), the texture information of the SAR image is introduced into DCGAN, and the prior information of SAR image distribution is introduced by using the Monte Carlo algorithm so that the generated SAR image can be closer to the distribution of the real image. In addition, the discriminator of PatchGAN and structural similarity loss are used to capture the details and structural similarity of SAR images, and the overall semantic information of images can be obtained through the pyramid network structure to improve the similarity with the original images. The training dataset includes MSATR and SAR-I data. In Baier et al. (2021), the semantic features and raster image generate the RGB and SAR images using an improved DCGAN. Firstly, spatially adaptive normalization is applied to construct a generator to fuse the semantic features and depth information. Moreover, the ResNet blocks and concatenation operations are applied to erase the block of checkerboard. The architecture of the proposed network is shown in Fig. 8 The datasets of GeoNRW and IEEE GRSS Data Fusion Contest 2020 (DFC2020) are applied for validation. Finally, U-Net is applied to evaluate the performance of the generated image in the task of image segmentation, and the image generated by the proposed model can achieve the mIoU of 0.2326, Fid of 0.0078, and pixel accuracy of 0.7692.

2.2 cGAN-based methods

By comparing with the traditional DCGAN, the cGAN-based model can generate SAR images with specific categories by introducing the category information. In Qin et al. (2022), an improved cGAN, namely conditional Wasserstein deep convolutional generative adversarial network with gradient penalty (CWDCGAN) is proposed for image expansion, which can generate SAR images of specific categories. The innovation point is, conditional information is introduced into the network to control the image categories generated, and Wasserstein distance and gradient penalty are used for optimization. Moreover, SVM and Resnet-18 are combined to enhance the generalization ability and recognition performance of the ATR system. Finally, the recognition accuracy of 96.81% is obtained. The architecture of the generator is shown in Fig. 9.

There are two parallel pipelines for noise upsampling and condition information upsampling. The noise matrix is concatenated with the feature map sampled on the condition information matrix after upsampling. Different from the traditional cGAN, that is, the

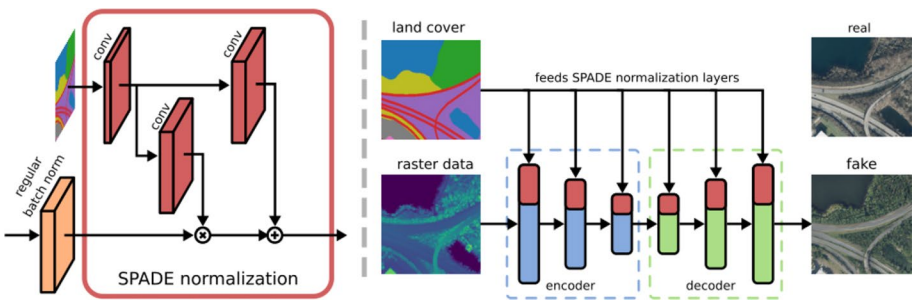


Fig. 8 The architecture of the proposed model [in Baier et al. (2021, Fig. 1)]

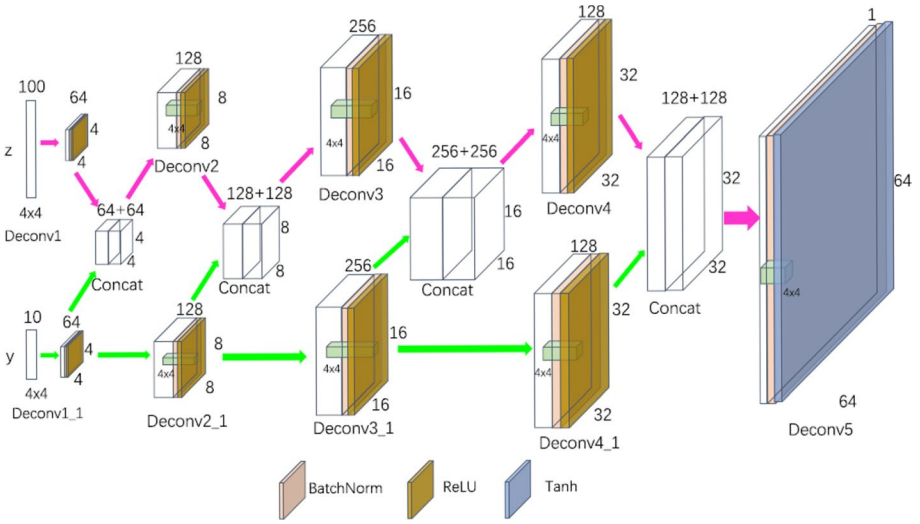


Fig. 9 The architecture of generator in CWDCGAN [inQin et al. (2022, Fig. 4)]

conditional information is concatenated at every step in the training process, instead of being concatenated at the first step. This modification can ensure that category information can be applied at every layer in the process of training, rather than being forgotten. After image generation, a common convolutional network is applied as a discriminator to output the judgment of the results. The gradient penalty is added to the loss function to guarantee stability while training. At the same time. root mean square promotion (RMSProp) is applied as an optimizer. The MSTAR dataset is applied to validate the generation performance of the proposed model, and seven indices are applied for quality evaluation. The mean value achieves 0.1096, variance achieves 0.0238, information entropy achieves 0.9022, linear index of fuzziness (LIF) achieves 0.0244, mean gradient achieves 0.3932, the gray level difference (GLD) achieves 0.1137, and FID achieves 30.4716, which perform the best by comparing with other models. After that, in the process of ATR, ResNet is applied for feature extraction, and SVM is applied for classification. The dataset with 80% original images and 20% generation images achieves 97.98% accuracy. In Liu et al. (2021a), an improved cGAN is applied for semi-supervised SAR data augmentation in the tasks of ATR, called semi-supervised conditional generative adversarial network with a multi-discriminator (SCGAN-MD) Firstly, the random noise and the label are fed into the cGAN-based generator. Then, the generated images are fused with real images to train a multi-discriminator. After that, the trained multi-discriminator is applied to label the unlabeled SAR images. Finally, the newly labeled images are applied for CNN training in the task of ATR. The MSTAR dataset is applied in this experiment. The result shows that the proposed SCGAN-MD can achieve the recognition accuracy of 99.13%, 98.43%, 96.12%, and 91.18% when using all, 80., 40, and 20 real labeled SAR images in the dataset, respectively.

Different from the traditional cGAN, the label information is not included in the input of the discriminator of ACGAN but is considered as the optimization target for discriminator training. In Kong and Zhang (2021), the original SAR images with their corresponding labels are applied for ACGAN training. Then, the generated SAR images are added to the

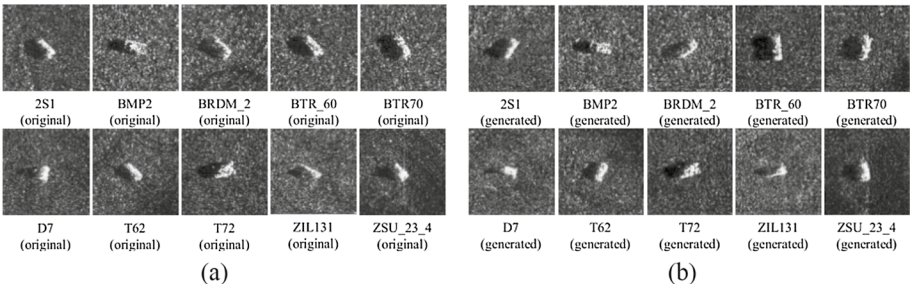


Fig. 10 The generation results of the proposed approach. **a** Original images; **b** generated images[in Kong and Zhang (2021, Fig. 5)]

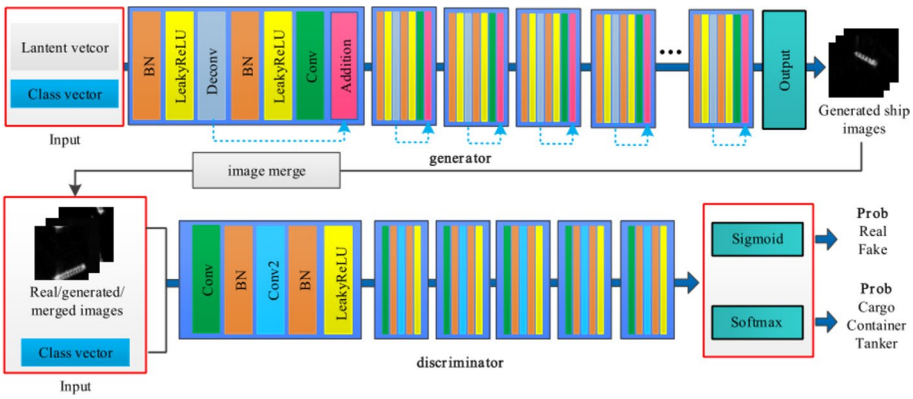


Fig. 11 The architecture of the proposed model [in Li et al. (2020a, Fig. 2)]

original SAR images as a new dataset. The generated results are shown in Fig. 10. After that, the ResNet34 is applied as a classifier for SAR image recognition. The result shows that the proposed approach can achieve an accuracy of 99.38% in the task of target recognition, which is 2.35% higher than using the original dataset without data augmentation.

In Li et al. (2020a), an ACGAN is proposed to improve the generation performance of the network and the classification accuracy for generated images. The proposed generator is modified based on the existing auxiliary classifier GAN (ACGAN), including two improvements. The first one is the transposed convolutional operation is applied for up-sampling instead of using linear interpolation, and the size of output images can be enlarged. The second one is, an extra layer is added at the end of the generator to generate larger images. In addition, the residual convolutional operations are applied to erase the texture information of the different types of images. Furthermore, weighted loss of WGAN-GP and Wasserstein is applied in the discriminator, and LeakyReLU is applied as the activation. And a softmax is added to discriminate the category of input images. The architecture of the proposed network is shown in Fig. 11. The GF-3 dataset is applied in this experiment. After image generation, a pre-trained VGG16 is fine-tuned and applied for classification and Multi-scale SSIM (MS-SSIM) is applied for evaluation. Finally, the proposed model achieves the MS-SSIM of 0.935 for Cargo, 0.884 for Container, and 0.908 for Tanker in intersource, and 0.431 for Cargo, 0.312 for Container, and 0.178 for Tanker in intraclass.

2.3 InfoGAN-based methods

In Xie et al. (2021), an improved InfoGAN is applied for SAR data augmentation, and samples are generated purposefully to solve the problem of insufficient sample diversity. The input of the generator consists of two parts, the first part is a set of random noise, and the other is the condition information in latent space. Information theory is applied to express mutual information between latent variables and generators. However, mutual information is difficult to maximize directly, so it is necessary to define a CNN (Q network) to fit it to obtain the lower bound, making the network can optimize continuously. The Q network shares parameters with the discriminator and outputs the judgment category. The MSTAR dataset is applied for training and testing. Finally, to evaluate the generation performance of the proposed model, the recognition accuracy after adding generated images is calculated to verify the effeteness of the augmented data, which achieves 99.02%. The generation result is shown in Fig. 12.

2.4 WGAN-based methods

To solve the problem of the vanishing and collapse of gradient, WGAN and WGAN-GP are designed to improve the stability of training. In Bao et al. (2019), an end-to-end GAN-based model is proposed for SAR image generation. There are three networks are applied, including DCGAN, weight clipping WGAN, and WGAN-GP. For DCGAN, the fractional-strided convolutional and stride convolutional layers are applied to replace the traditional pooling operation, hence the network can learn the sampling approach itself. Then, the fully connected layers are removed. For WGAN-GP, a network with four residual blocks is designed to avoid the problem of mode collapse. The MSTAR dataset is applied for training and validation. The result shows that the WGAN-GP with four residual blocks has the optimal performance for SAR image generation, and the weight clipping WGAN is easily achieving local minimum and hard to converge. In Lu et al. (2019b, 2019a), WGAN is applied for SAR image data augmentation by using the MSTAR dataset, which is more stable for convergence. In Lu et al. (2019b), YOLOv3 is applied to verify the availability of

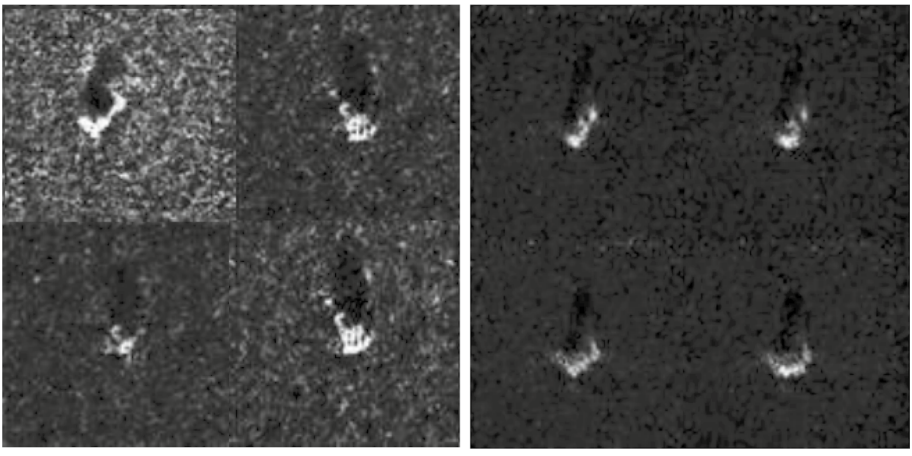


Fig. 12 The generation result by controlling the continuous component in the latent variable [in Xie et al. (2021, Fig. 10)]

generated samples. The generated samples are added in different proportions as the training set. In the target detection results, WGAN achieves a confidence level of 0.808, while DCGAN only achieves 0.6625. At the same time, the recognition accuracy using the dataset including 500 generated images and 1000 generated images is improved by 2% and 3.4% respectively. In addition, in Lu et al. (2019a), the Lipschitz constraint is enforced by using gradient penalty to avoid mode collapse. The generation result shows that the images generated by WGAN have no obvious consistency features and can generate more types of SAR samples, which is much better than DCGAN. In Cui et al. (2019), the WGAN-GP is introduced to generate SAR images for ATR. Firstly, the original WGAN-GP is applied for SAR image generation. Then, a support vector machine (SVM) is pre-trained to select the generated images with high confidence (> 90%), and the rest of them are discarded. Then, an azimuth discriminator is designed including target segmentation, edge detection, and azimuth discrimination. The target segmentation is realized using iterative thresholding, and the edge can be extracted based on the critical value. Finally, the maximum bounding rectangles of the targets are extracted and then the azimuth can be calculated, and the generated SAR images with specific azimuth can be selected. The generated images based on the proposed method achieve a mean value of 182.82, a variance of 405.96, ENL of 87.82, and RR of 0.444. The generation result is shown in Fig. 13.

In He et al. (2019), the quadratic convolutional operation is applied to extract the features of multiplicative speckles in the SAR images, whose architecture is shown in Fig. 14. Then, two GANs with different types are connected parallelly, including PWGAN and CNN-PGAN. The PWGAN consists of several generator and discriminator pairs, which are the same as the number of categories, and the trained model can generate the SAR image with a specific category. The PWGAN is designed based on DCGAN and WGAN-GP, and the loss function can be defined as the loss of WGAN-GP and the mean statistical loss for each category. Different from PWGAN, in CNN-PGAN, WGAN-GP is not combined with DCGAN, and a shared discriminator is designed to determine the category of the SAR image, which is constructed using a pre-trained CNN model. Finally, the proposed PWGAN and CNN-PGAN can achieve the classification accuracy of 80.36% and 78.13% in Dataset1 and 76.88% and 80.0% in Dataset2.

However, the original WGAN-GP is developed to improve the stability and robustness while training a model, but the model cannot generate multi-category images and the training process is uncontrollable. In Du et al. (2021c, d), an improved model

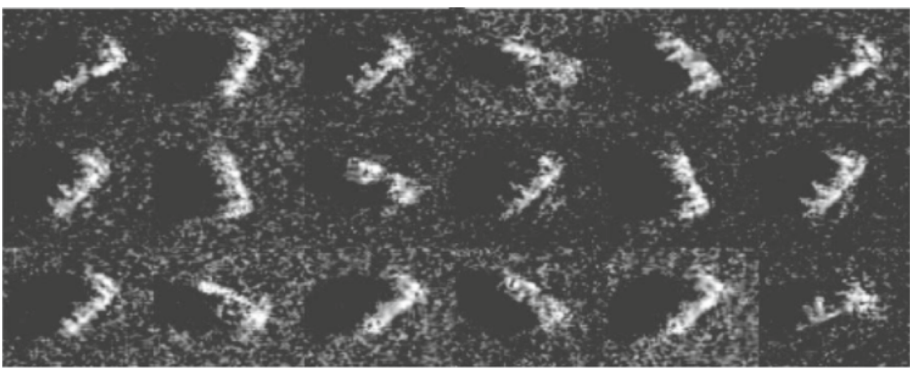


Fig. 13 The generation results of the proposed approach[in Cui et al. (2019, Fig. 9)]

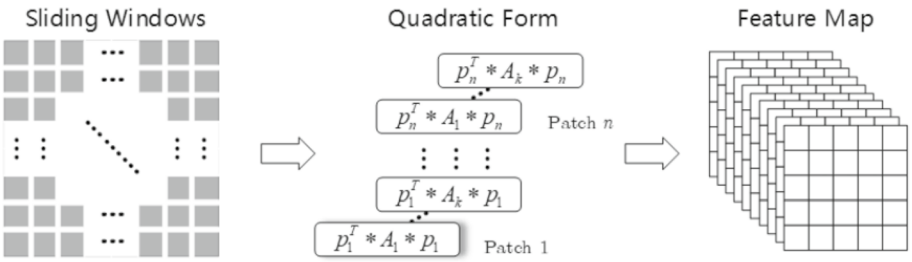


Fig. 14 The architecture of the quadratic convolutional operation [in He et al. (2019, Fig. 3)]

based on WGAN-GP is proposed for multi-category SAR image generation, which consists of four parts: an encoder, a generator, a discriminator, and a pre-trained classifier (ResNet18). Firstly, the feature representations of the received real SAR image are extracted. Then, the extracted features are concatenated with a sequence of Gaussian noise and the category features of input images. After that, the mixed feature is processed by the encoder and the generator can output the final generation result. Finally, the generated image is received by the discriminator and pre-trained ResNet18, and GAN can be trained by minimizing the loss of the classifier. As for the loss functions of the proposed method, the loss function of the generator is the same as the original WGAN-GP, but the loss function of the discriminator is modified as the cross entropy between the discrimination result and the classification result calculated by the pre-trained classifier. The MSTAR dataset is applied for training and validation. The comparison between real images and the images generated by different models is shown in Fig. 15. To evaluate the generation performance of the proposed model, cross and hybrid tests are applied for evaluation. In the cross-test, the generated images are applied for training, and the real images are applied for the test, however, in the hybrid test, the generated images are hybrid with the real images to improve the classification performance. The accuracy of the cross-test and the hybrid test is 93.17% and 98.91%, respectively, which performs best than the other models. Meanwhile, the author verifies that the generation model without the pre-trained classifier can maintain the diversity of generated images, which proves the effectiveness of category constraints. In Du et al. (2021c), a similar model namely multi-constraint GAN, is proposed by the same team, however, the pre-trained classifier is changed from ResNet18 to ResNet16. Then, the signal-to-clutter ratio (SCR) is applied to evaluate the generation performance and multiscale structure similarity (MS-SSIM) is applied to evaluate the diversity of the images generated by the proposed model. The MS-SSIM of the proposed model is 0.158, which performs optimally compared with other models.

Besides, in Luo et al. (2020), an improved progress growing of GAN (PGGAN) is applied to generate the minority class SAR images to alleviate imbalanced image distribution. Firstly, the minority category images can be expanded by using the Auto-Augment proposed by Google Brain, which is constructed using Bayesian optimization to replace traditional reinforcement learning. Then, PGGAN is applied to balance the training process of the generator and discriminator. The architecture of the model is shown in Fig. 16. Furthermore, the loss function in WGAN-GP is applied to improve stability while training. The dataset of MSTAR is applied in this experiment for training and validation. Finally, the recognition accuracy can be improved by about 11.68%, which can achieve 97.36% in total.

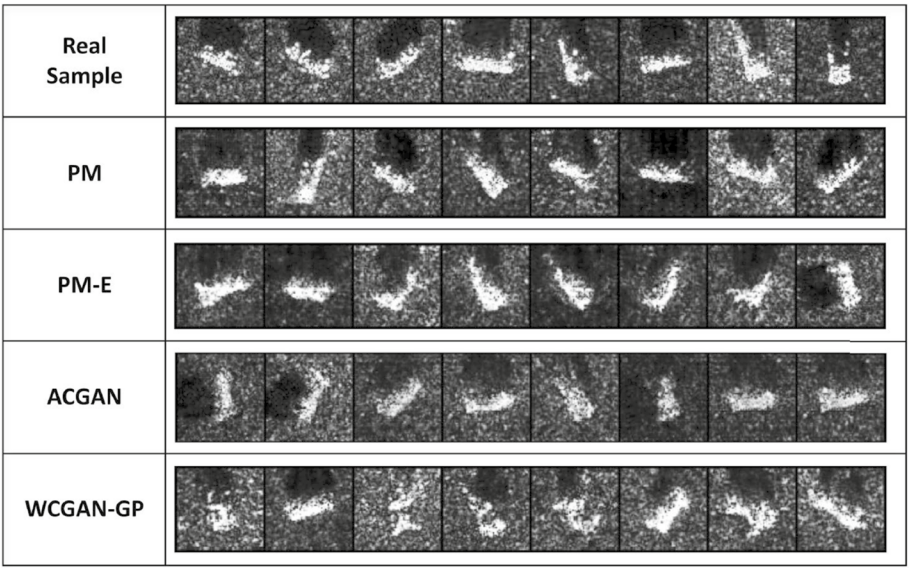


Fig. 15 The generation performance of the proposed model and different models [in Du et al. (2021d, Fig. 3)]

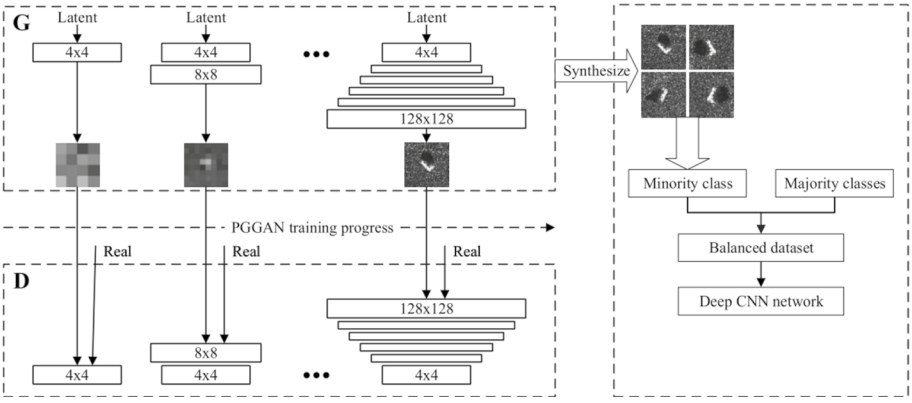


Fig. 16 The architecture of the applied PGGAN [in Luo et al. (2020, Fig. 2)]

2.5 CycleGAN-based methods

CycleGAN is usually applied in the task of domain transformation, however, it can also be applied for data augmentation by fusing the segmented target and the background. In Zhang (2021), CycleGAN and semantic segmentation are applied for SAR data augmentation. Firstly, the median filter, Gaussian filter, and bilateral filter are compared for pre-processing of SAR images, and the median filter is selected to reduce speckle noise and increase the contrast between the target and the background. However, by comparing the generation performance of traditional sample augmentation technology based on texture

modeling, image augmentation technology based on CycleGAN, and image augmentation technology based on attention CycleGAN, the existing approaches cannot meet the requirement of the high quality in the task of image generation. The main idea of this paper is, the target in the image can be extracted using semantic segmentation, and GAN can be applied to expand the image of the background. The segmented target can be fused with generated background, meeting the requirement of data augmentation. Hence, the DeepLabV3 is applied to segment the target and background in the optical images. Then, CycleGAN is applied for domain transformation between the segmented target and the background. After that, the generated target and background are permutated and combined as a new SAR image. Moreover, a sample quality evaluation approach based on object detection is proposed, which is constructed based on MobileNetv3 and YOLOv4, and the availability of generated samples can be verified through different training set ratios (real images: generated images). Finally, in a training dataset consisting of 300 real images and 3000 generated images, the accuracy achieves 90.6%, which verifies that the generated samples play a supporting role in the training of the deep neural network. The example of generated images is shown in Fig. 17.

However, in the process of model iteration, the noise features in SAR images are constantly amplified, which leads to unclear target contours in the generated images and is not conducive to subsequent image segmentation tasks. In Zhao et al. (2020), an improved CycleGAN is applied for data augmentation of SAR images to improve the segmentation performance. Therefore, a median filter based on the sliding window is used for denoising. After that, U-Net is applied for SAR image segmentation including two modifications. First, the cosine announcing cycle is applied to change the learning rate. Second, the focal loss is applied here to replace the traditional Dice loss. Finally, intersection over union (IoU) is applied to evaluate the segmentation quality. The results of the images enhanced by some simple methods such as flipping and rotation are compared. The DOTA and the NWPU VHR-10 dataset are applied for training and evaluation. To evaluate the generation performance, mean, variance, standard deviation (Std), the equivalent number of looks (ENL), and radiometric resolution (RR) are applied, which achieves 49.99, 922.64, 30.37, 2.60, and 2.06 for land area and 36.46, 682.34, 26.12, 1.90 and 2.34 for ship target, respectively.

2.6 Pix2PixGAN-based methods

In Long et al. (2019), Pix2PixGAN is applied for data augmentation and an improved SSD is designed for ship detection in SAR images. Firstly, the edges of SAR images in the dataset can be extracted using the Canny detector. The architecture of the generator is the

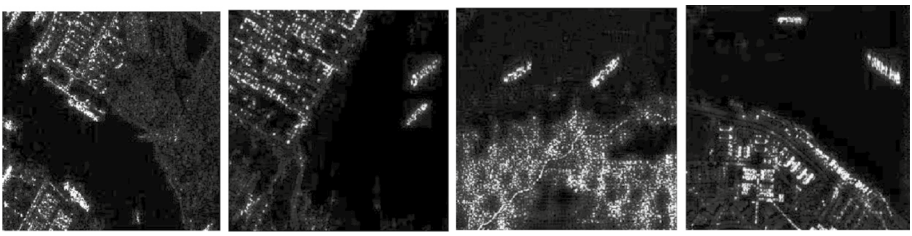


Fig. 17 The generation results of SAR images in the complex environment [in Zhang (2021, Fig. 5.10)]

same as that of Pix2PixGAN. Then, in the design of the discriminator, the Markov method is applied to calculate the mean value of all patches. The generated images are shown in Fig. 18. Moreover, the Inception blocks are applied to improve the SSD, which can extract the features in different receptive fields. Finally, the detection accuracy after adding the generated images is increased from 80.3 to 82.2% by using the improved SSD.

2.7 Auto-encoder-based methods

In Wang et al. (2018a), an improved autoencoder is developed based on Wasserstein distance (WAE) and applied for data augmentation in the task of ATR. The architecture of the proposed WAE is shown in Fig. 19. Firstly, random noise is combined with conditional information as the input of the encoder. Then, the noise and conditional information are passed through the decoder to learn the distribution and the reconstructed data can be obtained. Moreover, a discriminator is trained to distinguish the real and generated image. Furthermore, the residual blocks are applied in the encoder and decoder for better extraction of the features. And the conditional information consists of the category and azimuth information. The dataset of MSTAR is applied to measure the performance of the proposed model. Finally, the FID of 62 can be obtained using the proposed WAE and the recognition accuracy of 94.8% can be calculated.

In Huang et al. (2022), a variational autoencoder (VAE) is applied here as the generator of the WGAN-GP, namely VAE-WGANGP. Furthermore, the loss function is improved to solve the problem of erratic training, which contains four parts. The first part is the difference between the generated image and the real image namely feature reconstruction loss, showing the consistency of spatial correlations and quality. The second part is the KL divergence loss between the prior distribution and the Gaussian distribution. The third one is the adversarial loss between the real image and generated image calculated by the discriminator. The last part is gradient penalty loss, which is applied to improve the stability and performance of the proposed model. In the task of ATR, the proposed method achieves an accuracy of 94.9%, which performs optimally.

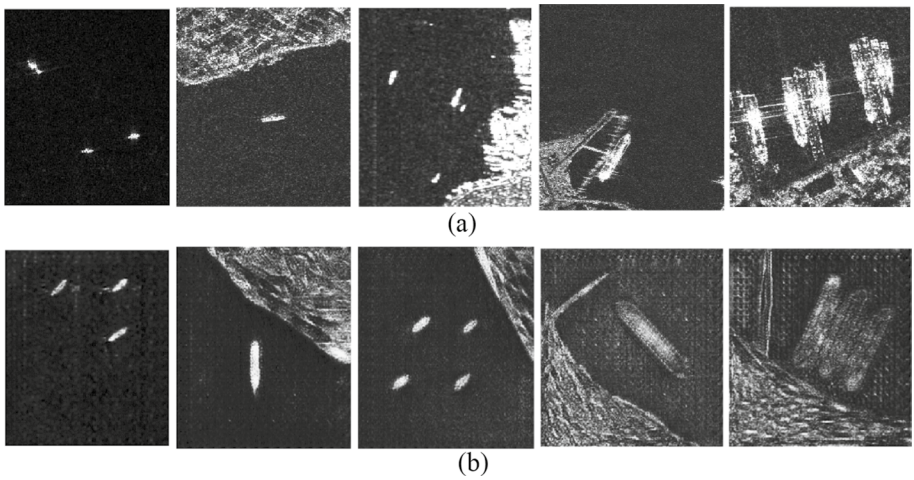


Fig. 18 The generated images. **a** The real images; **b** the generated images [in Long et al. (2019), Table 2)]

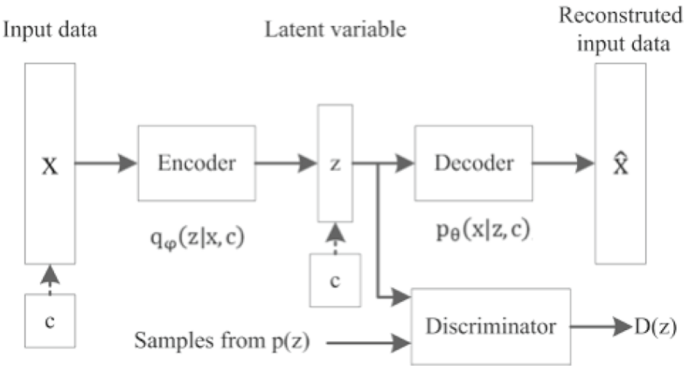


Fig. 19 The architecture of the proposed WAE [in Wang et al. (2018a, Fig. 1)]

In Du et al. (2021a), a generator is proposed based on the conditional probabilistic model to generate the ISAR images with high resolution. Firstly, a similarity measurement based adversarial variational autoencoder is proposed, which can receive a set of gaussian noise as the input and learn the mapping from the noise to a high-resolution ISAR image. Then, the sqrt-cosine similarity is applied to extract the region with a larger value in the ISAR image. In addition, KL divergence is applied to replace the original L2 loss to constrain the distribution using conditional information. Moreover, conditional information is introduced to learn the co-features in ISAR images and improve the generation performance. The structure of the proposed model is shown in Fig. 20.

Then, a CNN-based model is applied to obtain sufficient statistics of its approximate posterior probability distribution and SNR is used to evaluate the generation performance. The generation result is shown in Fig. 21.

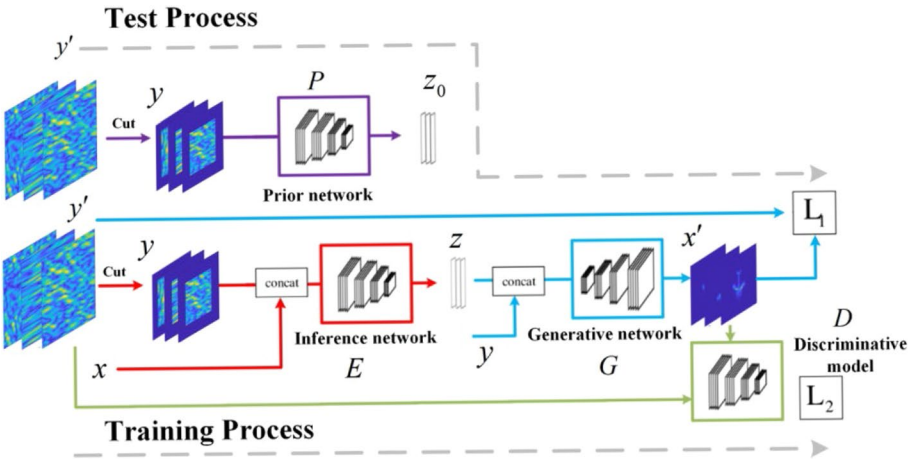


Fig. 20 The architecture of the proposed model [in Du et al. (2021a, Fig. 2)]

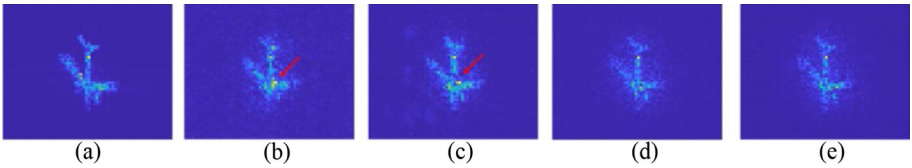


Fig. 21 The generation results using different model. **a** Ground truth; **b** CV-CNN; **c** GAN; **d** the proposed model without conditional information; **e** the proposed model with conditional information [in Du et al. (2021a, Fig. 9)]

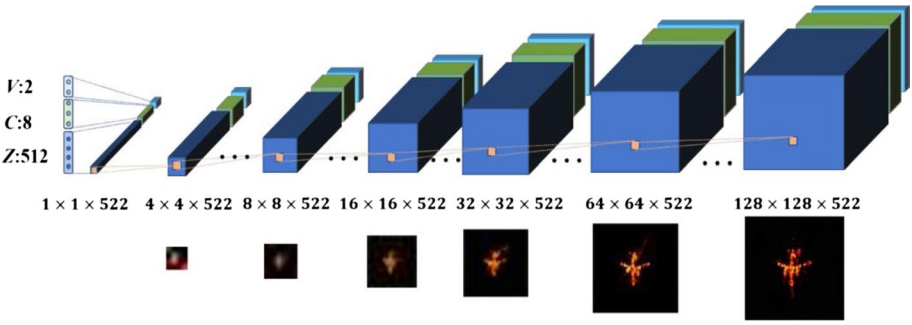


Fig. 22 The architecture of the generator [in Zhou et al. (2021, Fig. 1)]

2.8 ISAR image generation

In Zhou et al. (2021), cGAN is applied for ISAR image generation. The architecture of the generator is shown in Fig. 22. Firstly, the category and azimuth information is applied as the conditional vector and fused with random noise. Then, the features of the real ISAR images and generated images are extracted and distinguished using the discriminator. Meanwhile, a label predictor and an azimuth predictor are trained to predict the category and the azimuth of the generated images, respectively. Finally, the proposed method can achieve an SSIM higher than 0.7.

3 SAR/ISAR image generation for image translation

At present, the processing of remote sensing images tends to be realized using multimodal analysis. Multimodal analysis of the same area or target through SAR and optical imaging can complement the missing information of each mode. For example, SAR images focus on the structural features inside the target, while optical images focus on the details and texture features of the target surface. With the development of computer vision technology, the GAN-based models can be applied in various image styles transformation tasks, such as the translation between the optical image and oil painting image (Liu 2021), and the colorization of the sketches (Li et al. 2021). The additional condition information can be introduced into the GAN to make the model generate images of a specific style. Hence, through the translation of SAR and optical images, the abundant structural information of SAR

images can be extracted, which can be captured in all-weather, and the speckles in SAR images can be suppressed. Meanwhile, the SAR images can be translated into more intuitive optical images, which can be interpreted using the rapidly developed optical image processing approaches.

3.1 cGAN-based methods

In Niu et al. (2018); Enomoto et al. (2018), cGAN is applied in the task of SAR to optical image translation. The U-Net is applied as a generator due to the encoder-to-decoder architecture and L1 loss is applied as the loss function. In Niu et al. (2018), the result achieves the SSIM of 0.2695 and 0.2611, MSE of 0.0461 and 0.0281 in the tasks of optical to SAR and SAR to optical translation, respectively. In Enomoto et al. (2018), the proposed approach achieves a PSNR of 28.67 and SSIM of 0.08533.

However, the original cGAN cannot generate satisfactory translation images. In order to retain the abundant structure information of SAR images, the L1 norm is applied as a loss function to restrict the generation result closer to the original image. But the generation results only using the L1 norm are blurred and unsatisfactory. In Li et al. (2020b), the structure similarity index measure (SSIM) is applied and combined with the L1 norm to represent the structure and texture features between the generated image and the original image, which can help to generate the image with better structural information. Moreover, U-Net is applied as a generator to transfer the features between the input and the output. The result shows that networks like dualGAN and CycleGAN can reconstruct the texture and structural features between the SAR and optical images, but the color feature is missed. However, a network like pix2pixGAN can retain color information, but the details of texture and structure are missed. By comparing with other models, the proposed cGAN based on the L1 norm and SSIM has the optimal performance in the task of SAR-to-optical image translation, which achieves the SSIM of 0.353 and PSNR of 17.6747. The translation result is shown in Fig. 23.

Moreover, in Doi et al. (2020), an improved cGAN is applied to solve the problem of extensive color deviation. Firstly, a region classifier is pre-trained to learn the region information to map the information from the distribution of SAR images to the distribution of the optical images. Then, the features extracted by the pre-trained region classification network are concatenated with the features in the generator and transferred to the discriminator at the same time. For the discriminator, spectral normalization is applied to improve the stability of training. L1 loss and adversarial loss are combined to measure the difference between the real image and the generated image. The architecture of the proposed model is shown in Fig. 24. The SEN1-2 dataset is used for training and validation. Finally, the proposed model is compared with Pix2PixGAN and can achieve a PMSE of 49.3, PSNR of 15.74, and SSIM of 0.04177.

Besides, in Yang et al. (2022b), an improved CGAN (ICGAN) is applied for the task of SAR-to-optical image translation, addressing the problem of blurred contour, the disappearance of texture, and wrong color simultaneously. There are three main modifications for the problem of contour, texture feature, and color, respectively. Firstly, a parallel feature fusion generator is designed, including an up-sampling pipeline and a down-sampling pipeline. The convolutional kernels with small size are applied in the up-sampling pipeline for speckle noise erasing, and the convolutional kernels with 7×7 and 3×3 are applied in the down-sampling pipeline to extract the texture and semantic features, respectively. After that, mirror filling is applied for padding to replace the original zero padding, which has

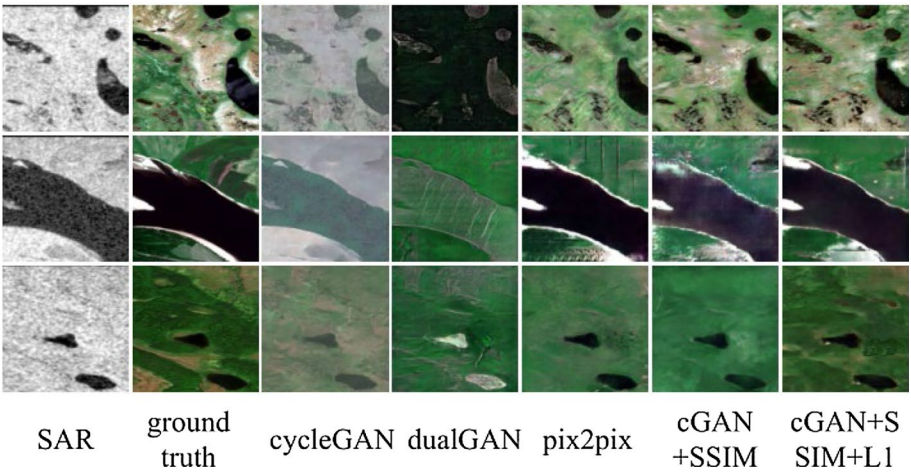


Fig. 23 The comparison result of several models [in Li et al. (2020b, Fig. 2)]

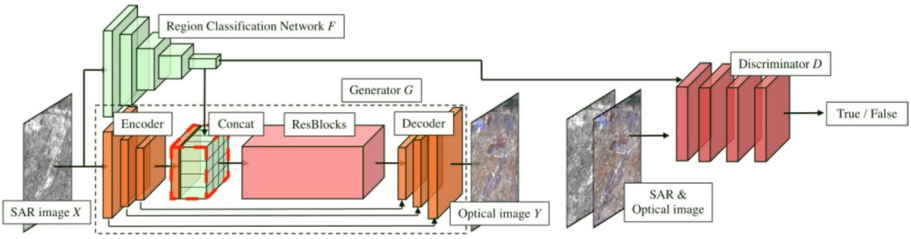


Fig. 24 The architecture of the proposed model [in Doi et al. (2020, Fig. 2)]

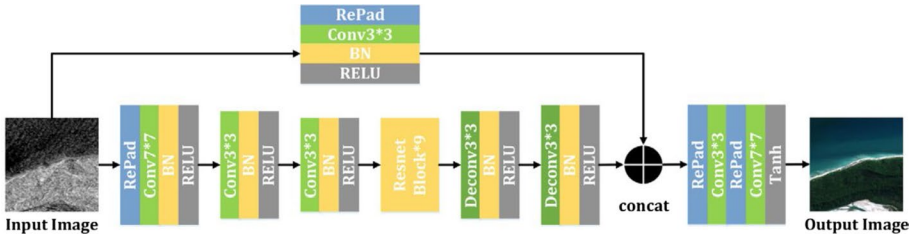


Fig. 25 The architecture of the generator in ICGAN [in Yang et al. (2022b, Fig. 3)]

better performance and clearer contour of the generated images. The architecture of the proposed generator is shown in Fig. 25. Then, a multi-scale discriminator is designed to address the weak texture feature, which consists of three different classifiers with different receptive fields. Moreover, the MSE loss between the real images and the generated images is combined with L1 loss and adversarial loss, addressing the problem of color inconsistency. The dataset of SEN1-2 is applied for training and validation. Finally, VGG19 and ResNet50 are applied to evaluate the performance of the proposed model. The proposed ICGAN can achieve a PSNR of 18.8111, SSIM of 0.3959, MSE of 0.1296, and NIQE of

46.6666. For classification, the ICGAN achieves accuracy of 97.73% and 97.97% using VGG19 and ResNet50, respectively.

3.2 DCGAN-based methods

In addition, DCGAN can replace the original GAN as the new backbone of the generation architecture due to the excellent feature extraction capability of the convolutional operation, and there exist some works using the modified DCGAN for image translation. In Zhang et al. (2020), a feature-guided SAR to optical image translation approach is proposed, considering the specific feature of multimodal images, such as the high-frequency noise for SAR image and optical image in some complicated environments. The architecture of the proposed model is shown in Fig. 26.

The generator consists of three parts, including a SAR image encoder, a SAR-to-optical image translator, and an optical image decoder. VGG-based model is applied for the construction of a SAR image encoder, which can extract the deep features of the SAR image with continuous down sampling. Then, nine ResNet blocks are applied for the translation from SAR to optical images. The optical image decoder includes several upsampling operations using transposed convolution blocks, which can output the translated results with the same size as input images. In the discriminator, a pyramid architecture based on PatchGAN is applied to discriminate high-resolution and low-resolution images respectively and calculate the output matrix and label matrix. This multi-scale discriminator can help the network focus on global features and local features of the images. Besides, a VGG19 is pretrained to extract the modal and texture features between the SAR ad optical images. At the same time, feature matching is performed before each max pooling layer to stimulate the

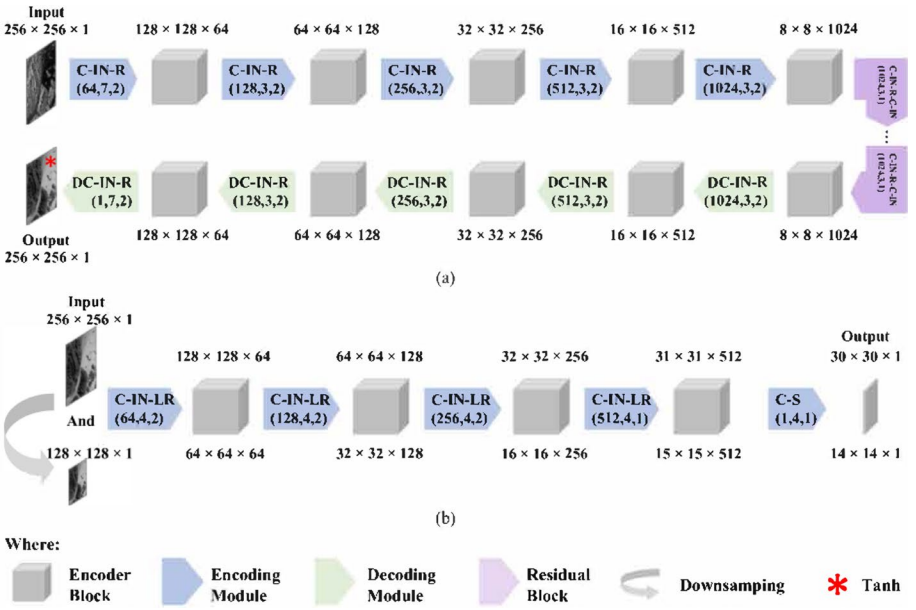


Fig. 26 The architecture of the proposed model. **a** Generator; **b** discriminator [in Zhang et al. (2020, Fig. 2)]

generator can output the images with the modal information and texture features of the real images. Finally, discrete cosine transform (DCT) is applied to eliminate high-frequency noise in SAR images, and the DCT L1 loss between the real images and generated images is calculated for optimization. The SEN1-2 dataset is applied for training and evaluation of the generation performance of the proposed model. The proposed model achieves SSIM of 0.61, FSIM (Zhang et al. 2011) of 0.79, PSNR of 21.57, and MSE of 0.01, which performs optimally by comparing with other models.

In Fu et al. (2021), an improved DCGAN is applied for translation between SAR and optical images. Firstly, a DCGAN with a multi-scale residual connection is designed, which can be applied for SAR-to-optical and optical-to-SAR image translation with an encoder-to-decoder architecture. The architecture of the proposed generator is shown in Fig. 27. The application of multi-scale residual connection can fuse the feature maps of current and former output, which can increase the receptive field of the decoder. After that, the L1 loss is fused with adversarial loss to translate the image with abundant detailed information. The proposed model can achieve the optimal SSIM of 0.2595 and 0.2799, PSNR of 15.9171 and 15.5820, FID of 53.0067 and 154.7532 for the task of optical-to-SAR and SAR-to-optical image translation, respectively.

In Guo et al. (2021), an edge-preserving convolutional GAN is designed for SAR-to-optical image translation, namely EPCGAN. The complexity and speckle noise of the noise is neglected using the existing GAN, and the generation performance is unsatisfactory. Firstly, the structural information is extracted as a feature map, and the content-adaptive convolutional filter is applied for feature map convolution operation. The weight of content adaptive filters is defined as the weight multiplication of the standard convolutional filter and edge-preserving convolutional filter calculated by the content component. Then, the Sobel operator is applied to realize the edge-preserving convolutional operation for edge enhancing of the optical images. After that, the gradient of the images is extracted as the texture features, whose deep features can be extracted using convolutional operation. In addition, a novel gradient branch is designed, which can simulate the gradient of optical images using the content information, and the gradient

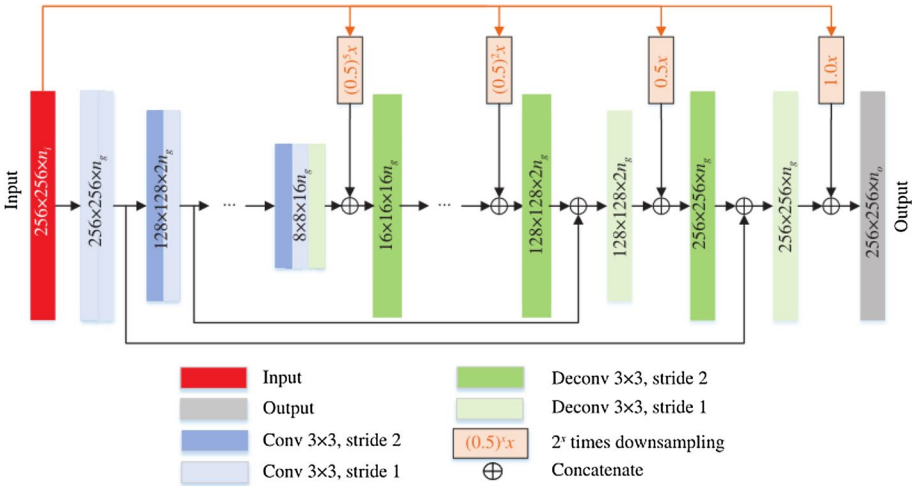


Fig. 27 The architecture of the proposed generator [in Fu et al. (2021, Fig. 2)]

can be fed back to the generation network for structural feature restoration. Hence, the proposed model can improve the structural feature and details of the translated optical images. The architecture of the proposed EPCGAN is shown in Fig. 28.

In the process of training, MSE loss is added to measure the difference between the real and generated images for better translation performance. The SEN1-2 dataset is applied for training and validation. Finally, the proposed approach can achieve the PSNR of 19.3627, MSE of 0.0151, and SSIM of 0.4771, the comparison of translation is shown in Fig. 29.

In Shao et al. (2022), an improved DCGAN is designed to solve the problem of color distortion and detail loss in the tasks of SAR to optical image conversion. The generator consists of an encoder, a SAR-to-optical conversion model, and a decoder for the optical image. In the encoder, atrous spatial pyramid pooling (ASPP) is applied to extract the multi-level features of the image. Then, ResNet is applied to construct the SAR-to-optical conversion model and the transposed convolutional operation is applied for optical image decoding. The architecture of ASPP is shown in Fig. 30, the dilation convolutional operation with a kernel size of 1×1 and 3×3 is applied to extract the feature of the image in the different receptive field. The SEN1-2 dataset is used for validation, and PSNR and SSIM are used for evaluation, which is increased by about 20%. The proposed method can achieve the PSNR of 24.956, 25.315, 26.688, and 26.241, and SSIM of 0.810, 0.836, 0.861, and 0.838 for forest, farmland, city, and desert, respectively.

In Yu et al. (2022), an improved DCGAN is designed for SAR to optical image translation with abundant texture features and color fidelity, namely HFGAN. A heterogeneous fusion generator is designed to fuse the global detailed features, including a coarse-scale network, a fine-scale transformation network, and a gated fusion network. The coarse-scale network is inspired by the architecture of U-Net, which can retain the features between the encoder and the decoder, and the fine-scale translation network is constructed using residual blocks to extract the detailed features in the image. After that, a gated fusion network is designed to combine the extracted features, whose architecture is shown in Fig. 31.

The gated fusion network can improve the generation ability of the important part adaptively with a self-attention mechanism, and the attention map can be calculated by a series of concatenation and convolution operations. In addition, the architecture with a multi-scale receptive field and spectral normalization is designed as the discriminator, which can help to fuse the local and global features and improve the stability for

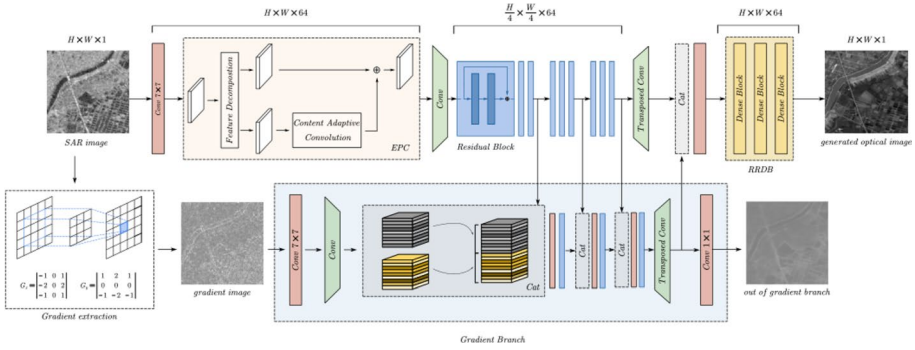


Fig. 28 The architecture of the proposed EPCGAN [in Guo et al. (2021, Fig. 4)]

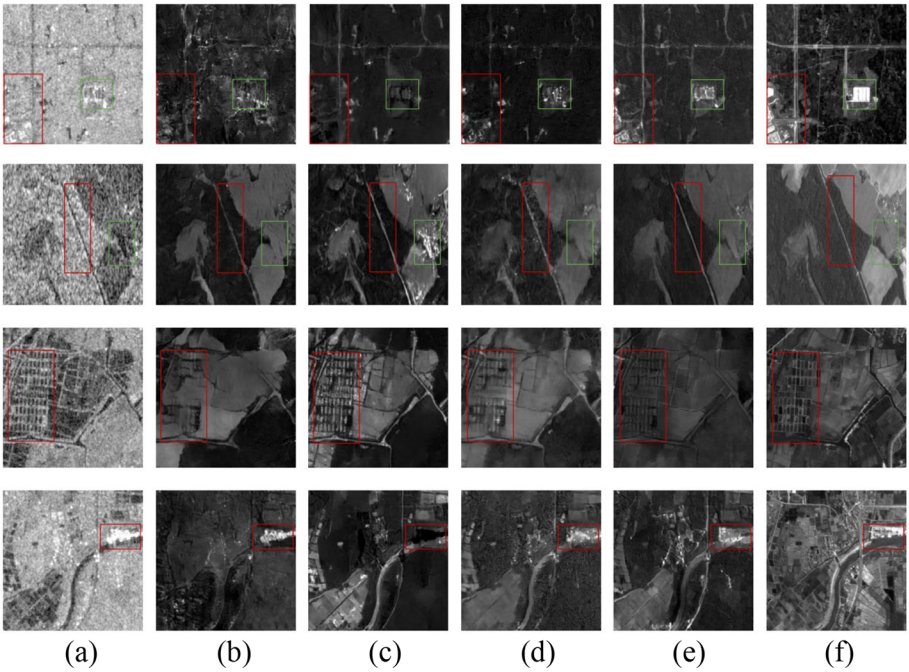


Fig. 29 The comparison of translation results. **a** SAR image; **b** Pix2PixGAN; **c** CycleGAN; **d** s-CycleGAN; **e** EPCGAN; **f** optical ground truth [in Guo et al. (2021, Fig. 5)]

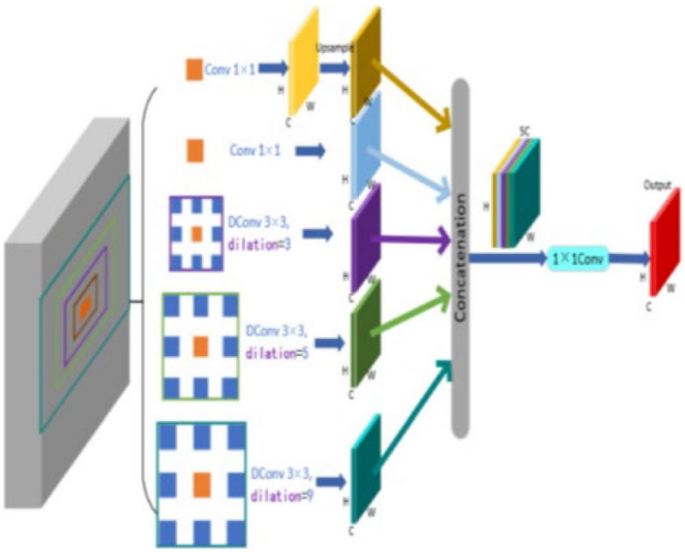


Fig. 30 The architecture of the proposed ASPP [in Shao et al. (2022, Fig. 3)]

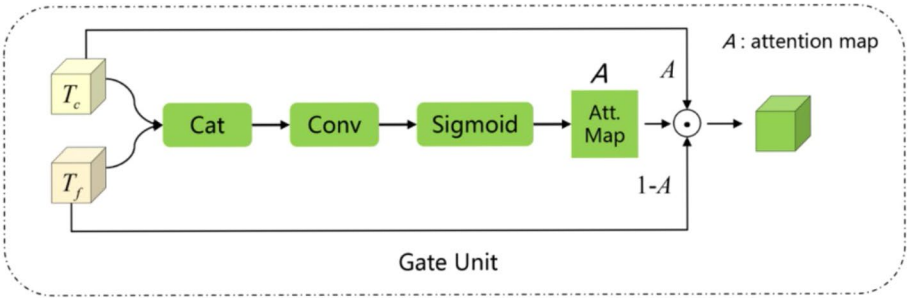


Fig. 31 The architecture of gated fusion network [in Yu et al. (2022, Fig. 3)]

training. The SEN1-2 dataset is applied for training and testing. The proposed model achieves a PSNR of 11.8467, SSIM of 0.0617, IS of 1.00532, FID of 170.1058, and spectral angle mapping (SAM) of 1.56183.

In Zhou (2020), SAR images are generated by image inpainting, and a random mask method is proposed to improve the diversity of generated samples. The main improvements are the local binary pattern (LBP) is applied to improve the loss function of the discriminator in WGAN-GP, hence the network can learn the texture features of images; the random mask and image completion are applied to generate a SAR image, the masked image and the masked part are inputted for training, and the masked image is applied for testing. The training dataset includes MSATR and SAR-I data. The proposed approach can achieve the classification accuracy of 96.592%, 95.210%, 91.582%, 88.457% and 68.421% using the dataset with 40, 20, 10, 5 and 1 real images. The example of generated images is shown in Fig. 32, showing the satisfactory generation performance of the proposed method.

3.3 Pix2PixGAN-based methods

In Lange (2019); Dietrich-Sussner et al. (2021); Toriya et al. (2019), the traditional Pix2PixGAN is applied for SAR-to-optical image translation. In Lange (2019), the Pix2PixGAN and enhanced super-resolution generative adversarial networks (ESRGAN) are applied and compared for the optical image to SAR image translation tasks. The labeled optical images can be translated into labeled SAR images. Then ResNet50 and VGG19 are pre-trained and applied to evaluate the translation performance of the proposed model. The SEN1-2 dataset is used for training and testing. The result shows that the image generated by Pix2PixGAN performs better than the image generated by ESRGAN, and VGG is better than ResNet50 in the task of SAR image classification. In Dietrich-Sussner et al. (2021), the masks are applied to replace the random noise as the input of the generator, which can generate the predicted SAR images. Finally, the proposed Pix2PixGAN with the convolutional kernel size of 5×5 can achieve the optimal SSIM of 0.8492. In Toriya et al. (2019), Pix2PixGAN is applied for SAR-to-optical image translation and alignment. First, the paired SAR and optical images are inputted to the generator, and the fake optical images can be generated. Then, the wrong corresponding points can be removed according to the scale value and the gradient direction information of the corresponding points. The SEN1-2 dataset is applied for training and validation. The result of image alignment using the proposed method is shown in Fig. 33.

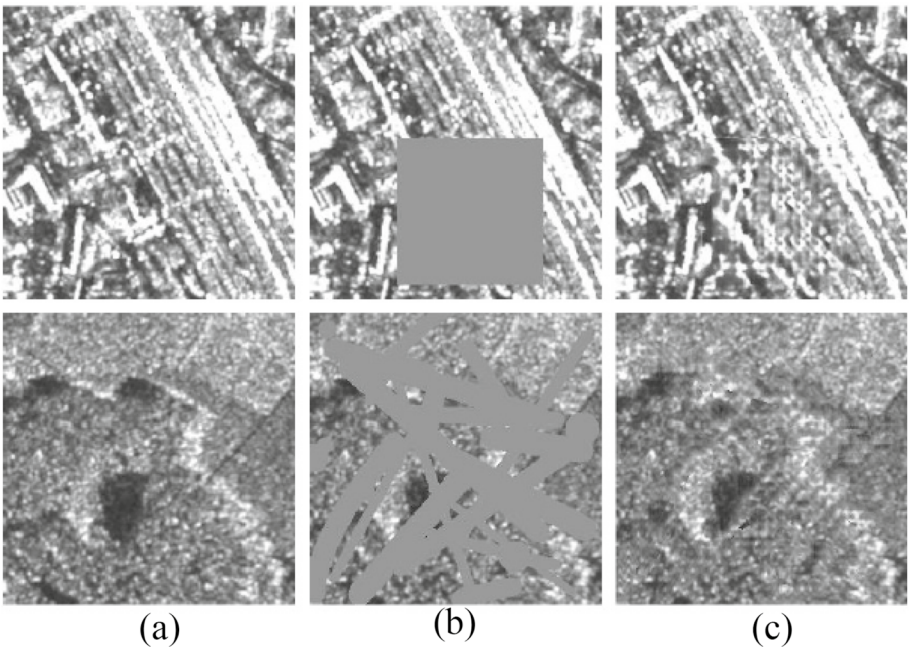


Fig. 32 The generation results of the proposed method. **a** Real images; **b** masked images; **c** generated images [in Zhou (2020, Fig. 5.7)]

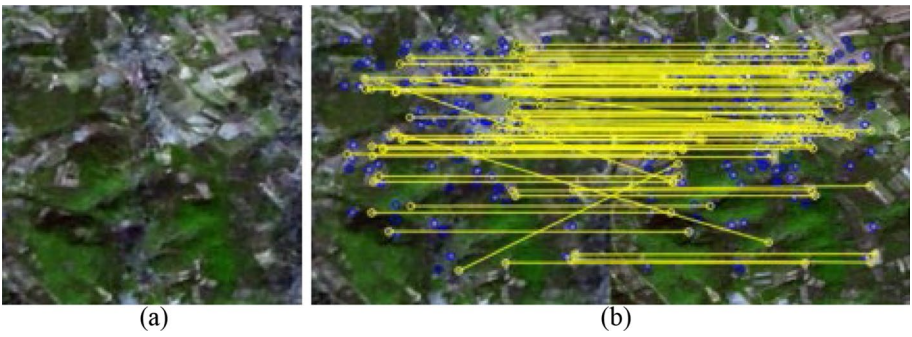


Fig. 33 The result of image generation and alignment. **a** The generated optical image; **b** the result of image alignment [in Toriya et al. (2019, Fig. 5)]

In the tasks of SAR-to-Optical translation, CycleGAN, and Pix2PixGAN are most applied. However, CycleGAN can retain the structural and texture information, but the category information of different targets is usually missed. In contrast, Pix2PixGAN can retain detailed inner information of different categories, but the generated image is always blurred. To solve the problems above, in Zuo and Li (2021); Zhang et al. (2021b), the texture features are extracted to improve the translation performance. In Zuo and Li (2021), a histogram of orientated phase congruency (HOPC) is designed as the image intensity descriptor of different modalities, which can extract the structural information between

SAR and optical images efficiently. The WHU-SEN-City dataset is applied for training and validation. The PSNR, SSIM, and chrominance feature-similarity index (FSIMc) of the proposed method achieve 11.203, 0.379, and 0.727, respectively. The generation result is shown in Fig. 34.

In Zhang et al. (2021b), the GLCM feature is extracted and Pix2PixGAN is applied to evaluate the influence of edge and polarization information for the results of SAR-to-optical image translation. Firstly, the optical and SAR images are both segmented into small patches. Then, the color information of the optical images and texture information of SAR images are extracted as feature vectors for training. After that, Pix2PixGAN is applied for SAR-to-optical image translation using the image pairs including SAR images with three types of polarization approaches and optical images. Finally, a CNN is applied for the classification of generated images and to evaluate the translation performance. The dataset of Sentinel-1 is applied in this experiment for analysis. The gray level concurrence matrix (GLCM) and the edge detected by the Canny operator are extracted as feature vectors, Then, an improved Pix2PixGAN is applied as a generator to be compatible with SAR images in single-channel and double-channel, as well as optical images with multiple spectral. And PatchGAN is applied as discriminator. After translation, the result shows that the Pix2PixGAN with edge information can restore the boundary of the target better and make it clearer. The translation result is shown in Fig. 35. In addition, in the task of classification, the double-channel SAR image with the condition of GLCM and Canny can achieve the most similar distribution with the real images, which has 20.11% of vegetation (15.09% in real image), 11.86% of water bodies (12.41% in real image) and 68.03% of building land (72.50% in real image).

Due to the paired SAR-optical images being hard to the acquisition, the networks requiring image pairs cannot be widely applied. Hence, there is a novel model designed to relax the limitation of SAR-optical image pairs, which are necessary while training a traditional Pix2PixGAN. In Ji et al. (2020), a cycle-consistency GAN is designed for SAR image colorization. Firstly, a one-hot vector is applied to represent the different region, like the water bodies and the desert. Then, the image is concatenated with the vector as the input of the generator. The proposed model consists of two pipelines, including forward and backward cycles with two generators and two discriminators. In the process of the forward cycle, the first generator is applied to translate the real SAR image to a fake generated optical image. Then, the second generator can translate the generate an optical image to a fake generated



Fig. 34 The generation result of the proposed method. **a** SEN-1 SAR image; **b** SEN-2 optical image; **c** Generation image [in Zuo and Li (2021, Fig. 3)]

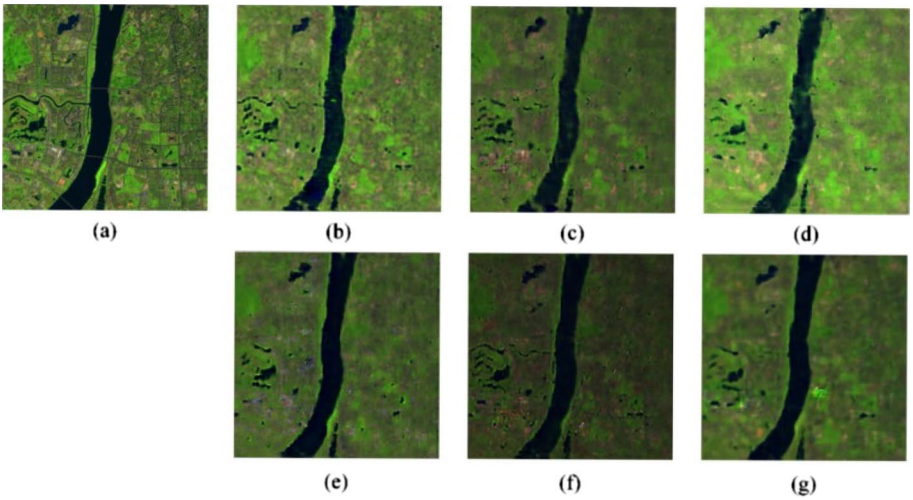


Fig. 35 The generation results. **a** Real image; **b** VV SAR image with GLCM; **c** VH SAR image with GLCM; **d** double-channel SAR image with GLCM; **e** VV SAR image with GLCM and Canny; **f** VH SAR image with GLCM and Canny; **g** double-channel SAR image with GLCM and Canny [in Zhang et al. (2021b, Fig. 5)]

SAR image. After that, the first discriminator is applied to distinguish the real and generated optical images. In the process of the forward cycle, the second generator translates the real optical image into a fake SAR image, and the first generator can translate the fake SAR image into a fake optical image. Then, the second discriminator is applied to distinguish the real and generated SAR images. The architecture of the proposed model is shown in Fig. 36.

The loss function is designed as the fusion of adversarial loss, cycle-consistency loss, and multidomain classification loss. The cycle-consistency loss does not require paired images, and the number and order of images are arbitrary, which is defined as the L1 loss between the real SAR image and the fake SAR image generated by generator 2, applying for the coloring task of the SAR image. In addition, to prevent incorrect coloring, a multidomain classification loss is used, including a classification loss of a real image for optimizing the discriminator and a classification loss of a fake image for optimizing the generator. The SEN1-2 dataset is applied in this experiment. The results show that although the generation performance is slightly worse than that of Pix2PixGAN, the pairing image is not applied, so the proposed method is more practical. Compared with CycleGAN, the

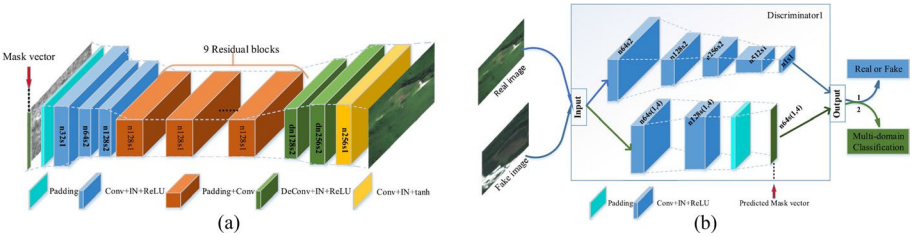


Fig. 36 The architecture of the proposed model. **a** Generator; **b** discriminator [in Ji et al. (2020, Figs. 2, 3)]

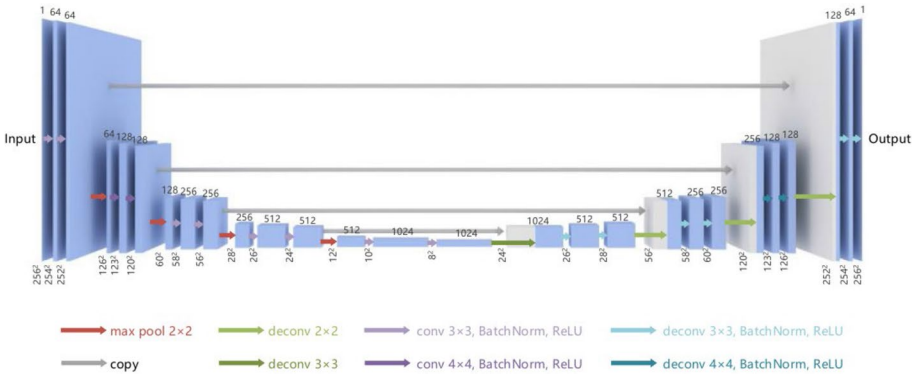


Fig. 37 The architecture of the proposed generator [in Sun et al. (2022, Fig. 2)]

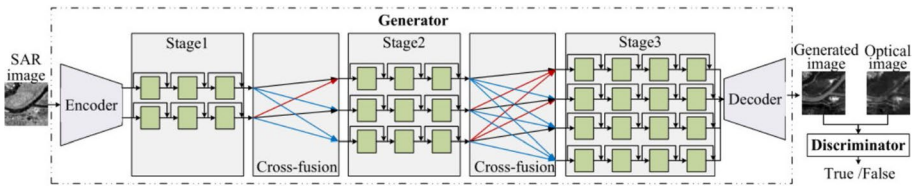


Fig. 38 The architecture of the proposed model [in Wei et al. (2022, Fig. 1)]

proposed network can generate an image with detailed information. Finally, the proposed approach can achieve a PSNR of 19.132, SSIM of 0.149, cosine similarity (COSIN) of 0.882, and FID of 95.901.

Most of the existing models use the existing pixel-level approaches, such as ResNet and U-Net, for SAR-to-optical image translation, which may cause the loss of features and blurred edges due to the semantic information lost in the generated image. Hence, the architecture of the generator can be modified to improve the translation performance. In Sun et al. (2022), the classical U-Net is improved, using an inhomogeneous convolutional kernel to replace the padding operation to maintain the original size of the input images and avoid adding the unnecessary edge information. The architecture of the proposed generator is shown in Fig. 37. After that, PatchGAN is applied to construct a discriminator to combine the local and global discrimination of whether the input image is real or fake. Moreover, LeNet is applied for the SAR image recognition task, which has simple architecture without complicated trick settings, hence it is more suitable to evaluate the performance of the proposed model by comparing it with other deeper networks. The simulated SAR and optical images are applied in this experiment. After translation, the optical images translated by the proposed model achieve a PSNR of 21.4738 and SSIM of 0.7420. And the category accuracy of ATR is 0.7508 (high than 0.6854 using original SAR images) and the orientation accuracy of 0.9533 (high than 0.9221 using original SAR images).

In Wei et al. (2022), an improved Pix2PixGAN based on feature cross-fusion inference is applied for the translation of SAR to optical image, which can retain the high-resolution and low-resolution features of the image and improve the performance of image translation. The architecture of the proposed model is shown in Fig. 38.

The proposed model consists of a multi-scale generator, including two, three, and four multi-scale branches in each step from high-resolution to low-resolution down-sampling layers, respectively. The multi-scale layers can reduce the size and increase the channels of the feature maps while down-sampling. Moreover, the down-sampling based on convolutional operations can fuse the feature maps of high resolution with low resolution, and the nearest-neighbor interpolation up-sampling is applied to fuse the feature maps of low resolution with high resolution. After that, a two-layer Gaussian pyramid is applied to construct the discriminator to distinguish real and fake images. In the design of the loss function, besides the adversarial loss based on the least-squares loss, the perceptual loss based on VGG19 and the feature matching loss are applied to achieve optimal performance. The SEN1-2 dataset is applied to evaluate the translation performance. Finally, the proposed approach achieves an MSE of 1902.5892, PSNR of 16.5955, and SSIM of 0.3004, showing satisfactory translation performance. The translation result is shown in Fig. 39.

In Shi et al. (2022), a Pix2PixGAN-based network is designed for SAR-to-optical image translation, namely generate-validate adversarial networks (GVANs). Firstly, a GAN with a cyclic structure is applied, including SAR-to-optical and optical-to-SAR translation channels. In addition, a validated image generator is applied to further optimize the blurred edge features. Then, an embedded layer based on U-Net is proposed to improve the feature extraction ability. The first layer of the generator is replaced by a whole U-Net to improve the adaptability of multidomain images. After that, dilation convolution layers are applied to reduce the influence of speckle noise, and four kinds of dilation convolution layers with different down-sampling rates are applied to extract the multi-scale texture features. In the decoder, the transposed convolutional operations are applied for up-sampling. The dataset of SEN1-2 is applied in this experiment. Finally, the proposed approach can achieve the PSNR of 23.007, 15.515, and 20.231, SSIM of 0.696, 0.306, and 0.586 in three regions of the dataset, whose translation result is shown in Fig. 40.

3.4 CycleGAN-based methods

The main difference between SAR and optical imaging is, SAR images focus on the physical properties of the target surface, and optical images focus on the details of the structure.

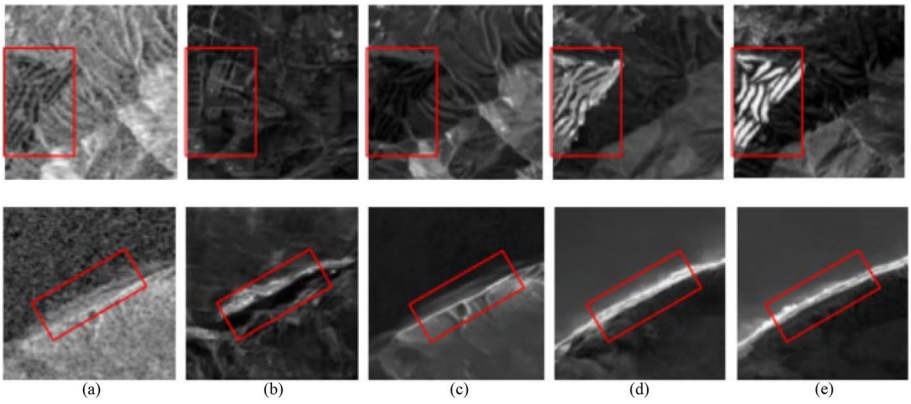


Fig. 39 The translation comparison of different models. **a** SAR images; **b** Pix2Pix; **c** CycleGAN; **d** proposed model; **e** optical images [in Wei et al. (2022, Fig. 4)]

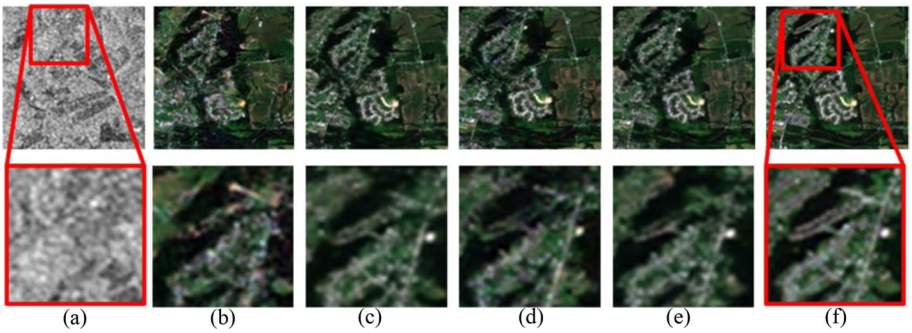


Fig. 40 The comparison of translation results. **a** SAR image; **b** Pix2PixGAN; **c** improved network; **d** GVAN; **e** real optical image [in Shi et al. (2022, Fig. 4)]

However, the traditional CycleGAN cannot meet the requirement of satisfactory translation between SAR and optical images, hence, several improvements to the original CycleGAN are applied for better translation results. In Fuentes Reyes et al. (2019), an adapted CycleGAN is applied for the tasks of SAR-to-optical image translation. Firstly, the learning rate and epochs are modified to achieve the same translation performance with fewer training times. Then, the Xavier algorithm proposed in Glorot and Bengio (2010) is applied to avoid parameters collapsing. Moreover, the size of the patch is set as 512×512, balancing the computational time consumption and translation quality. Furthermore, intensity cropping of the SAR images is applied to derive near-Gaussian distributions. Finally, DeepLabv3+ is applied to evaluate the translation performance using three indices of segmentation evaluation. The proposed model achieves IoU of 35.64%, Precision of 61.63%, and Recall of 45.78%. In Luo and Pi (2022), U-Net is combined with T-Net for SAR-to-optical image translation with high quality. By comparing the traditional CycleGAN with two generators, the proposed model has a single generator, because the images generated by CycleGAN tend to approximate the intermediate state between the SAR and optical domain. The architecture of the proposed generator is shown in Fig. 41.

The generator consists of two pipelines, including U-Net and T-Net. U-Net can help to remain the semantic information from encoder to decoder, and T-Net can help to fuse the semantic information of the bottom features and pixel-wise information of the top features,

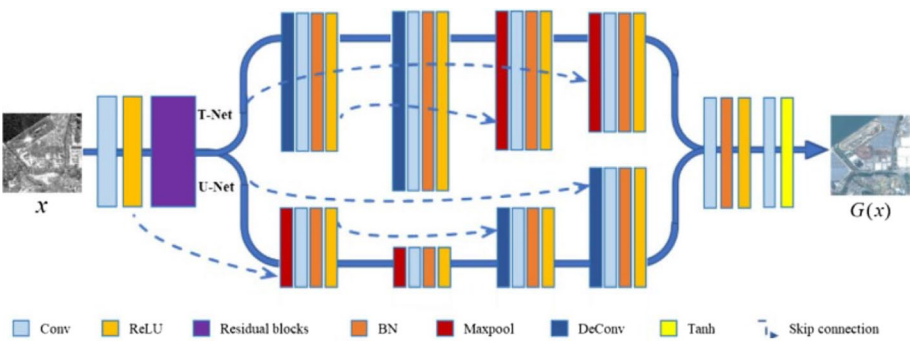


Fig. 41 The architecture of the proposed generator [in Luo and Pi (2022, Fig. 4)]

which can be applied to restore the lost features caused by max-pooling operation in the U-Net. After that, PatchGAN is applied to construct the discriminator. The SEN1-2 dataset is applied in the experiment. The generation result is shown in Fig. 42. Finally, the proposed model can achieve 0.73 in SSIM, 22.31 in PSNR, and 205.75 in FID.

In addition, the loss function can be improved to retain more detailed structural information. In Hwang et al. (2020), the generator is designed based on CycleGAN and least square GAN (LSGAN). The generated images cannot be continuously optimized when they are classified as real using the existing traditional GAN, and the generated images are repeated. Hence, the LSGAN is applied to punish the generated images whose distributions are far away from the real images. And SSIM loss is applied as the final loss function to retain the detailed texture and color information. Moreover, a network based on VGG-19 is applied to generate a perceptual loss. The generator is designed based on Pix2PixGAN, including the architecture of U-Net and PatchGAN. The WHU-SEN-City dataset is applied for training and validation. The proposed method can achieve the PSNR of 11.4347, 19.1925, and 9.8455, and SSIM of 0.2728, 0.5651, and 0.2258 in three regions. The result shows that the proposed method can translate the optical images with more structural details and few speckle noises. Besides, in Luo and Pi (2022), the perceptual loss and similarity loss are added in the traditional loss function to generate optical images with more structure and texture details.

In Wang et al. (2019b); Yang et al. (2022a), a supervised method is designed for the SAR-to-optical translation task. In Wang et al. (2019b), a cycle-consistent GAN (S-CycleGAN) is designed. The traditional CycleGAN is applied to learn the translation between two image domains by introducing the cycle consistency loss. Hence, the main idea of the design of Pix2PixGAN is applied in this paper to retain the same information between the SAR and optical images. And the MSE loss is added to the loss function for better translation. The original SAR images are split into small patches and the generated optical image patches can be obtained by using the trained S-CycleGAN. Then, the generated patches can be stitched together as a complete optical image. The dataset of SEN1-2 is applied here for training and validation. Finally, the proposed method can achieve a PSNR of 11.203, SSIM of 0.377, and FSIM of 0.727. In Yang et al. (2022a), an unsupervised approach is applied for SAR to optical image translation. The detailed architecture of the proposed model is shown in Fig. 43.

Firstly, a generator including a complicated encoder and a simple decoder is proposed, aiming to extract the abundant detailed information of the images in the encoder, and the principal information can be filtered using the decoder. Moreover, skip connection operations are applied in the encoder to fuse the features extracted by deep and shallow layers. Then, a multi-scale discriminator is proposed to improve the ability to discriminate the real image and generated image, including a local

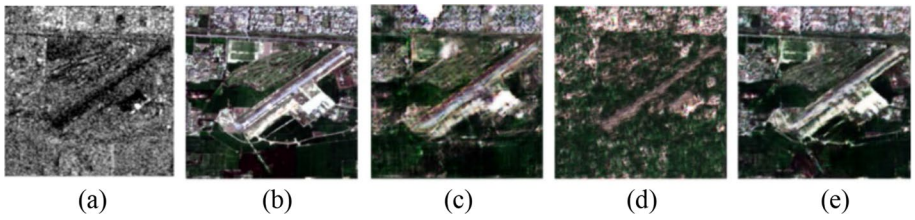


Fig. 42 The comparison of translation results. **a** SAR image; **b** Optical image; **c** U-Net; **d** V-Net; **e** U-Net + V-Net [in Luo and Pi (2022, Fig. 9)]

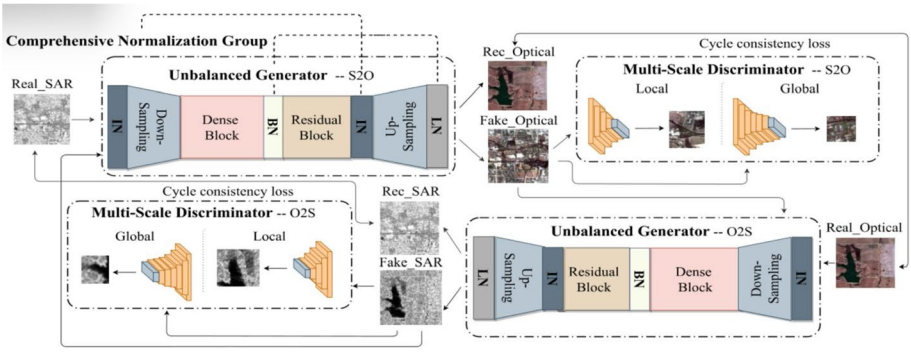


Fig. 43 The architecture of the proposed model [in Yang et al. (2022a, Fig. 1)]

discriminator and a global discriminator for local and global information extraction, which are designed based on PatchGAN. The global discriminator includes five down-sampling layers for a bigger receptive field, and the local discriminator includes three down-sampling layers to ensure the overall style of the generated image could be similar. In addition, based on the contrast experiment, the author found that LN layer can change the style features of the image (used in the residual network), while BN layer can obtain enough information of the feature map in the network (used in the full connection layer), and IN layer focuses on the change of image texture (used for up sampling). Therefore, a comprehensive normalization group, which consists of LN, BN, and IN, is proposed to ensure the color and texture information remain in the same distribution for the unsupervised generated SAR-to-optical image. The SEN1-2 and WHU-SEN-City dataset are applied for training and validation. Finally, PSNR, SSIM, FID, natural image quality evaluator (NIQE), and learned perceptual image patch similarity (LPIPS) are applied for evaluation. The proposed method can achieve PSNR of 11.22 and 10.93, SSIM of 0.1311 and 0.0862, NIQE of 25.9 and 21.5, FID of 103.1 and 86.8, LPIPS of 0.602 and 0.571 for SEN1-2 and WHU-SEN-City dataset, respectively. In addition, the translation results of the models are evaluated through visual comparison, indicating that the proposed model can solve the problem of color texture detail distortion.

3.5 ISAR image translation

In addition, due to the satisfactory translation performance of the Pix2PixGAN, it can be applied for translation between the ISAR masks and images for target segmentation. As an improved cGAN, Pix2PixGAN is applied in Li et al. (2019a) to generate the ISAR images. Firstly, the optical images are translated into gray-scale images to eliminate the influence of the color information. Then, the noise in the images are removed using pre-processing. After that, the ISAR images are used as the input of the generator, which can output the corresponding optical images after training. The generation result is shown in Fig. 44. Finally, the generation results are paired with the original real ISAR images as a new dataset for training and validation. The combined dataset can achieve the ATR accuracy of 96.67%, while the original ISAR dataset can achieve the accuracy of 87.50.

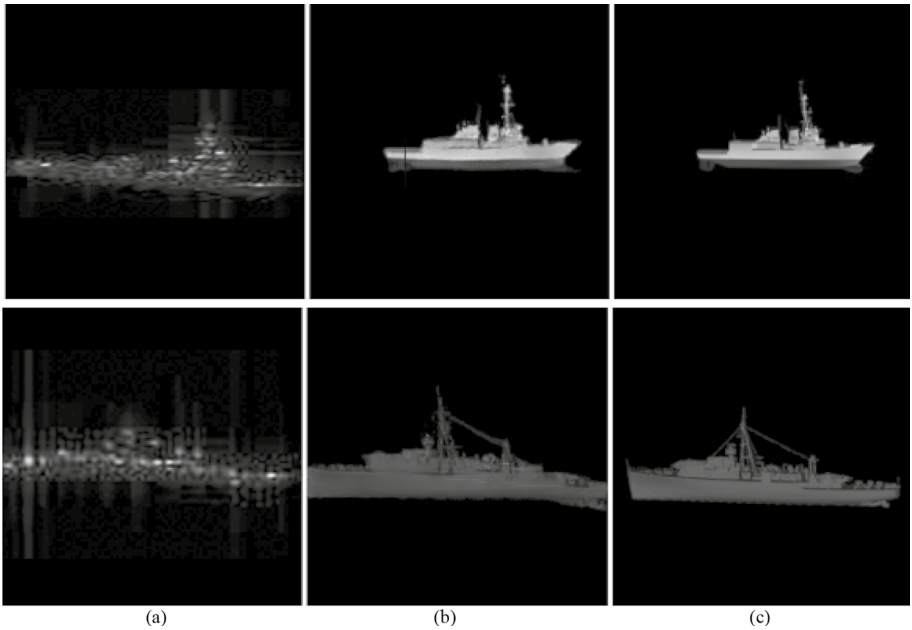


Fig. 44 The generation result using Pix2PixGAN [in Li et al. (2019a, Fig. 5)]

In Du et al. (2021b), the generator can receive a real ISAR image and a set of random noise as the input. Then, the real mask and generated mask can be paired with the real ISAR image and distinguished by the discriminator, which is constructed based on the

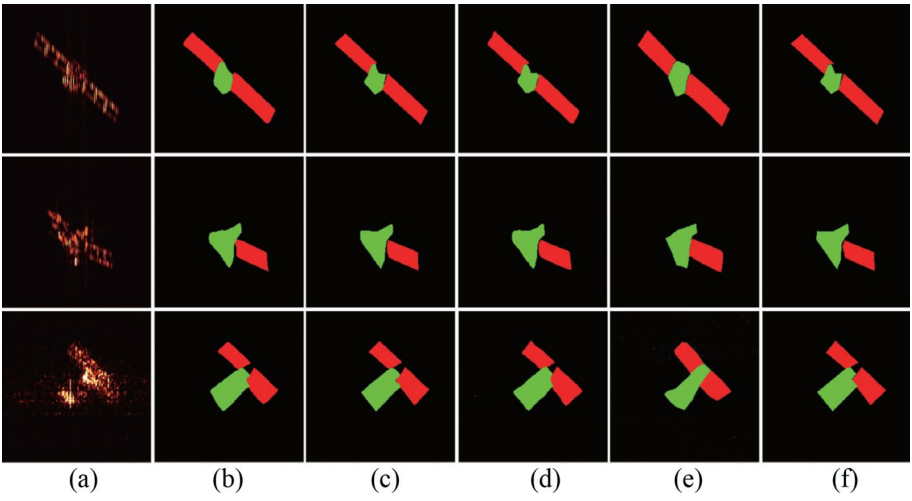


Fig. 45 The comparison of segmentation results. **a** The original ISAR image; **b** FCN; **c** SegNet; **d** DeepLabV3+; **e** the proposed method; **f** ground truth [in Du et al. (2021b, Fig. 8)]

Markov discriminator. The segmentation result is shown in Fig. 45. Finally, the proposed approach can achieve the MIoU of 77.95%.

4 SAR/ISAR image generation for image enhancement

In this section, the image generation approaches based on deep learning techniques for the task of SAR image enhancement are summarized, including contrast enhancement, super-resolution reconstruction, and dehazing of SAR images.

4.1 SRGAN-based methods

In Wang et al. (2018b), an improved GAN is applied for super-resolution reconstruction for SAR images. The architecture of the proposed SRGAN is shown in Fig. 46. Different from the traditional GAN, a low-resolution SAR image is applied to replace the random noise as the input of the generator, and the generator can output the generated super-resolution SAR image. Then, the real high-resolution image is distinguished from the generated super-resolution image using the discriminator. After that, the perceptual loss is combined with adversarial loss as the loss function. Finally, the proposed approach can achieve an MSE of 0.0014 and SSIM of 0.9083.

In Ai et al. (2021), a modified super-resolution generative adversarial network (SRGAN) is applied for contrast enhancement of ship targets in SAR images, namely ISRGAN. ISRGAN can retain the principal information of SAR images without prior information from the SAR imaging system. At the same time, the residual dense network (RDN) is used to combine the local and global features of the image to improve the image quality. In addition, the BN layer in the network is removed, which will damage the image structure and reduce the contrast between the background and the target. In the generator, RDN is applied to concatenate the convolutional features of different channels to obtain the local and global features of the image. At the same time, the

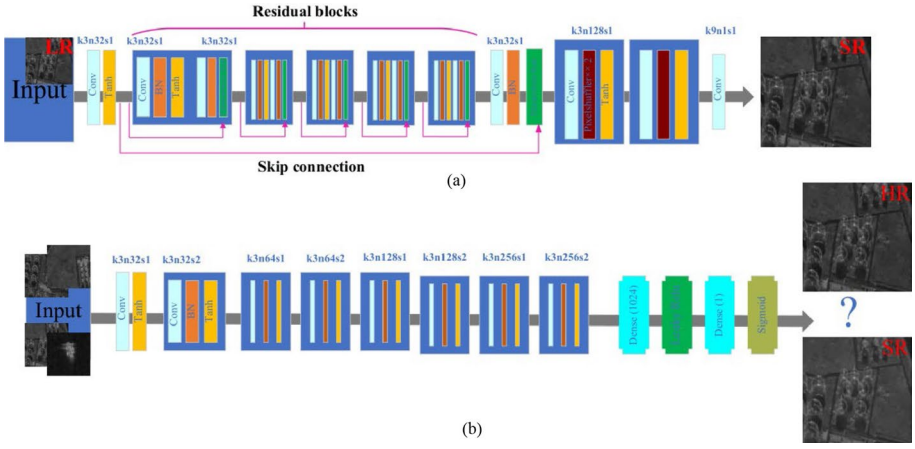


Fig. 46 The architecture of the proposed SRGAN. **a** The generator; **b** the discriminator [in Wang et al. (2018b, Fig. 1)]

features of different scales are fused using the residual connection, which can transfer the semantic information from the shallow network to the deep network, and the feature extraction ability and final generation effect of the proposed model are enhanced. At the same time, the content loss is added to the loss function of the original GAN, which is calculated from the original super-resolution and the output high-resolution image. The HRSID dataset is applied for validating the enhancing performance of the proposed model. Finally, the ratio of the intensity of the ambiguous region to the intensity of the target region (ATTR) achieves optimal in the compared four regions, which achieves − 1.1611, 1.2578, − 3.0703, and 3.7269, respectively. The performance of contrast enhancement is shown in Fig. 47.

4.2 DCGAN-based methods

In Gu et al. (2019), an improved DCGAN is applied in the tasks of SAR image super-resolution construction, namely noise-free GAN (NFGAN). Firstly, the low-resolution SAR images after bicubic interpolation are defined as the low-resolution version of the high-resolution SAR images. Different from the random noise input in traditional DCGAN, the low-resolution SAR images are inputted into the generator. Then the generated super-resolution SAR images are distinguished from the real high-resolution images in the discriminator. After that, a despeckling network is designed with the same architecture as the reconstruction network, including an encoder and a decoder. Then, the speckle noise is added in clean optical images, which are applied to pre-train the despeckling model. Moreover, the MSE loss between the generated super-resolution and real high-resolution images is added to the loss function. Finally, the proposed NFGAN can achieve a PSNR of 16.2466 and SSIM of 0.4823. In Zheng et al. (2019), a self-normalizing GAN (SNGAN) based approach is proposed for SAR image super-resolution reconstruction. The traditional GANs for super-resolution reconstruction are usually applying ReLU for activation and batch normalization (BN) to maintain the stable training state. However, the application of BN can normalize the features of the image, which is more proper to be applied in optical image classification. And SeLU is applied for activation in the generator. The architecture of the proposed model is shown in Fig. 48. Then, the TerraSAR and MSTAR datasets are applied to the experiment. Finally, the proposed model achieves a PSNR of 32.7646 and SSIM of 0.9240 using the TerraSAR dataset and can achieve the recognition accuracy of 97.69 and 94.47 under the standard operating condition (SOC) and the extended operating condition (EOC).

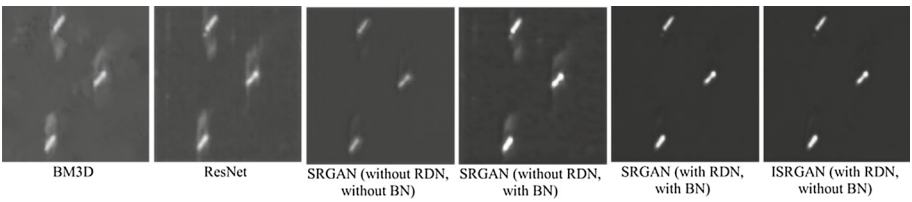


Fig. 47 The performance of the proposed model for contrast enhancement [in Ai et al. (2021, Fig. 7)]

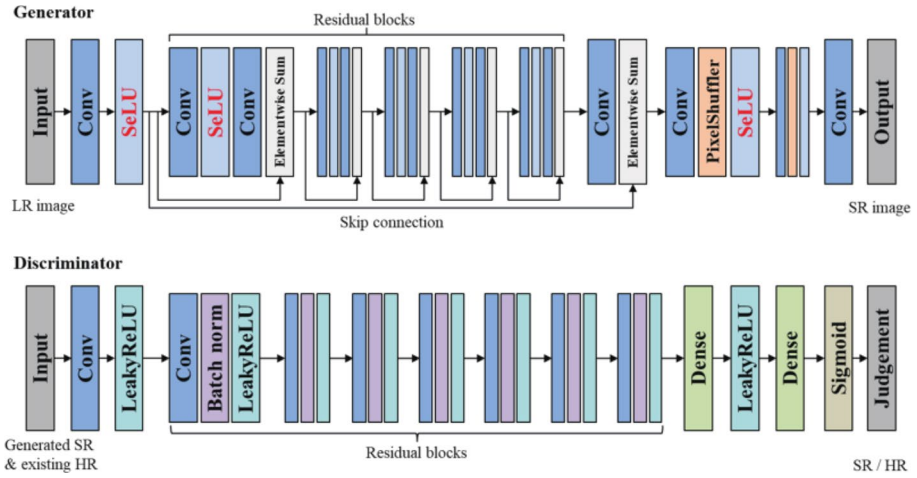


Fig. 48 The architecture of the proposed SNGAN [in Zheng et al. (2019, Fig. 1)]

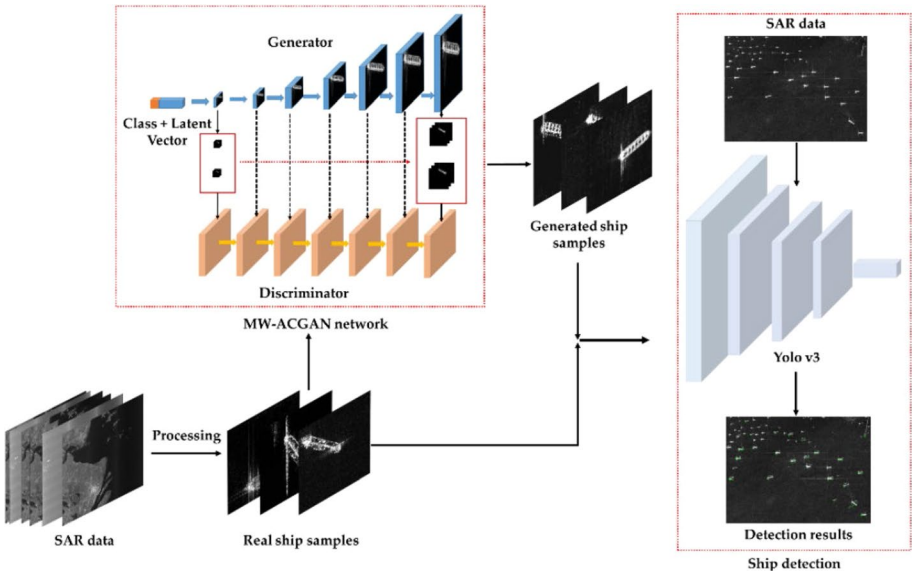


Fig. 49 The architecture of the proposed MW-ACGAN [in Zou et al. (2020, Fig. 2)]]

4.3 cGAN-based methods

In Zou et al. (2020), an improved network is designed based on WGAN-GP and ACGAN to generate the multi-scale high-resolution SAR images for the task of ship detection, namely Multiscale Wasserstein ACGAN (MW-ACGAN). Firstly, the proposed MW-ACGAN can generate the SAR images of ships steady by using Wasserstein distance. The architecture of the model is shown in Fig. 49, which consists of generator blocks and discriminator blocks. The generator block is constructed with two convolutional layers, two PixNormalization

layers, and two LeakyReLU activation layers. The discriminator block is constructed with two convolutional layers and two LeakyReLU layers. In the traditional ACGAN, JS divergence is applied to measure the distribution difference between the generated and real images. However, the application of JS divergence may lead to the vanishing gradient of the generator due to the over-training of the discriminator. Hence, Wasserstein distance is applied to replace the JS divergence and measure the difference between the real and generated SAR images. Moreover, the multi-scale information is applied to meet the requirement of multiple-resolution SAR image generation, which cannot be achieved using the traditional ACGAN. To this end, a multi-scale dataset is established first. Then, the convolution operation with the kernel size of 1×1 is applied to squeeze the channels of the feature maps, which are further changed to three channels. After that, the modified feature maps are distinguished using the corresponding discriminators, and the multi-scale Wasserstein loss and gradient penalty loss are applied for optimization. Finally, Yolov3 is applied for ship detection in the SAR dataset, which is conducted using the generated images and real images, and an accuracy over 0.94 is obtained.

In addition, the loss function and the fused features can be modified to achieve a better super-resolution performance. In Zhang et al. (2021a), a novel approach based on density-based spatial clustering merging with edge penalty (DBSCAN) is proposed for rapid superpixel generation of SAR images, which includes the process of rapid clustering and feature merging. Firstly, Gaussian filtering is applied to suppress the texture and speckle information. Then, the edge features can be extracted. After that, the traditional DBSCAN is modified for superpixel generation of the SAR images with speckle noise. Firstly, the nonedged pixel is selected as the seed pixel, and the similarity between the seed pixel and the four neighbor pixels is calculated. Then, the neighbor pixel with high similarity is set as a new seed pixel, and the procedure is iterated continuously. After the fast clustering, the superpixels with the highest similarity can be merged. Finally, the proposed method can achieve an accuracy of 90.15% by using 6000 generated superpixels. In Xiang et al. (2019), a superpixel generation approach for SAR images, which is inspired by Chen et al. (2017); Dhillon et al. (2007) is proposed. The proposed approach is based on linear feature clusters and edge constraints for superpixel generation. Firstly, the local gradient ratio pattern (LGRP) feature of each pixel, which is used to be applied for SAR automatic target recognition is extracted. Then, a feature-ratio-based detector with a Gauss-shaped window is applied to replace the traditional rectangular window. Therefore, the edge intensity map and the edge of the SAR image are extracted. After that, the weighted *K*-means clustering with edge constraint is designed to approximate the optimal cost. The proposed method can achieve the precision of 0.86, 0.84, and 0.83, and recall of 0.88, 0.86, and 0.81 for three different datasets. In Jing et al. (2021), the same authors propose a content-sensitive superpixel generation method for SAR images. The main contributions include three parts. First, a content-sensitive superpixel initialization method is proposed. The superpixel content complexity is expressed by computing entropy, and initialization is constructed based on this to keep the same distribution as the prior information to the greatest extent possible. Then, the generalized likelihood ratio (GLR) is used to compare speckles in SAR images. Finally, considering the inherent continuity information of adjacent pixels, the information is applied to enhance the connectivity of superpixels, and further effectively generate high-quality superpixels. Instead of using the traditional clustering method based on local weighted *K*-means clustering described in Xiang et al. (2019), a contraction-expansion search strategy (CES) is proposed, which explicitly uses the continuity information contained in adjacent pixels and enhances the connectivity of superpixels without any post-processing steps. The problem of losing details in the sparse area of the image described

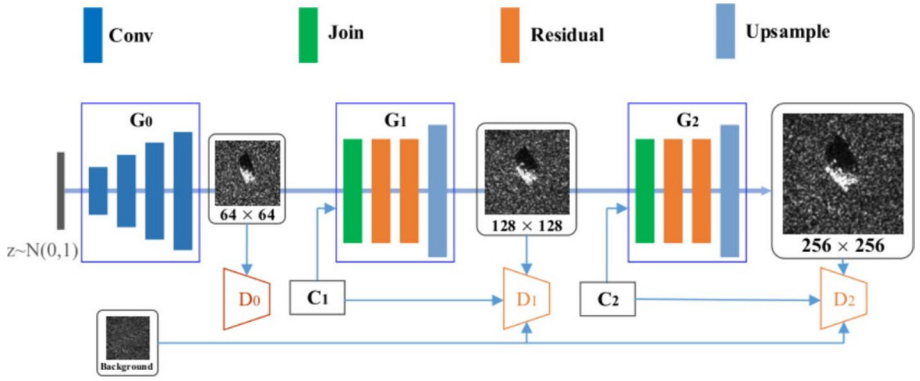


Fig. 50 The architecture of the proposed hierarchical GAN [in Huang et al. (2019, Fig. 1)]

in Xiang et al. (2019) is solved. With the help of CES, the proposed method can obtain superpixels with lower computational cost and higher edge adhesion. Finally, the proposed model achieves a boundary recall (BR) of 0.98 and an under-segmentation error (UE) of 0.11. In Huang et al. (2019), a hierarchical GAN is designed for high-resolution SAR image generation to solve the problem of small samples. The proposed generator contains three up-sampling stages, which are shown in Fig. 50. In the first stage, a set of normal distribution noise is applied for SAR image generation with the size of 64×64, and a discriminator is applied here for alternate training. In the second stage, the low-resolution SAR image generated in the first stage is combined with a condition vector, and several residual blocks are applied to learn the multi-mode representation. After that, the SAR image with the size of 128×128 is generated and transferred to the discriminator for distinguishing. The loss contains the global adversarial loss, the background loss, and the conditional loss. The generation approach in the third stage is the same as in the second one, and the SAR image with the size of 256×256 is generated. The MSTAR dataset is applied for training and validation. Finally, a pre-trained classifier is applied to evaluate the generation performance of the proposed model. The proposed method can achieve the accuracy of 94.17%, 95.83%, and 95.71% for T72, BMP2, and BTR70, respectively.

Besides, cGAN can also be applied for corrupted image restoring and dehazing. In Bermudez et al. (2019), an improved cGAN is applied to generate the corrupted optical images by using the corresponding SAR images. Firstly, the paired SAR and corresponding optical images are selected by using a sliding window. Then, the cGAN is trained using the selected image pairs. After that, the cloudy patches of the optical images can be generated by using the corresponding SAR images and the cloudy-free images at another time as the input. Moreover, the generated cloudy-free patches can be fused with the original cloudy optical images. The dataset of SEN1-2 is applied here for training. Finally, the proposed method can improve the accuracy and F1-score by 7.7% and 8.6% in the task of image classification. In Huang et al. (2020), an improved cGAN is applied to design an end-to-end dehazing network. Different from the traditional cGAN, the SAR images and cloudy RGB images are applied as conditions and input of the generator, respectively. In the training step, the convolutional operations are applied for feature extraction and an encoder-to-decoder network is applied to extract the detailed features of cloudy RGB images. And the extracted features are combined to generate haze-free images. After that, the loss function in the least square generative



Fig. 51 The comparison of different models for dehazing. **a** Input images; **b** ground truth; **c** DCP; **d** DehazeNet; **e** SAR-opt-cGAN; **f** The proposed model [in Huang et al. (2020) Fig. 4)]

adversarial network (LSGAN) is applied here both in the generator and discriminator. As for the discriminator, a 70×70 PatchGAN is applied for discriminator construction. A new dataset is constructed including cloudy RGB images, SAR images, and corresponding RGB haze-free images. The proposed approach achieves the PSNR and SSIM of 24.1638 and 0.9061 in thin fog images, 25.3111 and 0.9264 in moderate fog images, and 25.0731, and 0.8640 in thick fog images, respectively. The comparison of different models for image dehazing is shown in Fig. 51.

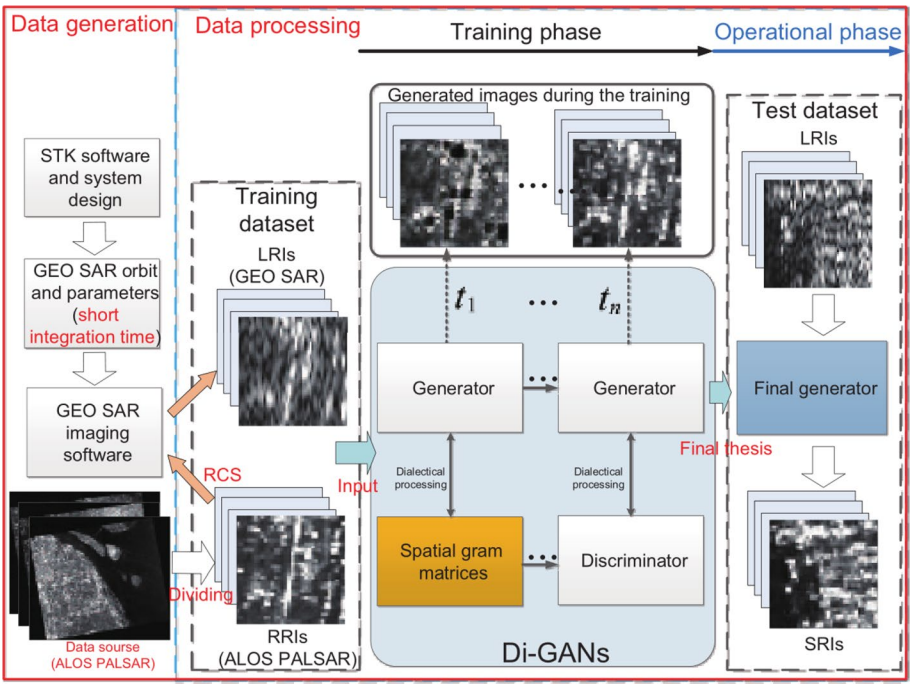


Fig. 52 The flowchart of the proposed method [in Li et al. (2019c, Fig. 1)]

4.4 Pix2PixGAN-based methods

In Li et al. (2019c), the dialectical-GAN (DI-GAN) is applied for SAR image super-resolution generation. The flowchart of the proposed approach is shown in Fig. 52, which can generate the super-resolution images using the low-resolution and high-resolution images as the input. In the generator, U-Net is applied as backbone to fuse the content and texture information of different resolutions. After that, PatchGAN is applied to construct the discriminator and the Wasserstein distance is applied as the loss function. The proposed SRGAN can achieve the MSE of 0.0524, 0.0014, and SSIM of 0.9661 and 0.9083 in the training and testing process. The result shows that the proposed model can generate super-resolution SAR images with a quality comparable to high-resolution SAR images.

In Gao et al. (2020), an improved Pix2PixGAN is applied for cloud removal by fusing the SAR and optical images. Firstly, a modified U-Net is applied to translate the real SAR image into a simulated optical image in an object-to-object manner. Though the simulated optical images cannot be considered real optical images due to the loss of texture information, they are more suitable for translation tasks by comparing with the original SAR images, and the simulated images can change the contrast and intensity randomly for better robustness. After that, a modified Pix2PixGAN is applied as the fusion network to fuse the simulated image, SAR image, and corrupted image to generate the image with detailed texture information, whose architecture is shown in Fig. 53. Moreover, the perceptual loss is added as the global loss function and the cloud mask is added as the local loss function. Finally, the proposed model can achieve the SSIM of 0.9135 and 0.9060, correlation coefficient (CC) of 0.9642 and 0.9721, SAM of 2.8158 and 3.1621, RMSE of 6.9184 and 9.7865 in Dataset A and Dataset B, respectively.

In Yanshan et al. (2022), a novel model, namely an optical-guided super-resolution network (OGSRN) is designed based on Pix2PixGAN for super-resolution SAR image generation using the co-registered high-resolution optical image. The main innovation includes three parts. First, the information of the SAR images with low resolution and optical images with high resolution can be fused. Then, an enhanced residual attention module (ERAM) is designed for spatial and channel attention extraction. Finally, a network is designed for SAR to optical image translation. The proposed OGSRN includes two stages. First, an improved U-Net ($8 \times$ up sampling) and a residual attention module are designed to enhance the presentation ability of the network, and high-resolution SAR images can be generated from low-resolution SAR. Then, a style translation network is designed to convert the generated SAR images into optical images. The difference between the generated optical image and the real optical image is calculated to optimize the whole network and finally achieve the purpose of generating high-resolution SAR images from low-resolution SAR images. The architecture of spatial and channel attention mechanism is shown in Fig. 54.

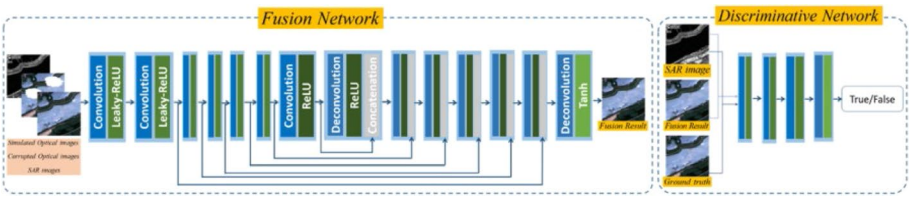


Fig. 53 The architecture of fusion network [in Gao et al. (2020, Fig. 3)]

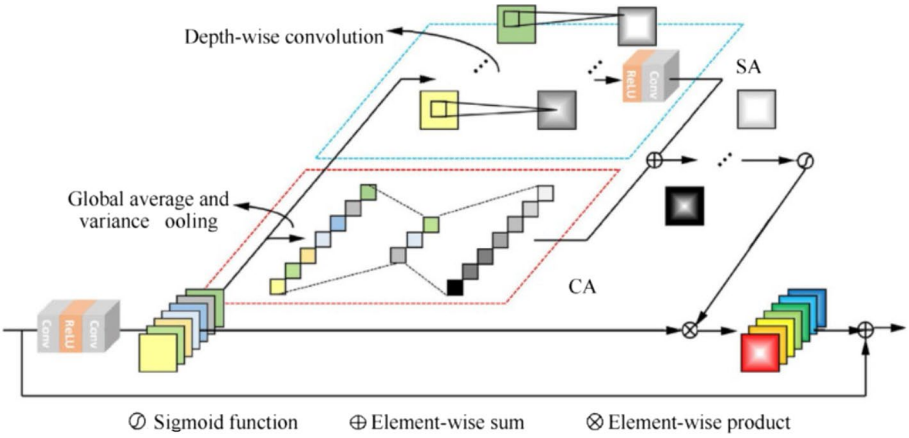


Fig. 54 The architecture of spatial and channel attention mechanism [in Yanshan et al. (2022, Fig. 3)]

The convolution layer is used to transform a single-layer SAR image into a multi-channel image, and the spatial attention and channel attention are fused using average pooling and variance pooling. L1 loss is applied here instead of L2 loss, which has better convergence performance and is helpful for super-resolution reconstruction. SAR to Optical translation model includes three parts: an encoder, a decoder, and a residual network. The encoder can be applied for downsampling the SAR image continuously, reducing the image size, and increasing the image channels. Then, the residual network contains 6 residual blocks, which can receive the feature maps produced by the encoder, and output optical images of the same size. The decoder is designed for upsampling using transposed convolutional operations. Then, PatchGAN is used to make the generated image approximate to the real image, and L1 loss is used to help translate the texture information of the SAR image into the optical image domain. The SEN1-2 dataset is applied for training and validation. The proposed method can achieve a PSNR of 45.39 and SSIM of 0.992, which performs optimally by comparing with other methods.

4.5 ViT-based methods

With the development of the architecture of the transformer, the approach based on ViT can be applied for super-resolution reconstruction. In Smith et al. (2022), an improved vision transformer (ViT) architecture is applied for SAR image super-resolution. The proposed algorithm operates on images recovered by the range migration algorithm (RMA) and produces high-fidelity images of intricate targets. Inspired by the MobileViT applied for image classification, a model namely Mobile-SRViT is proposed for image super-resolution, whose architecture is shown in Fig. 55.

Firstly, several MobileNetV2 blocks are applied to model the global and local features. Then, several MobileViT blocks are applied for feature fusion. The loss function is performed using pixel2pixel L1 metric. Finally, the proposed model achieves a PSNR of 37.61 and SSIM of 0.9909.

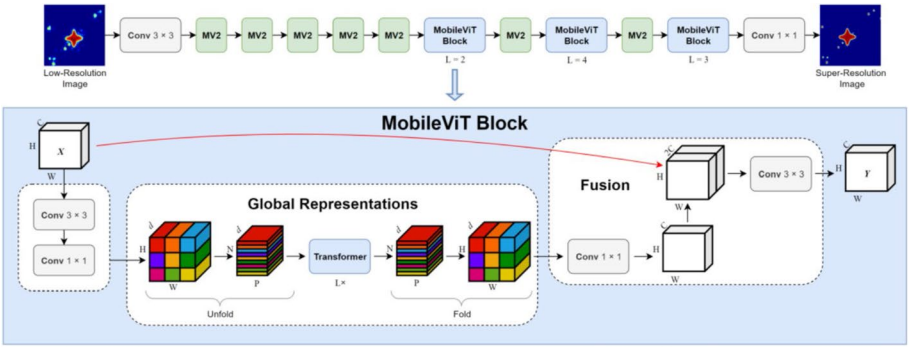


Fig. 55 The architecture of the proposed Mobile-SRViT [in Smith et al. (2022, Fig. 2)]

4.6 ISAR image enhancement

The deep neural networks are also applied for ISAR image enhancement. Different from optical images with detailed color and texture features, ISAR images focus on structural features and contrast information. In Xiao et al. (2021), a sparse feature extraction method based on dilated convolution and standard convolution is proposed to improve the reconstruction quality of the ISAR images. The proposed network consists of three parts, including the coarse feature extraction block, backbone part, and feature fusion part. Firstly, the coarse feature extraction block includes the convolutional operations and ReLU activation. Then, the dilated convolution and ResNet are combined as a sparse convolution block to extract the local and global features of the ISAR images. After that, the regularization loss is applied to replace the traditional MSE loss to reduce the sidelobe and retain the consistency between the reconstructed images and the original images. Finally, the proposed model achieves an entropy of 12.31, FID of 3.13, PSLR of − 41.92, and ISLR of − 36.36. The comparison of the results reconstructed by the different models is shown in Fig. 56. Similarly, the improvement of the loss function is also realized in Qin and Gao (2020), using GAN for super-resolution reconstruction of ISAR images. The absolute loss is combined with adversarial loss as the loss function in the generator. Besides, in Wang et al. (2022b), the L1 loss is applied to replace the MSE loss to measure the difference between the super-resolution and high-resolution images. Meanwhile, a pre-trained VGG19 is

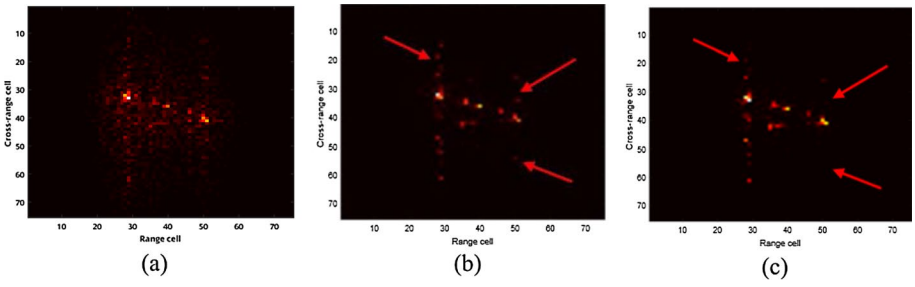


Fig. 56 The comparison of the results using different model. **a** The original image; **b** the proposed method without sparse convolution; **c** the proposed method with sparse convolution [in Xiao et al. (2021, Fig. 7)]

applied to extract the feature loss for better reconstruction. Finally, the proposed method can achieve the PSNR of 26.6683, SSIM of 0.8511, and IE of 1.6324.

Because the traditional methods cannot refocus each pixel of the ISAR image of high-speed moving targets, the real and imaginary parts of the radar image contain a lot of information. Hence, in Yuan et al. (2022a), a model based on Pix2PixGAN is applied to refocus the ISAR image. In addition, an instance normalization (IN) layer is proposed to calculate the higher-order information of defocusing, and an adaptive weighting loss function is proposed to accelerate the calculation process. First, the dual-channel image with real and imaginary parts is used in the generator to output a single-channel image. Then, IN layer is applied to replace the BN layer of the original network, which can accelerate the convergence of the network, and prevent over-fitting by calculating the covariance of the real part and the imaginary part. In addition, an adaptive weighted loss is proposed, and its architecture is shown in Fig. 57. First, the unfocused image is combined with the ground truth and the generated refocused image into an image pair, and both are input into the discriminator. Then, the adaptive loss is calculated, including the local and global loss extracted by the different numbers of convolution kernels before and after the pooling operation. Finally, the proposed model can achieve an entropy of 1.2934, NMSE of 0.4905, and SSIM of 0.7989.

In Yuan et al. (2022b), an improved attention GAN is applied for ISAR imaging. Firstly, the generator is constructed using two modified U-Net, and the Inception blocks are applied for the convolutional operation to extract the multi-scale global and local features of the ISAR images. Then, the spatial and channel attention blocks are combined to extract vital information and improve imaging quality. Finally, the proposed method can achieve the image entropy (IE) of 7.4387, PSNR of 49.5087, and RMSE of 0.8533. The result of imaging is shown in Fig. 58.

In Qin et al. (2019), the residual network is applied to generate the high-resolution ISAR images using the original low-resolution ISAR images. In the proposed model, the first convolution layer is applied to extract the features of the low-resolution ISAR image, then, 12 residual layers are applied for feature mapping. Finally, a convolution

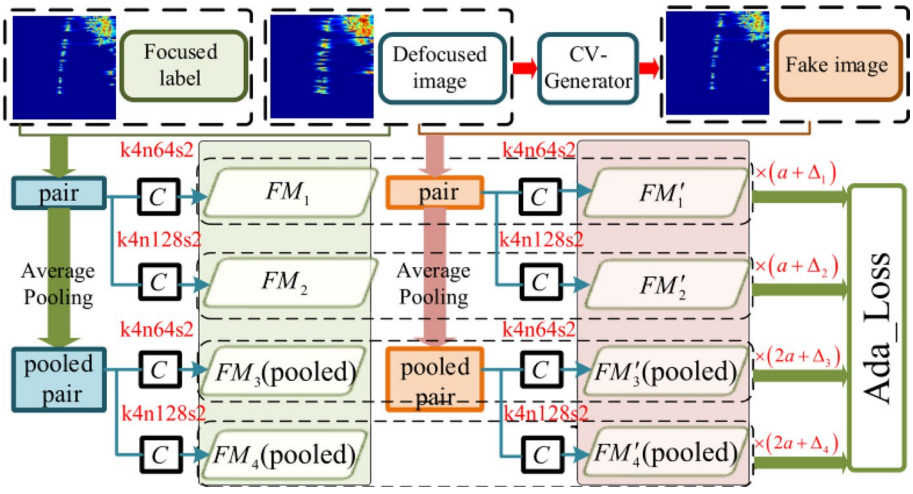


Fig. 57 The architecture of the proposed adaptive weighted loss [in Yuan et al. (2022a, Fig. 3)]

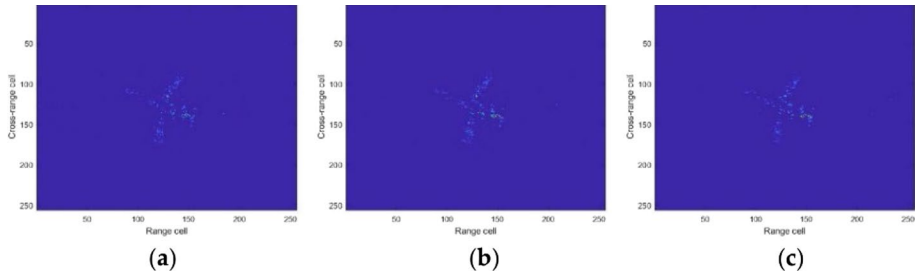


Fig. 58 The imaging result using the proposed method. **a** U-Net; **b** the proposed method; **c** ground truth [in Yuan et al. (2022b, Fig. 10)]

layer is applied to output the high-resolution images. Besides, the contrast information in the ISAR image is more important than the color and texture information, hence, the BN layers are removed to retain the abundant contrast information. In addition, the author proposes a new evaluation index *IMV* to evaluate the high-resolution generation performance of the proposed model. However, the training process using the traditional adversarial loss is easy lacks stability, and is hard to converge. Hence, in Li et al. (2022a), WGAN-GP is firstly applied for super-resolution ISAR image generation. Moreover, U-Net is applied in the generator to generate the super-resolution ISAR images and retain the strong scatting points in the low-resolution images. Besides, the SSIM-based loss is added to the loss function of the generator. After generation, the VGG16 is applied to evaluate the generation performance of the proposed model. Finally, the proposed mode can achieve a PSNR of 33.1675, SSIM of 0.9678, and classification accuracy of 81.8%.

In Li et al. (2022b), a transformer-based GAN is applied for ISAR image generation to solve the problem of the limited dataset with high quality. Firstly, the BiCubic interpolation method is applied to adjust the small ISAR images to bigger images. Then, the LeWin transformer blocks are applied to construct the generator with a structure like U-Net. After that, PatchGAN and LeWin blocks are used to construct the patch and global discriminator. The architecture of the proposed whole network is shown in Fig. 59. Moreover, the perceptual loss is added with MSE loss and adversarial loss, which can be calculated by comparing the difference between the feature mapping between the real and generated ISAR images. Finally, the proposed method can achieve the IE of 2.4356 and 2.5214, SSIM of 0.8163 and 0.8047, and TCR of 91.2041 and 89.4582 by using 50% and 25% sampling ratios, respectively.

In Jie (2020), the author proposed a comprehensive approach for multi-source ISAR image fusion and recognition. Firstly, the Pix2PixGAN is applied for image dehazing, which can be trained by using the simulated ISAR images and can achieve an SSIM higher than 0.98. Then, the CycleGAN is applied for translation from the ISAR image to the optical image. Besides, the U-Net with residual blocks is added as the generator of the CycleGAN, and PSNR, SSIM, and MSE are applied for evaluation. Moreover, a GAN with a multi-scale discriminator, Pix2PixHD, is applied for high-resolution ISAR image reconstruction, which can improve the detailed contour information of the generated images. Finally, a CNN is designed for ISAR image recognition.

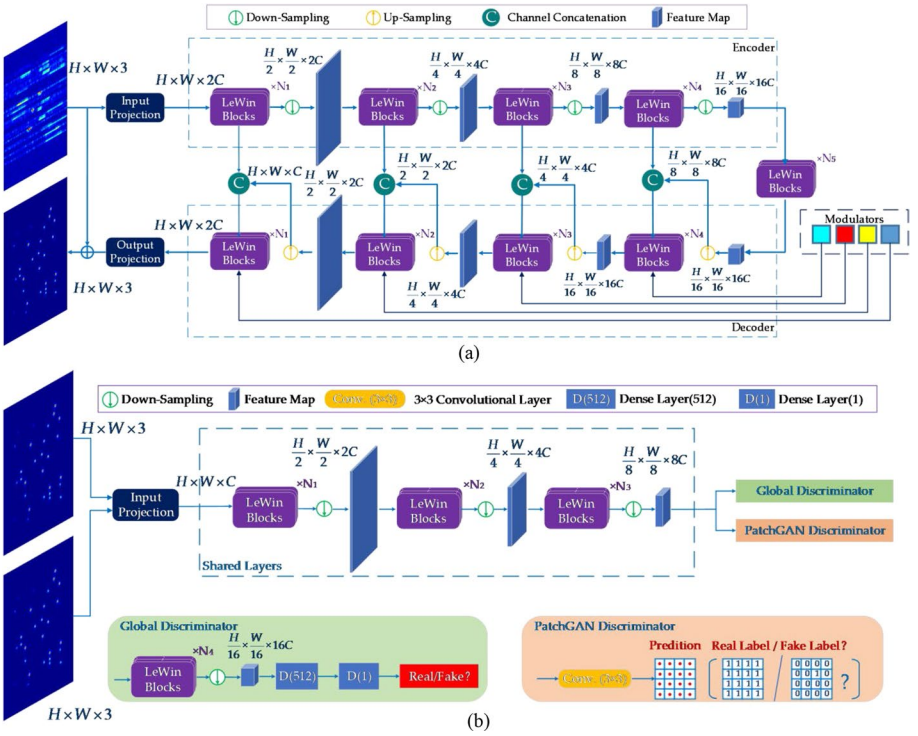


Fig. 59 The architecture of the proposed method. **a** Generator; **b** discriminator [in Li et al. (2022b, Figs. 3, 4)]

5 SAR/ISAR image generation for azimuth interpolation

The scattering characteristics of targets are different at different observation angles, with complex shapes and significant azimuthal anisotropy. Due to the difficulties of SAR/ISAR imaging, so it's urgently required the SAR/ISAR image dataset with all azimuths (Wang et al. 2022a). The traditional approaches of constructing target multi-azimuths SAR/ISAR image datasets generally consist of laboratory semi-physical simulation, computer simulation, and actual data acquisition (de Almeida et al. 2018). However, it's hard to describe the detailed structure of the target accurately using the first two methods, and it's expensive by using the last method. Hence, with the development of deep generative models, there exist several novel methods for azimuth interpolation of SAR/ISAR images, which are summarized as follows.

5.1 DCGAN-based methods

In Zhang et al. (2018), an improved DCGAN is applied to generate the SAR image with a specific azimuth. First, the sparse model is applied to linearly combine the missing azimuth images, and the dataset can be augmented by using two SAR images with known azimuth. Then, in the discriminator, the SAR image is segmented, and the binary image is obtained.

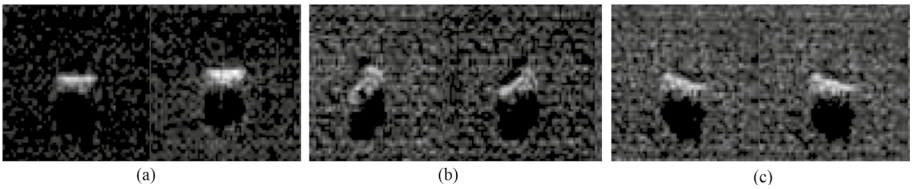


Fig. 60 The generated SAR image with specific azimuth. **a** 94.763°; **b** 148.199°; **c** 104.036° [in Zhang et al. (2018, Fig. 4)]

And the minimum bounding box can be calculated after edge detection. Finally, the azimuth of the target can be calculated in the discriminator. The result is shown in Fig. 60.

In Wang et al. (2022c), a self-attention GAN (SAGAN) is proposed based on DCGAN for the azimuth interpolation of SAR images. By comparing with the original GAN, the sine and cosine of the azimuth are introduced as condition labels and concatenated with random noise after up-sampling in the generator. In the discriminator, the image and azimuth information is concatenated to discriminate whether the image is real or fake and if the azimuth is matched or not. Moreover, the self-attention mechanism is applied to reflect the correlation between the features in the whole image for better generation. The spectral normalization is applied in the weight matrix of each layer of the generator and discriminator, which can make the discriminator meet the Lipschitz condition. At the same time, spectral normalization can also avoid gradient anomalies caused by excessive parameters of the generator to improve stability and efficiency. The dataset contains the images captured in every five degrees of two aircraft. The mean, variance, dynamic range, equivalent number of views, radiation resolution, and multi-scale structural similarity (MSSSIM) are applied for comprehensive evaluation. The result shows that the mean value, variance, dynamic range, equivalent number of views, radiation resolution, and multi-scale structure similarity of the proposed model are closest to the real image, which is better than DCGAN.

The existing SAR image generation approaches can only generate images around the specific azimuth, which lack features for target recognition. Hence, a novel GAN is proposed to generate the precise SAR image based on two given known azimuths. In Wang et al. (2022a), an improved GAN is applied to generate the SAR image with a specific azimuth. The architecture of the proposed model is shown in Fig. 61. Firstly, the generator can

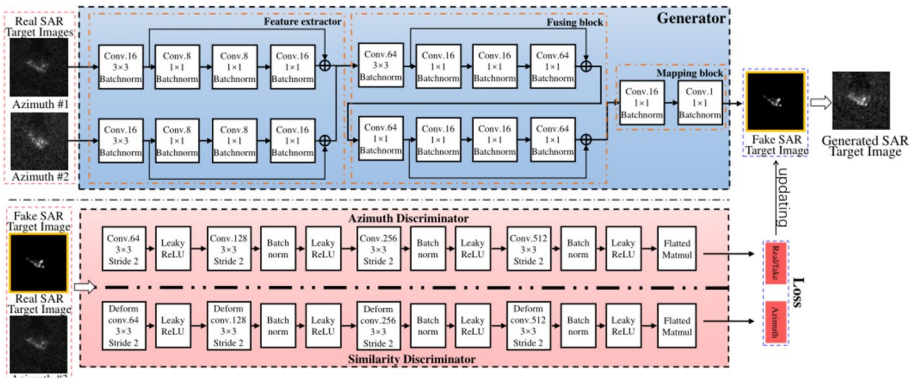


Fig. 61 The architecture of the proposed model [in Wang et al. (2022a, Fig. 2)]

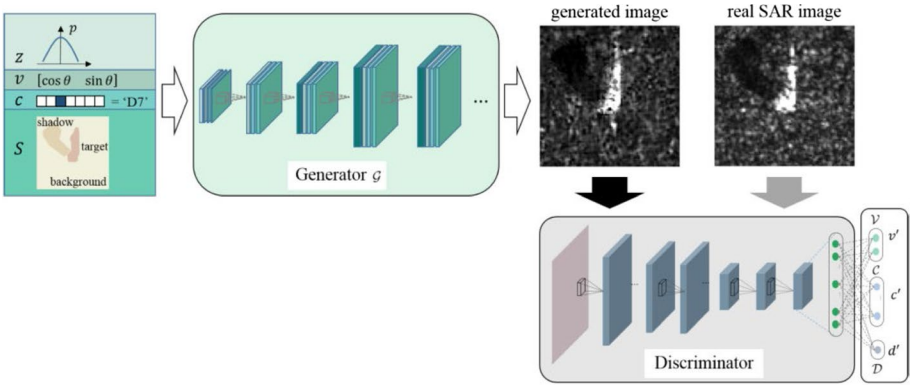


Fig. 62 The architecture of the proposed AAE [in Song et al. (2021, Fig. 3)]

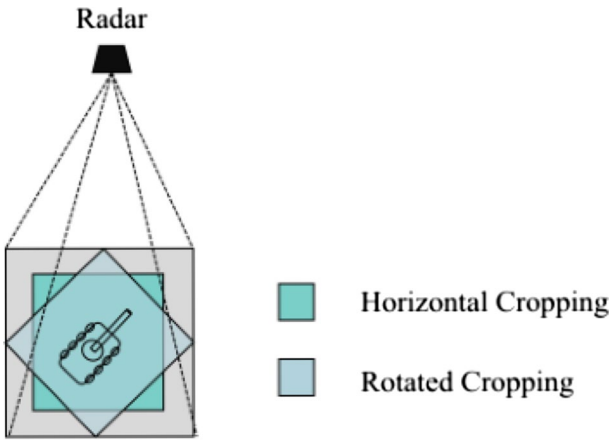


Fig. 63 The principle of proposed rotated cropping operation[in Song et al. (2021, Fig. 5)]

extract the features of the target based on two input SAR images, which can be fused by using multiple residual blocks. Then, two discriminators are designed, including a similarity discriminator and an azimuth discriminator, which can be applied to distinguish the real and generated images and measure the azimuth difference between the generated image and the target image, respectively. For the design of the loss function, MSE is applied to replace the logarithmic error for stable training, and Wasserstein distance is applied in the generator to achieve the optimal model. The dataset of MSTAR is applied here for training and testing. Finally, the recognition accuracy using the dataset generated by the proposed method achieves 98.22%, which is better than using the dataset without augmentation.

5.2 Auto-encoder-based methods

In Song et al. (2019, 2021), an adversarial autoencoder (AAE) is designed to solve the problem of the insufficient dataset in the tasks of the few shot-learning (FSL), which can be applied to generate the all-azimuth SAR image.s The architecture of the proposed AAE is shown in Fig. 62.

Similar to traditional GAN, AAE also consists of a generator and a discriminator. The input of the generator includes four parts, the first part is a set of random noise with normal distribution; the second one is the azimuth information represented by sine and cosine; the third one is the category information of the target; the last on is the semantic map of the image, including the part of the background, shadow area, and target area. Moreover, the original amplitude data of the image is applied for training and testing, instead of the image after histogram equalization. Besides, as an improved method, in Song et al. (2021), due to the mapping between the rotation matrix and rotation angle is not continuous, a rotated cropping operation is applied for the region of interest (RoI) cropping, instead of horizontal cropping, which is shown in Fig. 63. The application of rotated cropping can release the problem of over-fitting and can improve the representation capability of the proposed network. After that, the architecture is grown incrementally by referring to the progressive growing GAN (PGGAN), which can help the discriminator obtain the features in different scales. The MSTAR dataset is applied for validation of the proposed model. For the design of the loss function, the most applied KL divergence is applied for optimization, and L1 loss is added to minimize the difference between the real image and the generated image. Besides, L2 loss is applied to optimize the prediction results of azimuth and category. The proposed model achieves the normalized cross-correlation (NCC) of 0.9368, mean gradient structural similarity (MGSM) of 0.7843, and normalized gradient structural correlation (NGSC) of 0.9224.

Before that, the author proposed a deep generative model-based full-aspect SAR image generation method in their prior research (Song and Xu 2017). The zero-shot or one-shot learning can map the features from low dimension to high dimension, and the relationship between the two images can be defined based on the distance between them. Firstly, the category information is inputted into a fully-connected network, which can construct a continuous two-dimensional target space based on the orientation invariant features. Then, the category feature can be concatenated with the two-dimensional azimuth feature. After that, the combined feature can be inputted into the generator, which consists of three fully-connected layers and a deep DNN, and the transposed convolutional operation is applied for up-sampling. Besides, the L2 loss between the generated and real SAR images is introduced as the loss function. Meanwhile, an interpreter model is trained, which is constructed based on the inverse process of the generator and can output the azimuth information of the input SAR image. The result shows that the generated image has less speckle noise than the real image, which is smoother, indicating that it can be applied to speckle reduction tasks in the future. The trained interpreter can achieve an accuracy of 96.8% in the task of target recognition and the error between estimated azimuth and true azimuth is 16°. The generation result is shown in Fig. 64.

5.3 Pix2PixGAN-based methods

In Wang et al. (2019a), Pix2PixGAN is applied to generate the missing azimuth in the SAR image dataset. Firstly, the real targets are modeled to obtain the simulated SAR images with all azimuths. Then, the simulated SAR images are applied as the training dataset for training a Pix2PixGAN, where the images are considered as the input and the azimuths are regarded as the conditional information to learn the mapping between the simulated images and the real SAR images. The generator is constructed using U-Net and the discriminator is constructed using PatchGAN. The MSTAR dataset is applied in this experiment. Finally, the proposed model can achieve the IS of 1.100797 and a correlation coefficient of 89.27%.

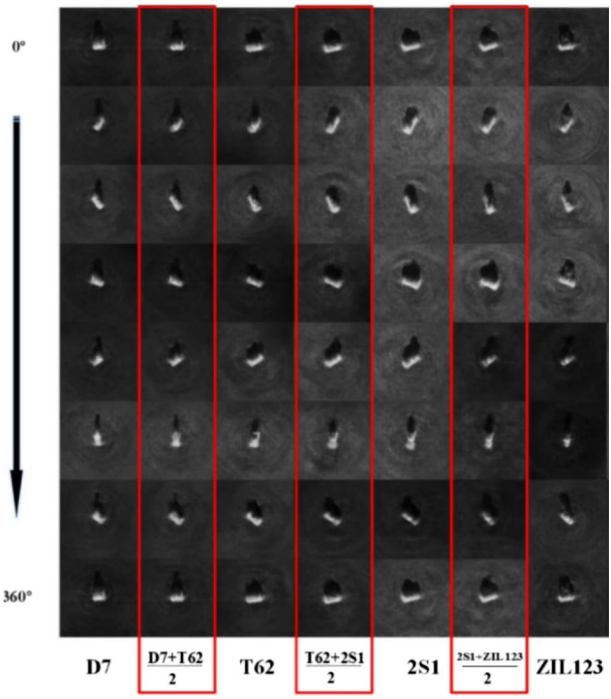


Fig. 64 The generation results [in Song and Xu (2017, Fig. 7)

5.4 InfoGAN-based methods

In Liang (2021), an improved InfoGAN is applied for multi-view SAR image generation. The cross entropy is applied as the loss function of the discriminator in the original InfoGAN, but for the generated image that is discriminated as real, even though it still has a large difference from the real image, it can no longer be further optimized. Therefore, the loss function of InfoGAN needs to be improved for better generation performance. The least square loss is applied to replace the cross entropy loss, and the network can be further optimized by punishing the image that is discriminated to be real but has low similarity with the real image, which is can help to improve the quality of generated images. At the same time, Jensen–Shannon (JS) divergence is replaced by Pearson's chi-square divergence, and LeakyReLU is applied as the activation function to replace the sigmoid of the last layer. The second part is the feature decoupling process for azimuth. A random noise is fixed by a trained generator, and ten groups of random noise are copied. Then, the noise from −1 to 1 with the step size of 0.2 is sampled using potential encoding interpolation. The results show that the azimuth values of the generated images are different with different noises. At the same time, AEGAN is applied to realize the transformation of data distribution in different domains. The images with a pitch angle of 17 in MSTAR are applied as a training dataset and the images with a pitch angle of 15 are applied as a test dataset. For a comprehensive evaluation, IS (9.137) and FID (75.23) is applied to evaluate the quality of generated images, and mean value (189.28), variance (409.21), equivalent number of views (87.55) and radiation resolution (0.441) are applied to measure the similarity of real images and generated images,

showing the best generation performance by comparing with DCGAN and original InfoGAN. However, the proposed approach still cannot generate images with specific azimuths. Moreover, AEGAN can perform distribution conversion on SAR images at different pitch angles, and can effectively improve the recognition accuracy of the model for SAR images at different pitch angles.

5.5 Probabilistic graphical model-based methods

In Zhai et al. (2019), a GAN based on a Bayesian network is proposed to generate the adjacent azimuth SAR images according to the existing azimuth SAR image for the same target. A continuous structure generation model is designed for SAR image generation. In a Bayesian network, continuous azimuth SAR images are set as dependent variables. According to the correlation between continuous variables in the Bayesian network, a SAR image of an adjacent azimuth can be calculated given a SAR image with a certain azimuth. A posteriori probability is approximated using neural networks. The dataset consists of images with azimuth from 0 to 359 of three targets. Firstly, the images of eight adjacent azimuths in the training set are combined. Then the combined features of the first four images and the respective features of each image are extracted. After that, a random variable is generated to represent the combined features of the four images, and an arbitrary variable is generated to represent the feature of the first image. Then, the features of the other three images are generated according to the relationship between features. Finally, the loss between the last four real images and the generated image is calculated. At the same time, a small number of images can be used to fine-tune the trained model to generate more images. The generation result is shown in Fig. 65.

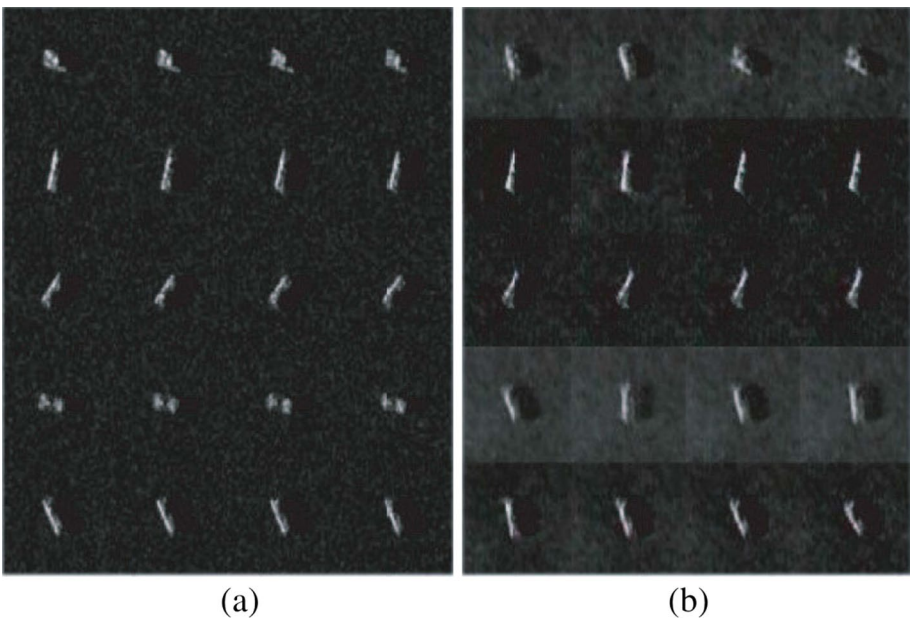


Fig. 65 The generation results of the proposed model. **a** Original images with four consecutive angles; **b** generated images with four adjacent angles [in Zhai et al. (2019, Fig. 4)]

6 SAR image generation for deceptive jamming

In Fan et al. (2020), an improved cGAN is applied for SAR deceptive jamming with high fidelity, namely deceptive jamming template generative adversarial network (DJTGAN). The loss of WGAN-GP is applied here to construct a new loss function. The condition vector of the generator contains four kinds of condition information: azimuth angle, elevation angle, target type and image resolution. At the same time, PatchGAN is introduced into the discriminator to output the probability that the discrimination result is real or false for each patch, and the final discrimination result is calculated by fusing the result of each patch. The framework of the proposed approach is shown in Fig. 66.

And MSTAR dataset is applied in the experiment. The generation result is shown in Fig. 67. Furthermore, linear index of fuzziness (LIF), average gradient (AG), mean square deviation (MSD) and gray level difference (GLD) are applied to quantitatively evaluate the edge details and contrast enhancement of the generated image. The result shows that the proposed DJTGAN can generate the SAR image as real as possible, which is proper to be applied as the deceptive jamming template.

In order to improve the convenience and readability comparing different works, the related works with their different tasks, datasets, models and evaluation indices are summarized in Table 1.

7 Analysis of deep learning based SAR/ISAR image generation

The application of deep generative models for SAR/ISAR image generation are summarized from Sects. 2 to 6. It can be seen that different models suit for different tasks. In order to discover the reasons why they can be applied in different fields and find out their unique

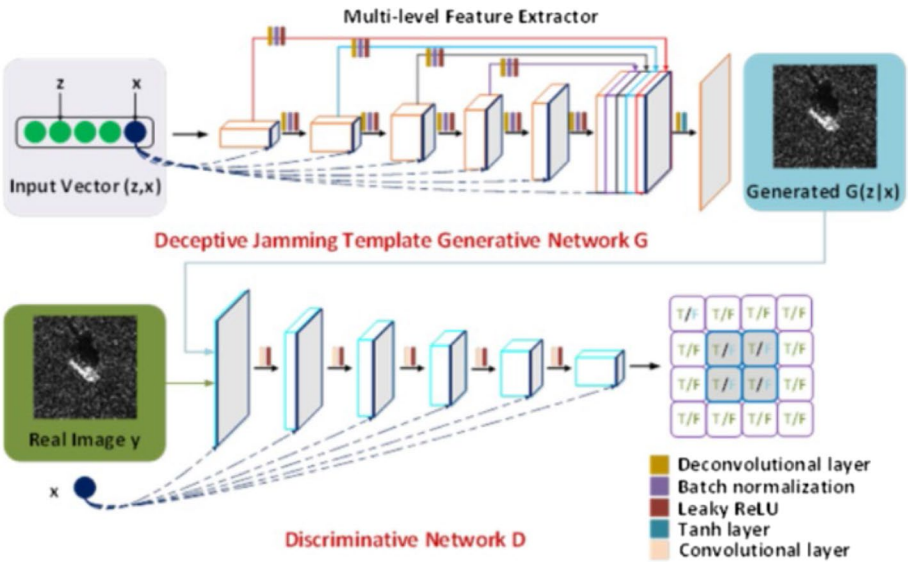


Fig. 66 The architecture of the proposed DJTGAN [in Fan et al. (2020, Fig. 2)]

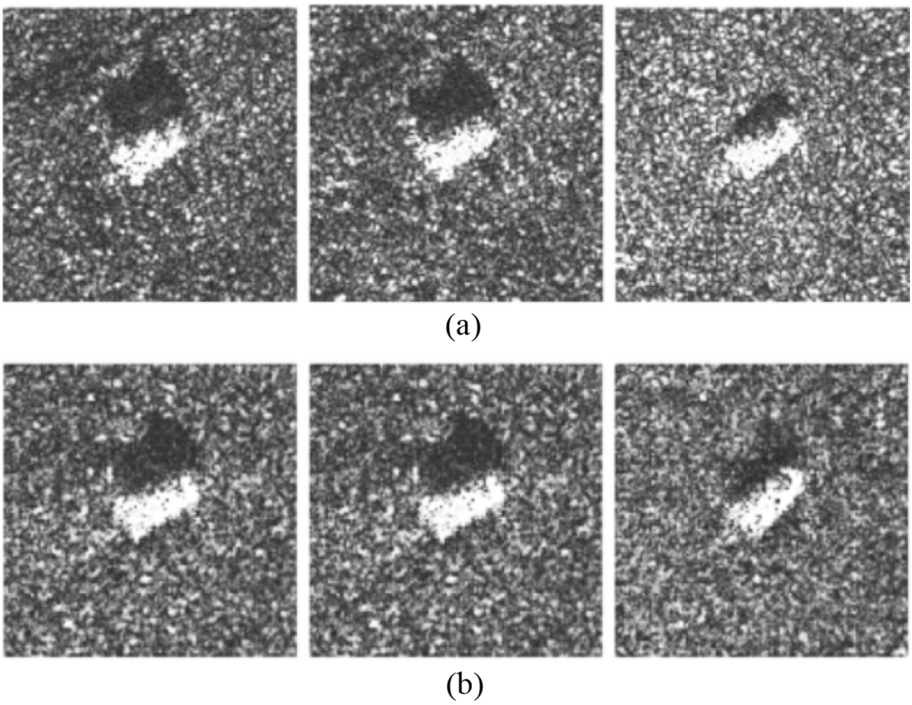


Fig. 67 The example of generated results. **a** Real images; **b** generated images [in Fan et al. (2020, Fig. 5)]

characteristics, the most used models in this review are analyzed. Furthermore, the potential approaches in the field of SAR image generation are proposed in this section.

7.1 Analysis of the task, network and dataset

The application frequency of different approaches for different tasks is shown in Fig. 68.

It can be found that the main tasks in the field of SAR/ISAR image generation are data augmentation, image translation, and image enhancement. In the contrast, the works about azimuth interpolation and deceptive jamming are relatively few. There exist many difficulties in the process of SAR/ISAR imaging, hence the SAR/ISAR images are hard to acquire by comparing with the optical images. The limited dataset may lead to inadequate training of the deep neural networks. Therefore, by comparing with the traditional data augmentation approaches, such as flipping and cropping, the application of the deep generative models can improve the quality and diversity of the generation samples. Furthermore, in the task of data augmentation, the auto-encoder, DCGAN, cGAN, and WGAN are the most used methods. By comparing with the vanilla GAN using the full-connected layers to construct the network, the convolutional layers are applied to construct the backbone of the improved DCGANs. Besides, WGAN uses the Wasserstein distance to improve the convergence performance and the similarity between the real and generated images. And cGAN is applied to generate the SAR/ISAR image of a specific category by introducing the condition information to the network. In the contrast, the works of CycleGAN, Pix2PixGAN, and InfoGAN are relatively few, and are not quite suitable for mere data augmentation.

Table 1 Summary of deep generative networks for SAR image processing

Related work	Task	Image type	Dataset	Model	Evaluation
Gao et al. (2018)	Data augmentation	SAR	MSTAR	DCGAN	Recognition Accuracy
Zhou (2020)	Data augmentation	SAR	MSATR and SAR-I	DCGAN	Recognition Accuracy and FID
Bhamidipati et al. (2020)	Data Augmentation	SAR	MSATR and DFC2020	DCGAN	mIoU, FID and pixel accuracy
Yan et al. (2022)	Data augmentation	SAR	Self Dataset	DCGAN	ENL, SSIM, FSIM, CSIM and recognition accuracy
Hwang and Shin (2021)	Data augmentation	SAR	MSTAR	DCGAN	Recognition Accuracy
Li et al. (2019b)	Data augmentation	SAR	Self Dataset	DCGAN	Detection Accuracy
Qin et al. (2022)	Data augmentation	SAR	MSTAR	cGAN	Mean Value, Variance and FID
Liu et al. (2021a)	Data augmentation	SAR	MSTAR	cGAN	Recognition Accuracy
Kong and Zhang (2021)	Data augmentation	SAR	MSTAR	cGAN	Recognition accuracy
Li et al. (2020a)	Data augmentation	SAR	GF-3	cGAN	SSIM
Zhou et al. (2021)	Data augmentation	ISAR	Self dataset	cGAN	SSIM
Xie et al. (2021)	Data augmentation	SAR	MSTAR	InfoGAN	Recognition Accuracy
Bao et al. (2019)	Data augmentation	SAR	MSTAR	WGAN	None
Lu et al. (2019b)	Data augmentation	SAR	MSTAR	WGAN	Recognition accuracy
Lu et al. (2019a)	Data augmentation	SAR	MSTAR	WGAN	None
Cui et al. (2019)	Data augmentation	SAR	MSTAR	WGAN	Mean value, variance, ENL, RR
He et al. (2019)	Data augmentation	SAR	Self dataset	WGAN	Recognition accuracy
Du et al. (2021d)	Data augmentation	SAR	MSTAR	WGAN	Recognition Accuracy
Du et al. (2021c)	Data augmentation	SAR	MSTAR	WGAN	SSIM
Luo et al. (2020)	Data augmentation	SAR	MSTAR	WGAN	Recognition Accuracy
Zhang (2021)	Data augmentation	SAR	Self Dataset	CycleGAN	Recognition accuracy
Zhao et al. (2020)	Data augmentation	SAR	DOTA and NWPU VHR-10	CycleGAN	Mean, variance, Std, ENL and RR
Long et al. (2019)	Data augmentation	SAR	Self Dataset	Pix2PixGAN	Detection Accuracy
Wang et al. (2018a)	Data augmentation	SAR	MSTAR	Auto-Encoder	FID, WAE and recognition accuracy
Huang et al. (2022)	Data augmentation	SAR	MSTAR	Auto-Encoder	Recognition Accuracy

Table 1 (continued)

Related work	Task	Image type	Dataset	Model	Evaluation
Du et al. (2021a)	Data augmentation	ISAR	Self dataset	Auto-encoder	SNR
Niu et al. (2018)	Image translation	SAR	Self dataset	cGAN	SSIM and MSE
Enomoto et al. (2018)	Image translation	SAR	Self dataset	cGAN	SSIM and PSNR
Li et al. (2020b)	Image translation	SAR	SEN1-2	cGAN	SSIM and PSNR
Doi et al. (2020)	Image translation	SAR	SEN1-2	cGAN	PMSE, SSIM and PSNR
Yang et al. (2022b)	Image translation	SAR	SEN1-2	cGAN	MSE, NIQE, SSIM, PSNR and recognition accuracy
Zhang et al. (2020)	Image translation	SAR	SEN1-2	WGAN	SSIM, FSIM, PSNR and MSE
Fu et al. (2021)	Image translation	SAR	Self dataset	WGAN	SSIM, PSNR and FID
Guo et al. (2021)	Image translation	SAR	SEN1-2	WGAN	SSIM, PSNR and MSE
Shao et al. (2022)	Image translation	SAR	SEN1-2	WGAN	PSNR and MSE
Yu et al. (2022)	Image translation	SAR	SEN1-2	WGAN	PSNR, SSIM, IS, FID and SAM
Zhou (2020)	Image translation	SAR	MSATR and SAR-I	WGAN	Recognition Accuracy
Lange (2019)	Image translation	SAR	SEN1-2	Pix2PixGAN	None
Dietrich-Sussner et al. (2021)	Image translation	SAR	SEN1-2	Pix2PixGAN	SSIM
Toriya et al. (2019)	Image translation	SAR	SEN1-2	Pix2PixGAN	None
Zuo and Li (2021)	Image translation	SAR	WHU-SEN-City	Pix2PixGAN	PSNR, SSIM and FSIMc
Zhang et al. (2021b)	Image translation	SAR	Self Dataset	Pix2PixGAN	The ratio of the generated area
Ji et al. (2020)	Image translation	SAR	SEN1-2	Pix2PixGAN	PSNR, SSIM, COSIN and FID
Sun et al. (2022)	Image translation	SAR	Self dataset	Pix2PixGAN	PSNR, SSIM, category accuracy and orientation accuracy
Wei et al. (2022)	Image translation	SAR	SEN1-2	Pix2PixGAN	PSNR, SSIM and MSE
Shi et al. (2022)	Image translation	SAR	SEN1-2	Pix2PixGAN	PSNR and SSIM
Li et al. (2019a)	Image translation	ISAR	Self dataset	Pix2PixGAN	Recognition Accuracy
Du et al. (2021b)	Image translation	ISAR	Self dataset	Pix2PixGAN	mIoU
Yuan et al. (2022a)	Image translation	ISAR	Self dataset	Pix2PixGAN	IE, NMSE and SSIM
Jie (2020)	Image translation	ISAR	Self dataset	Pix2PixGAN	PSNR, MSE and SSIM

Table 1 (continued)

Related work	Task	Image type	Dataset	Model	Evaluation
Fuentes Reyes et al. (2019)	Image translation	SAR	SEN1-2	CycleGAN	IoU, precision and recall
Luo and Pi (2022)	Image translation	SAR	SEN1-2	CycleGAN	SSIM, PSNR and FID
Hwang et al. (2020)	Image translation	SAR	WHU-SEN-City	CycleGAN	SSIM and PSNR
Wang et al. (2019b)	Image translation	SAR	SEN1-2	CycleGAN	SSIM, PSNR and FSIM
Yang et al. (2022a)	Image translation	SAR	SEN1-2 and WHU-SEN-City	CycleGAN	SNR, SSIM, NIQ, FID and LPIPS
Wang et al. (2018b)	Image enhancement	SAR	Self dataset	SRGAN	MSE and SSIM
Ai et al. (2021)	Image enhancement	SAR	HRSID	SRGAN	ATTR
Gu et al. (2019)	Image enhancement	SAR	Self Dataset	DCGAN	PSNR and SSIM
Zheng et al. (2019)	Image enhancement	SAR	TerraSAR and MSTAR	DCGAN	PSNR and SSIM
Yuan et al. (2022b)	Image enhancement	ISAR	Self Dataset	DCGAN	IE, PSNR and RMSE
Xiao et al. (2021)	Image enhancement	ISAR	Self Dataset	DCGAN	IE, FID, PSLR and ISLR
Qin and Gao (2020)	Image enhancement	ISAR	Self Dataset	DCGAN	None
Wang et al. (2022b)	Image enhancement	ISAR	Self Dataset	DCGAN	IE, PSNR and SSIM
Zou et al. (2020)	Image enhancement	SAR	Self Dataset	cGAN	Detection Accuracy
Zhang et al. (2021a)	Image enhancement	SAR	Self Dataset	cGAN	Recognition Accuracy
Xiang et al. (2019)	Image enhancement	SAR	Self Dataset	cGAN	Precision and Recall
Jing et al. (2021)	Image enhancement	SAR	Self Dataset	cGAN	Boundary recall and under-segmentation error
Huang et al. (2019)	Image enhancement	SAR	MSTAR	cGAN	Recognition Accuracy
Bermudez et al. (2019)	Image enhancement	SAR	SEN1-2	cGAN	Recognition Accuracy and F1-score
Huang et al. (2020)	Image enhancement	SAR	Self Dataset	cGAN	PSNR and SSIM
Li et al. (2019c)	Image enhancement	SAR	Self Dataset	Pix2PixGAN	MSE and SSIM
Gao et al. (2020)	Image enhancement	SAR	Self Dataset	Pix2PixGAN	SSIM, CC, SAM and RMSE
Yanshan et al. (2022)	Image enhancement	SAR	SEN1-2	Pix2PixGAN	PSNR and SSIM
Smith et al. (2022)	Image enhancement	SAR	Self Dataset	ViT	PSNR and SSIM
Li et al. (2022b)	Image enhancement	ISAR	Self Dataset	ViT	IE, SSIM and TCR

Table 1 (continued)

Related work	Task	Image type	Dataset	Model	Evaluation
Li et al. (2022a)	Image enhancement	ISAR	Self Dataset	WGAN	PSNR, SSIM and recognition accuracy
Qin et al. (2019)	Image Enhancement	ISAR	Self Dataset	ResNet	IMV (Self Designed)
Zhang et al. (2018)	Azimuth Interpolation	SAR	MSTAR	DCGAN	None
Wang et al. (2022c)	Azimuth Interpolation	SAR	Self Dataset	DCGAN	SSIM
Wang et al. (2022a)	Azimuth Interpolation	SAR	MSTAR	DCGAN	Recognition Accuracy
Song et al. (2019)	Azimuth Interpolation	SAR	MSTAR	Auto-Encoder	Recognition Accuracy
Song et al. (2021)	Azimuth Interpolation	SAR	MSTAR	Auto-Encoder	NCC, MGSM and NGSC
Song and Xu (2017)	Azimuth Interpolation	SAR	MSTAR	DNN	Recognition Accuracy
Wang et al. (2019a)	Azimuth Interpolation	SAR	MSTAR	Pix2PixGAN	IS and CC
Liang (2021)	Azimuth Interpolation	SAR	MSTAR	InfoGAN	IS, FID, Mean value, Variance, ENV and RR
Zhai et al. (2019)	Azimuth Interpolation	SAR	MSTAR	Probabilistic Graphical Model	None
Fan et al. (2020)	Deceptive Jamming	SAR	MSTAR	cGAN	LIF, AG, MSD and GLD

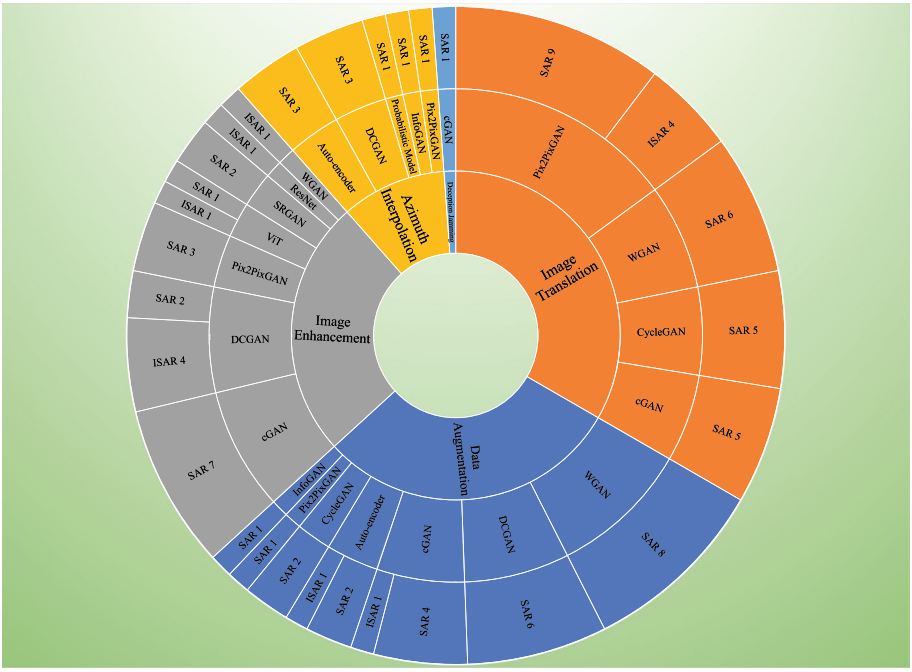


Fig. 68 The application frequency of the different approaches in the reviewed paper

Besides, in the task of image translation, most works are about SAR/ISAR to optical image translation, but few for optical to SAR/ISAR image translation, because of the difficulties of SAR/ISAR image acquisition. In this field, image-to-image translation models, such as CycleGAN and Pix2PixGAN are widely applied. Pix2PixGAN requires paired SAR/ISAR and optical images, while CycleGAN requires the images in different styles, which is not strict by comparing with Pix2PixGAN. The SAR/ISAR images have abundant structural information, while the optical images have detailed color and texture information, the translation between them can improve the quality and quantity of the dataset. Besides, SAR/ISAR has the capability of all-time and all-weather imaging, so the translation methods can fix the image corrupted by bad weather.

The task of image enhancement consists of contrast enhancement, super-resolution reconstruction, and dehazing of SAR/ISAR images, among which, the works about super-resolution reconstruction from low-resolution SAR/ISAR images are relatively abundant. To this end, many researchers modified the models, and the input of the generator is no longer a set of random noise, but a low-resolution SAR/ISAR image. Besides, the deep features can be extracted from the real high-resolution SAR/ISAR images by using deep models like the residual blocks, and the mapping from low-resolution to high-resolution can be learned, which can be applied to reconstruct the super-resolution images. In addition, the novel ViT-based models are also being applied in this field, which may have a better performance by using the self-attention mechanism by comparing with the traditional convolution-based models in the future.

Moreover, due to the difficulties of capturing the SAR/ISAR image of a target with all azimuth, the deep generative models can also be applied to the task of azimuth interpolation. And the electromagnetic scattering feature of the target in different azimuths is also

variant, hence, the traditional interpolation methods such as image rotation cannot meet the requirement of high accuracy in this field. Till now, there exist several excellent works for azimuth interpolation, such as the works in Song et al. (2019, 2021). The azimuth information can be added to the input of the generator, so the mapping between the images with the original azimuth and target azimuth can be learned. Besides, the probabilistic graphical model based on a Bayesian network can be applied for azimuth interpolation in Zhai et al. (2019). The adjacent azimuth image can be generated based on the correlation between continuous variables in the Bayesian network. As for the deceptive jamming, the related works are relatively few, showing that it will have an enormous application potential using the more advanced generative models in the future.

After that, we find that there are 2 papers for ISAR data augmentation, 4 papers for ISAR image translation, 7 papers for ISAR image enhancement, and no papers for azimuth interpolation and deceptive jamming of ISAR image. By comparing with SAR, the related works of ISAR image generation based on the deep generative models are relatively few. One of the vital reasons is the difficulties in capturing the ISAR images. The application frequency of the dataset for different generation tasks in this review is shown in Fig. 69.

Figure 69 shows that MSTAR is the most used dataset in the task of data augmentation, azimuth interpolation, and deceptive jamming, which includes multi-categories target images in different azimuth and angle of pitch. The MSTAR dataset is already widely applied in the tasks of ATR and SAR image classification. In addition, the MSTAR dataset is also suitable for azimuth interpolation because the given azimuth information can be applied as the condition information of the input of the generator. Besides, SEN1-2 is the most used dataset for image translation, including 282,384 SAR-to-optical image pairs. The SAR and optical images are precisely matching, which is suitable for the translation and coloring tasks. After that, in the task of image enhancement, the self-dataset is the most applied. Because both the MSTAR and SEN1-2 dataset contains no low-resolution

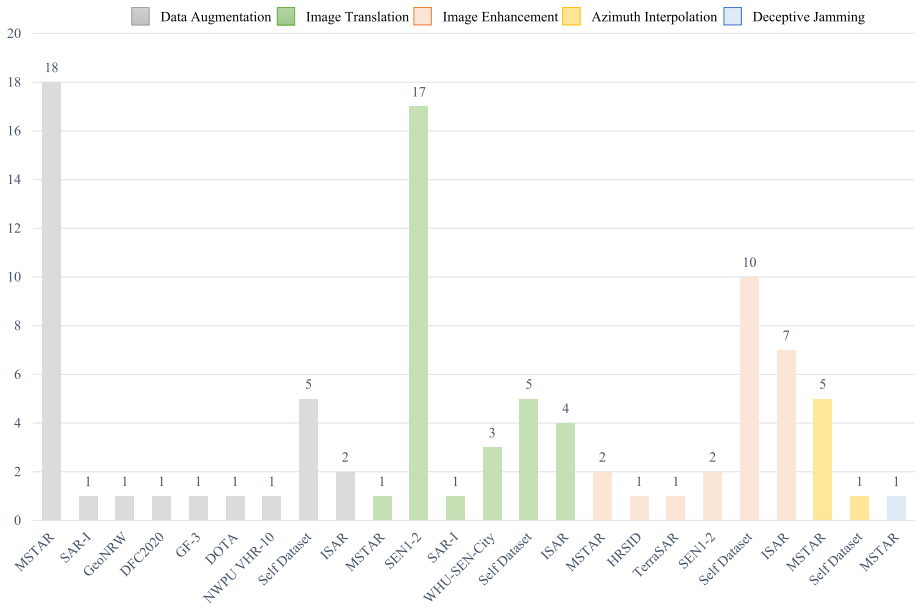


Fig. 69 The application frequency of datasets in the reviewed paper

and high-resolution SAR image pairs, most of the deep generative models are supervised models. Hence, the distance between the real high-resolution and generated super-resolution images is hard to calculate. So the self-made dataset is applied most in this field. Besides, the public dataset can also be applied after pre-processing, for example, the image after interpolation can be regarded as the high-resolution sample and the original image can be regarded as the low-resolution sample. It's worth noting that all works of ISAR image generation use their self-datasets, including the actual imaging and simulated data.

7.2 Analysis of the evaluation indices

The application frequency of the evaluation indices is summarized in Fig. 70. By reviewing the evaluation indices applied in the related works, we find that the recognition accuracy is the most used single evaluation index, which accounts for about 28.74%. It's easy and intuitive to train a classifier, which can then be applied to determine whether the proposed model can generate the image with a specific category or whether the generated images are similar enough to the real images. However, the architecture of the classifiers is variant in different works, so it's difficult to compare their recognition results. After that, the SSIM is always applied with PSNR and MSE, which account for about 32.19% and 12.65%, respectively. Firstly, the PSNR can be calculated by using MSE, which can be applied to measure the similarity between the generated and real SAR/ISAR images. Then, due to the value of PSNR is not intuitive to human feelings, SSIM can be introduced to measure the distance between two distributions together, which is more in line with the intuitive feelings of the human eye. It's worth noting that the category 'None' means no evaluation index is applied in related work, perhaps the reason is that the paper is relatively early and there are few evaluation indices that can be applied at that time. Another reason might be the generation results is hard to evaluate, such as the works about azimuth interpolation.

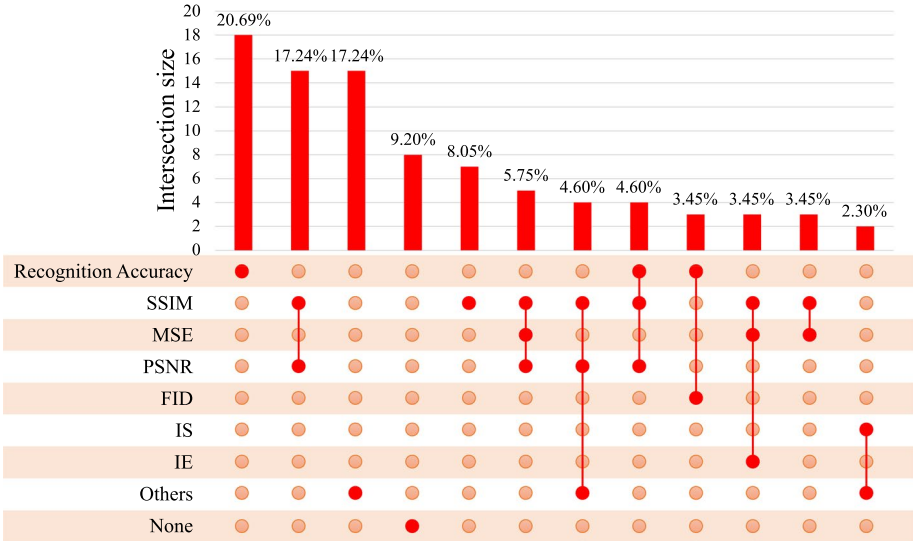


Fig. 70 The application frequency of the evaluation indices in the reviewed paper

Besides, there exist some evaluation indices that are relatively few applying, such as linear index of fuzziness (LIF), average gradient (AG), mean square deviation (MSD), and gray level difference (GLD) (Fan et al. 2020). And some evaluation indices are from other tasks, including image segmentation, image classification, and object detection, such as under-segmentation error (UE, F1-score, precision, and recall. The evaluation indices in the two cases above are summarized in category 'Others' for better readability.

7.3 Analysis of potential methods

By reviewing the related papers summarized in this survey, it can be found that there are two main generative models that can be applied for SAR/ISAR image generation tasks, that are GANs and auto-encoders. The traditional machine learning models can be classified as discriminant models and generative models. The discriminant models can be trained to learn decision functions directly from the data, such as decision trees and support vector machines (SVM). Besides, the generative models can be trained to learn the distribution of the data and model the data, such as Bayes-based models. VAE and GANs are kinds of generative models, which can learn the latent features of the data and have been greatly developed in recent years. VAE is to find a probability density in an explicit way and can achieve the optimal by maximizing the evidence lower bound (ELBO). And GAN is to find a balance in an adversarial way and needs no explicit probability density, which has a latent space to find the meaningful vector and achieve the attribute generation and domain translation. However, adversarial training can always cause the problem of mode collapse.

In recent years, with the development of deep generative models, diffusion model is designed and applied in many fields, such as computer vision (CV) and natural language processing (NLP). The architecture of the diffusion model is shown in Fig. 71.

Diffusion model is first proposed in Sohl-Dickstein et al. (2015), and widely applied after Ho et al. (2020). The diffusion model consists of a forward diffusion process and a reverse diffusion process. Briefly, the forward diffusion process is a Markov chain, which can add the Gaussian noise continuously to the original image till the distribution of data tends to a standard Gaussian distribution. While the reverse diffusion process learns to restore the distribution of the original data. The diffusion model can solve the unstable training problem of GAN, which can add the Gaussian noise to a real image systematically, while the GAN is applied to generate the image from a set of random noise and the coincidence between the generated and real distributions are relatively low.

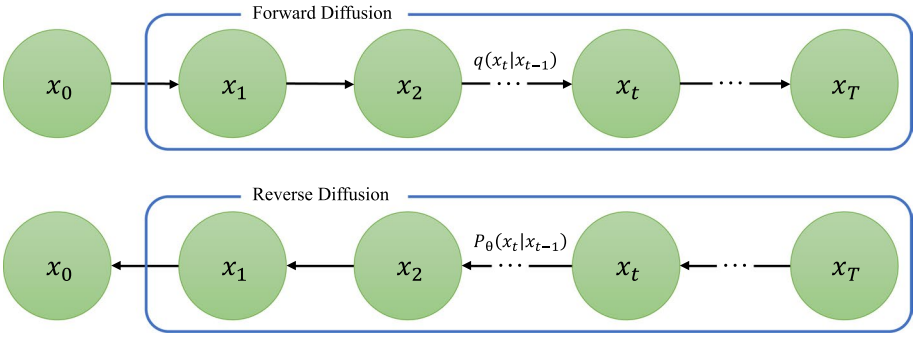


Fig. 71 The architecture of the diffusion model

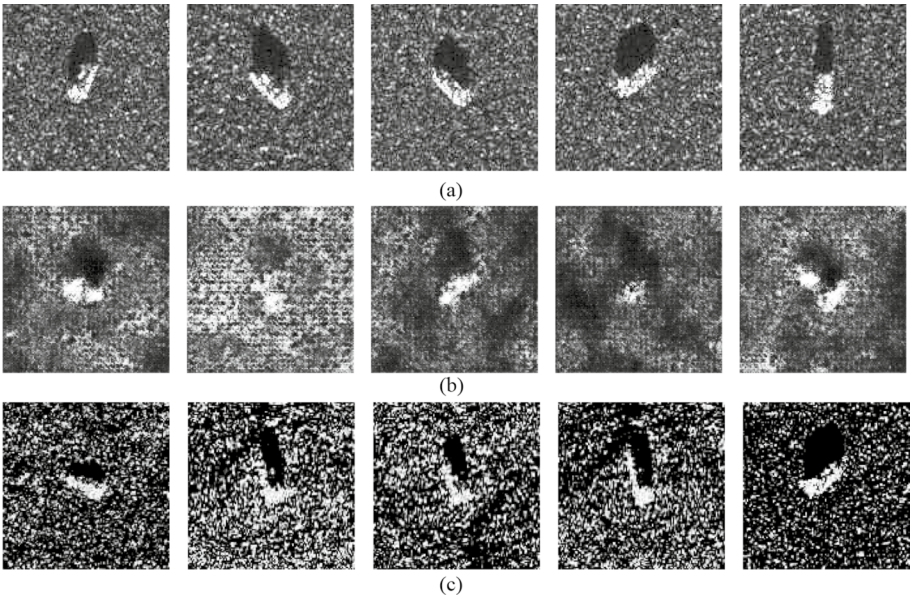


Fig. 72 The generation comparison of DCGAN and DDPM. **a** Real SAR images; **b** SAR images generated by DCGAN; **c** SAR images generated by DDPM

Table 2 The evaluation results of different generation methods

Model	MSE	PSNR	SSIM
DCGAN	7118.079	9.838	0.053
DDPM	10740.796	8.215	0.0182

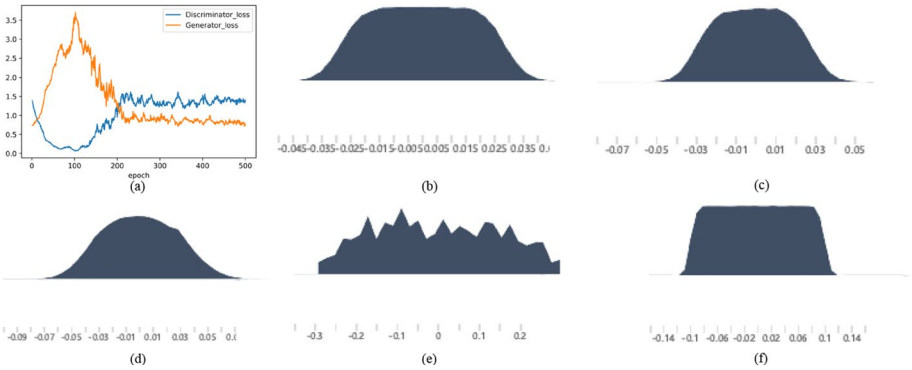


Fig. 73 The loss curves (**a**) and the visualization of generator weights (**b**–**e**) while training a DCGAN

To verify the feasibility of the diffusion model applied for SAR image generation, MSTAR is applied for training, and DDPM is applied for the generation. The dataset consists of 450 SAR images in the MSTAR dataset, and the model is trained for 500 epochs.

Besides, the DCGAN is applied for comparison in the same experimental setting, the generation results and evaluation indices are shown in Fig. 72 and Table 2. Besides, the loss curves and the visualization of weights while training a DCGAN is shown in Fig. 73.

It can be seen that the images generated by DDPM have seems better than DCGAN, and the speckle noise is clearer and more similar to the real SAR images. However, the evaluation results of DDPM are worse than that of DCGAN, and both of them cannot achieve satisfactory results, which may be caused by the small training dataset and epochs. Though the diffusion model cannot generate fine SAR images till now, the generation results show the probability and potential of diffusion models applied in this field. In addition, the loss curves of generator and discriminator oscillate drastically in Fig. 73, showing it's difficult for DCGAN to achieve optimal parameters for SAR image generation. And the weight of the last transposed convolutional layer of the generator converged between -0.14 to 0.14 finally.

However, by comparing with GANs, it's relatively difficult to add the conditional information for the diffusion-based models. In the recent works Choi et al. (2021); Liu et al. (2021b), a specific guidance function is introduced to control the output of each generation step, and the results meeting the specific condition can be obtained. Besides, in work Singh et al. (2022), the author claims that the diffused noise has a direct correlation to the saliency and orientation of the input image, which can be applied to generate the image with the same location and orientation as the input image. Hence, the inverse gradient is introduced to update the noise by calculating the gradient of the generated and real images. The generation result is shown in Fig. 74, showing that the generated images have the same orientation as the input images. Due to the wide application of SAR image azimuth interpolation, the proposed model can be modified and applied in this field in the future.

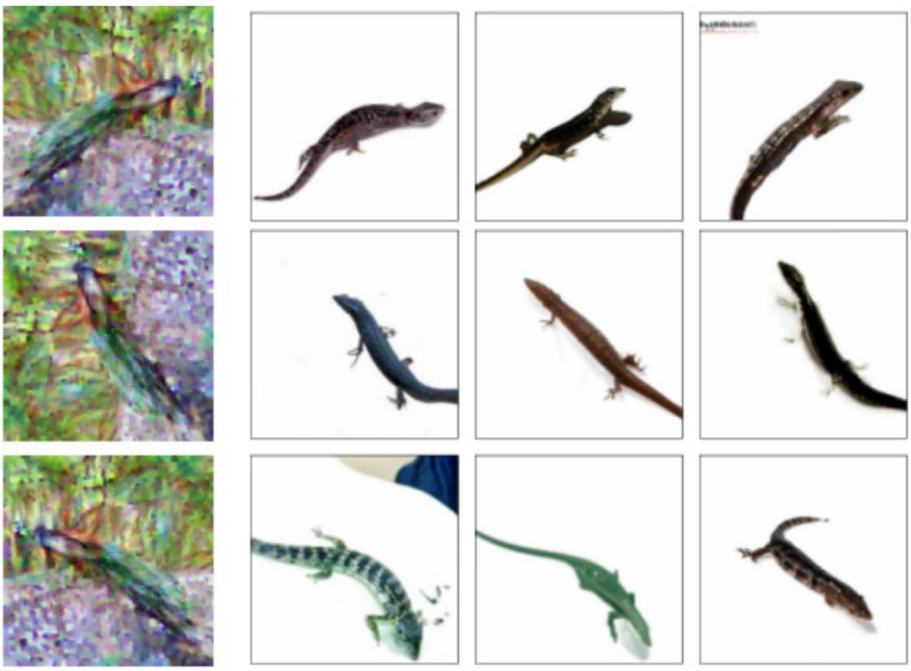


Fig. 74 The generation results of different orientations

8 Conclusion and future work

In this paper, a comprehensive survey of SAR/ISAR image generation approaches based on the deep generative models is proposed. The generation approaches are summarized and grouped by considering the different application fields, including SAR/ISAR data augmentation, SAR/ISAR image translation, SAR/ISAR image enhancement, azimuth interpolation, and deceptive jamming. After that, the related works are grouped based on the different deep generative models, such as the GAN, DCGAN, cGAN, InfoGAN, CycleGAN, Pix2PixGAN, and some other improved GANs. Besides, some other deep generative models, such as auto-encoder and probabilistic graphical models can also be applied for SAR/ISAR image generation. This review summarizes 87 related papers and 5 related survey papers from 2017 to 2022, and the systematic analysis of these papers is summarized as follows:

- For application fields, SAR/ISAR data augmentation, SAR/ISAR image translation, and SAR/ISAR image enhancement are the three most involved applications. In the tasks of SAR/ISAR image translation, the application of SAR-to-optical is much more than optical-to-SAR translation. And in the tasks of SAR/ISAR image enhancement, the reconstruction of super-resolution SAR/ISAR images is the most common application.
- For deep generative models, GANs and improved GANs are the most used generation networks. The conditional information can be easily introduced to the generator of GANs, and the satisfactory generated images can be obtained based on adversarial learning.
- For the applied dataset, MSTAR is the most used dataset in the task of data augmentation, and the SEN1-2 dataset is mostly applied for image translation. Besides, all datasets applied for ISAR image generation are self-simulated or self-captured images.
- For evaluation indices, recognition accuracy is the most used evaluation approach, which is intuitive and easy to be employed. After that, SSIM is always combined with PSNR and MSE to evaluate the generation performance of the models together.

The analysis indicates that there still exists challenges and opportunities in the field of SAR/ISAR image generation. For instance, though the GAN-based models are the most applied, the problem of mode collapse is still the main problem of these models. Hence, with the development of other deep generative models, such as diffusion-based models, the generation networks are not limited to GANs, and researchers tend to develop a novel model for the precise generation of SAR/ISAR images. Diffusion-based models are more stable than the traditional GAN-based models, which have no need for adversarial training (Ho et al. 2020). By comparing with GANs, the diffusion-based models have novel architecture based on the probability graph and can obtain abundant images using different sampling approaches (Choi et al. 2021). Besides, novel evaluation indices can be proposed in the future to measure the difference between the real and generated images systematically, which can also be applied as the objective functions to improve the generation performance. This review can be learned and referenced by SAR/ISAR image researchers, image generation researchers, and artificial intelligence researchers, and can provide the advanced idea for them.

Acknowledgements This work was supported in part by National Key Research and Development Program of China under Grant 2021YFB3100800, in part by the National Natural Science Foundation of China under Grants 62022091,62201588, and 61921001.

Author contributions JZ wrote the main manuscript text and ZL, WJ, YL and XL offered guidance and suggestions, and proofread the manuscript. XZ prepared Fig. 2, 3, 4. All authors reviewed the manuscript.

Data availability Due to this is a survey paper, we only use an open-access data in Sect. 7.3. The data that support the findings of this study are openly available in MSTAR at https://www.sdms.afrl.af.mil/index.php?collection=registration.

Declarations

Conflicts of interest We do not have any conflict of interest in this paper.

References

Ai J, Fan G, Mao Y, Jin J, Xing M, Yan H (2021) An improved SRGAN based ambiguity suppression algorithm for SAR ship target contrast enhancement. IEEE Geosci Remote Sens Lett 19:1–5

Arjovsky M, Bottou L (2017) Towards principled methods for training generative adversarial networks. arXiv:1701.04862

Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. In: International conference on machine learning, PMLR, pp 214–223

Baier G, Deschemps A, Schmitt M, Yokoya N (2021) Synthesizing optical and SAR imagery from land cover maps and auxiliary raster data. IEEE Trans Geosci Remote Sens 60:1–12

Bank D, Koenigstein N, Giryes R (2020) Autoencoders. arXiv:2003.05991

Bao X, Pan Z, Liu L, Lei B (2019) SAR image simulation by generative adversarial networks. In: IGARSS 2019-2019 IEEE international geoscience and remote sensing symposium, IEEE, pp 9995–9998

Barratt S, Sharma R (2018) A note on the inception score. arXiv:1801.01973

Bermudez JD, Happ PN, Feitosa RQ, Oliveira DA (2019) Synthesis of multispectral optical images from SAR/optical multitemporal data using conditional generative adversarial networks. IEEE Geosci Remote Sens Lett 16(8):1220–1224

Bernardi R, Cakici R, Elliott D, Erdem A, Erdem E, Ikizler-Cinbis N, Keller F, Muscat A, Plank B (2016) Automatic description generation from images: a survey of models, datasets, and evaluation measures. J Artif Intell Res 55:409–442

Bhamidipati SRM, Srivatsa C, Kanakapura Shivabasave Gowda C, Vadada S (2020) Generation of SAR images using deep learning. SN Comput Sci 1(6):1–9

Bi H, Bi G, Zhang B, Hong W, Wu Y (2019) From theory to application: real-time sparse SAR imaging. IEEE Trans Geosci Remote Sens 58(4):2928–2936

Ca L, Chen ZX, Yun S, Js C, Hasi T, Hz PAN (2019) Research advances of SAR remote sensing for agriculture applications: a review. J Integr Agric 18(3):506–525

Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE). Geosci Model Dev Discussions 7(1):1525–1534

Chang YL, Chiang CY, Chen K (2011) SAR image simulation with application to target recognition. Progr Electromagn Res 119:35–57

Chen H, Zhang Y, Wang H, Ding C (2012) SAR imaging simulation for urban structures based on analytical models. IEEE Geosci Remote Sens Lett 9(6):1127–1131

Chen J, Li Z, Huang B (2017) Linear spectral clustering superpixel. IEEE Trans Image Process 26(7):3317–3330

Chen X, Duan Y, Houthooft R, Schulman J, Sutskever I, Abbeel P (2016) Infogan: Interpretable representation learning by information maximizing generative adversarial nets. Advances in neural information processing systems 29

Chen J, Xie B, Zhang H, Zhai J (2019) Deep autoencoders in pattern recognition: a survey. In: Bio-inspired computing models and algorithms, World Scientific, pp 229–255

Choi J, Kim S, Jeong Y, Gwon Y, Yoon S (2021) Ilvr: Conditioning method for denoising diffusion probabilistic models. In 2021 IEEE. In: CVF international conference on computer vision (ICCV), pp 14347–14356

Cui Z, Zhang M, Cao Z, Cao C (2019) Image data augmentation for SAR sensor via generative adversarial nets. IEEE Access 7:42255–42268

de Almeida FQ, Younis M, Krieger G, Moreira A (2018) Multichannel staggered SAR azimuth processing. IEEE Trans Geosci Remote Sens 56(5):2772–2788

Denis L, Dalsasso E, Tupin F (2021) A review of deep-learning techniques for SAR image restoration. In: 2021 IEEE international geoscience and remote sensing symposium IGARSS, IEEE, pp 411–414
Dhillon IS, Guan Y, Kulis B (2007) Weighted graph cuts without eigenvectors a multilevel approach. IEEE Trans Pattern Anal Mach Intell 29(11):1944–1957
Dietrich-Sussner R, Davari A, Seehaus T, Braun M, Christlein V, Maier A, Riess C (2021) Synthetic Glacier SAR Image Generation from Arbitrary Masks Using Pix2Pix Algorithm. In: 2021 IEEE international geoscience and remote sensing symposium IGARSS, IEEE, pp 4548–4551
Ding J, Chen B, Liu H, Huang M (2016) Convolutional neural network with data augmentation for SAR target recognition. IEEE Geosci Remote Sens Lett 13(3):364–368
Doi K, Sakurada K, Onishi M, Iwasaki A (2020) GAN-Based SAR-to-Optical Image Translation with Region Information. In: IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, IEEE, pp 2069–2072
Du C, Xie P, Zhang L, Ma Y, Tian L (2021) Conditional prior probabilistic generative model with similarity measurement for ISAR imaging. IEEE Geosci Remote Sens Lett 19:1–5
Du L, Lyu G, Shi Y (2021) ISAR image semantic segmentation based on GAN. Radar Sci Technol 19(5):479–484
Du S, Hong J, Wang Y, Qi Y (2021) A high-quality multicategory SAR images generation method with multiconstraint GAN for ATR. IEEE Geosci Remote Sens Lett 19:1–5
Du S, Hong J, Wang Y, Xing K, Qiu T (2021d) Multi-category SAR images generation based on improved generative adversarial network. In: 2021 IEEE international geoscience and remote sensing symposium IGARSS, IEEE, pp 4260–4263
Enomoto K, Sakurada K, Wang W, Kawaguchi N, Matsuoka M, Nakamura R (2018) Image translation between SAR and optical imagery with generative adversarial nets. In: IGARSS 2018-2018 IEEE international geoscience and remote sensing symposium, IEEE, pp 1752–1755
Fan J, Wang Q, Liu G, Zhang L, Guo Z, Tong L, Peng J, Yuan W, Zhou W, Yan J et al (2019) Monitoring and analyzing mountain glacier surface movement using SAR data and a terrestrial laser scanner: a case study of the Himalayas North Slope Glacier Area. Remote Sens 11(6):625
Fan W, Zhou F, Zhang Z, Bai X, Tian T (2020) Deceptive jamming template synthesis for SAR based on generative adversarial nets. Signal Process 172:107528
Farmanov N, Amankulova K, Szatmári J, Sharifi A, Abbasi-Moghadam D, Mirhossein-Nejad M, Mucsi L (2023) Crop type classification by DESIS hyperspectral imagery and machine learning algorithms. IEEE J Selected Top Appl Earth Observ Remote Sens
Franceschetti G, Migliaccio M, Riccio D (1995) The SAR simulation: an overview. In: 1995 international geoscience and remote sensing symposium, IGARSS'95. Quantitative Remote Sensing for Science and Applications, IEEE, vol 3, pp 2283–2285
Fu S, Xu F, Jin YQ (2021) Reciprocal translation between SAR and optical remote sensing images with cascaded-residual adversarial networks. Sci China Inf Sci 64(2):1–15
Fuentes Reyes M, Auer S, Merkle N, Henry C, Schmitt M (2019) Sar-to-optical image translation based on conditional generative adversarial networks-optimization, opportunities and limits. Remote Sens 11(17):2067
Gao F, Yang Y, Wang J, Sun J, Yang E, Zhou H (2018) A deep convolutional generative adversarial networks (DCGANs)-based semi-supervised method for object recognition in synthetic aperture radar (SAR) images. Remote Sens 10(6):846
Gao J, Yuan Q, Li J, Zhang H, Su X (2020) Cloud removal with fusion of high resolution optical and SAR images using generative adversarial networks. Remote Sens 12(1):191
Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics, JMLR Workshop and Conference Proceedings, pp 249–256
Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2020) Generative adversarial networks. Commun ACM 63(11):139–144
Green R (1998) The sensitivity of SAR backscatter to forest windthrow gaps. Int J Remote Sens 19(12):2419–2425
Gu F, Zhang H, Wang C, Wu F (2019) Sar image super-resolution based on noise-free generative adversarial network. In: IGARSS 2019-2019 IEEE international geoscience and remote sensing symposium, IEEE, pp 2575–2578
Guo J, He C, Zhang M, Li Y, Gao X, Song B (2021) Edge-preserving convolutional generative adversarial networks for SAR-to-optical image translation. Remote Sens 13(18):3575
Hassan M, Bhagvati C (2012) Structural similarity measure for color images. Int J Comput Appl 43(14):7–12

He C, Xiong D, Zhang Q, Liao M (2019) Parallel connected generative adversarial network with quadratic operation for SAR image generation and application for classification. Sensors 19(4):871
Heygster G, Dannenberg J, Notholt J (2009) Topographic mapping of the German tidal flats analyzing SAR images with the waterline method. IEEE Trans Geosci Remote Sens 48(3):1019–1030
Ho J, Jain A, Abbeel P (2020) Denoising diffusion probabilistic models. Adv Neural Inf Process Syst 33:6840–6851
Huang H, Zhang F, Zhou Y, Yin Q, Hu W (2019) High Resolution SAR image synthesis with hierarchical generative adversarial networks. In: IGARSS 2019-2019 IEEE international geoscience and remote sensing symposium, IEEE, pp 2782–2785
Huang B, Zhi L, Yang C, Sun F, Song Y (2020) Single satellite optical imagery dehazing using SAR image prior based on conditional generative adversarial networks. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision, pp 1806–1813
Huang Y, Mei W, Liu S, Li T (2022) Asymmetric training of generative adversarial network for high fidelity SAR image generation. In: IGARSS 2022-2022 IEEE international geoscience and remote sensing symposium, IEEE, pp 1576–1579
Hwang J, Shin Y (2021) Image Data augmentation for SAR automatic target recognition using Triple-GAN. In: 2021 international conference on information and communication technology convergence (ICTC), IEEE, pp 312–314
Hwang J, Yu C, Shin Y (2020) SAR-to-optical image translation using SSIM and perceptual loss based cycle-consistent GAN. In: 2020 international conference on information and communication technology convergence (ICTC), IEEE, pp 191–194
Isola P, Zhu JY, Zhou T, Efros AA (2017) Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1125–1134
Ji G, Wang Z, Zhou L, Xia Y, Zhong S, Gong S (2020) SAR image colorization using multidomain cycle-consistency generative adversarial network. IEEE Geosci Remote Sens Lett 18(2):296–300
Jie W (2020) Multi-source fusion of ship targets recognition methods based on generative adversarial networks. Master's thesis, Harbin Institute of Technology
Jing W, Jin T, Xiang D (2021) Content-sensitive superpixel generation for SAR images with edge penalty and contraction-expansion search strategy. IEEE Trans Geosci Remote Sens 60:1–15
Jozdani S, Chen D, Pouliot D, Johnson BA (2022) A review and meta-analysis of generative adversarial networks and their applications in remote sensing. Int J Appl Earth Observ Geoinf 108:102734
Khoshboresh-Masouleh M, Alidoost F, Arefi H (2020) Multiscale building segmentation based on deep learning for remote sensing RGB images from different sensors. J Appl Remote Sens 14(3):034503–034503
Kingma DP, Welling M (2013) Auto-encoding variational bayes. arXiv:1312.6114
Kong J, Zhang F (2021) SAR target recognition with generative adversarial network (GAN)-based Data Augmentation. In: 2021 13th international conference on advanced infocomm technology (ICAIT), IEEE, pp 215–218
Korhonen J, You J (2012) Peak signal-to-noise ratio revisited: is simple beautiful? In: 2012 fourth international workshop on quality of multimedia experience, IEEE, pp 37–38
Kudryavtsev V, Kozlov I, Chapron B, Johannessen J (2014) Quad-polarization SAR features of ocean currents. J Geophys Res 119(9):6046–6065
Lalitha V, Latha B (2022) A review on remote sensing imagery augmentation using deep learning. Mater Today
Lange J (2019) Translation of remote sensing images for the classification of unlabeled SAR data using Deep Convolutional Generative Adversarial Networks. PhD thesis, Humboldt University of Berlin
LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
Li Z, Bao Z, Wang H, Liao G (2006) Performance improvement for constellation SAR using signal processing techniques. IEEE Trans Aerospace Electr Syst 42(2):436–452
Li G, Sun Z, Zhang Y (2019a) Isar target recognition using pix2pix network derived from cgan. In: 2019 international radar conference (RADAR), IEEE, pp 1–4
Li J, Qu C, Peng S, Jiang Y (2019) Ship detection in SAR images based on generative adversarial network and online hard examples mining. J Electr Inf Technol 41(1):143–149
Li Y, Ao D, Dumitru CO, Hu C, Datcu M (2019) Super-resolution of geosynchronous synthetic aperture radar images using dialectical GANs. Sci China Inf Sci 62(10):1–3
Li L, Wang C, Zhang H, Zhang B (2020a) SAR image ship object generation and classification with improved residual conditional generative adversarial network. IEEE Geosci Remote Sens Lett
Li Y, Fu R, Meng X, Jin W, Shao F (2020) A SAR-to-optical image translation method based on conditional generation adversarial network (cGAN). IEEE Access 8:60338–60343

Li X, Li H, Wang C, Hu X, Zhang W (2021) Visual-attention GAN for interior sketch colourisation. IET Image Process 15(4):997–1007
Li X, Zhang G, Cui H, Hou S, Wang S, Li X, Chen Y, Li Z, Zhang L (2022) MCANet: a joint semantic segmentation framework of optical and SAR images for land use classification. Int J Appl Earth Observ Geoinf 106:102638
Li W, Yuan Y, Zhang Y, Luo Y (2022) Unblurring ISAR imaging for maneuvering target based on UFGAN. Remote Sens 14(20):5270
Li JT, Bian Z, Guo LX (2022) Optimized complex object classification model: reconstructing the ISAR image of a hypersonic vehicle covered with a plasma sheath using a U-WGAN-GP framework. Int J Remote Sens 43(14):5306–5323
Liang M (2021) Research on multi-view SAR image target data augmentation based on generative adversarial networks. Master's thesis, University of Electronic Science and Technology of China
Liu Y (2021) Improved generative adversarial network and its application in image oil painting style transfer. Image Vision Comput 105:104087
Liu X, Huang Y, Wang C, Pei J, Huo W, Zhang Y, Yang J (2021a) Semi-supervised sar atr via conditional generative adversarial network with multi-discriminator. In: 2021 IEEE international geoscience and remote sensing symposium IGARSS, IEEE, pp 2361–2364
Liu X, Park DH, Azadi S, Zhang G, Chopikyan A, Hu Y, Shi H, Rohrbach A, Darrell T (2021b) More control for free! image synthesis with semantic diffusion guidance. arXiv:2112.05744
Long Y, Juan S, Xiang L (2019) Application of SAR ship data augmentation based on generative adversarial network in improved SSD. Acta Armamentarii 40(12):2488
Lu Q, Jiang H, Li G, Ye W (2019a) Data augmentation method of sar image dataset based on wasserstein generative adversarial networks. In: 2019 International conference on electronic engineering and informatics (EEI), IEEE, pp 488–490
Lu Q, Li G, Ye W (2019b) A SAR image data augmentation method based on generative adversarial network. In: China High Resolution Earth Observation Conference, Springer, pp 419–430
Luo Z, Jiang X, Liu X (2020) Synthetic minority class data by generative adversarial network for imbalanced sar target recognition. In: IGARSS 2020-2020 IEEE international geoscience and remote sensing symposium, IEEE, pp 2459–2462
Luo Y, Pi D (2022) SAR-to-optical image translation for quality enhancement. J Ambient Intell Hum Comput pp 1–16
Ma P, Li C, Rahaman MM, Yao Y, Zhang J, Zou S, Zhao X, Grzegorzek M (2022) A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches. Artif Intell Rev pp 1–72
Mirza M, Osindero S (2014) Conditional generative adversarial nets. arXiv:1411.1784
Mohammadi M, Sharifi A, Hosseingholizadeh M, Tariq A (2021) Detection of oil pollution using SAR and optical remote sensing imagery: a case study of the Persian Gulf. J Indian Soc Remote Sens 49(10):2377–2385
Niu X, Yang D, Yang K, Pan H, Dou Y (2018) Image translation between high-resolution remote sensing optical and SAR data using conditional GAN. In: Pacific Rim conference on multimedia, Springer, pp 245–255
Núñez J, Cincotta P, Wachlin F (1996) Information entropy. In: Chaos in Gravitational N-Body Systems, Springer, pp 43–53
Odena A, Olah C, Shlens J (2017) Conditional image synthesis with auxiliary classifier gans. In: International conference on machine learning, PMLR, pp 2642–2651
Perry R, Dipietro R, Fante R (1999) SAR imaging of moving targets. IEEE Trans Aerospace Electron Syst 35(1):188–200
Prickett M, Chen C (1980) Principles of inverse synthetic aperture radar/ISAR/imaging. In: EASCON'80; electronics and aerospace systems conference, pp 340–345
Qin D, Gao X (2020) Enhancing ISAR resolution by a generative adversarial network. IEEE Geosci Remote Sens Lett 18(1):127–131
Qin J, Liu Z, Ran L, Xie R, Tang J, Guo Z (2022) A target SAR image expansion method based on conditional wasserstein deep convolutional GAN for automatic target recognition. IEEE J Select Top Appl Earth Observ Remote Sens 15:7153–7170
Qingling L, Wei Y (2020) A survey of generative adversarial network applications for SAR image processing. Telecommun Eng 60(1):121–128
Qin D, Liu D, Gao X, Jingkun G (2019) ISAR resolution enhancement using residual network. In: 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), IEEE, pp 788–792
Qiongnan H, Weigang Z, yonggang L (2021) Review on SAR data expansion based on GAN. J Ordnance Equip Eng 42:31–38

Radford A, Metz L, Chintala S (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434
Renga A, Graziano MD, Moccia A (2018) Segmentation of marine SAR images by sublook analysis and application to sea traffic monitoring. IEEE Trans Geosci Remote Sens 57(3):1463–1477
Schlegl T, Seeböck P, Waldstein SM, Schmidt-Erfurth U, Langs G (2017) Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: Information Processing in Medical Imaging: 25th International Conference, IPMI 2017, Boone, NC, USA, June 25-30, 2017, Proceedings, Springer, pp 146–157
Shao Z, Zhang X, Zhang T (2022) GAN with ASPP for SAR Image to Optical Image Conversion. In: IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium, IEEE, pp 3355–3358
Sharifi A (2020) Flood mapping using relevance vector machine and SAR data: a case study from Aqqala, Iran. J Indian Soc Remote Sens 48(9):1289–1296
Sharifi A (2021) Yield prediction with machine learning algorithms and satellite images. J Sci Food Agric 101(3):891–896
Sharifi A, Mahdipour H, Moradi E, Tariq A (2022) Agricultural field extraction with deep learning algorithm and satellite imagery. J Indian Soc Remote Sens pp 1–7
Shen H, Lin L, Li J, Yuan Q, Zhao L (2020) A residual convolutional neural network for polarimetric SAR image super-resolution. ISPRS J Photogrammetry Remote Sens 161:90–108
Shi H, Zhang B, Wang Y, Cui Z, Chen L (2022) SAR-to-optical image translating through generate-validate adversarial networks. IEEE Geosci Remote Sens Lett 19:1–5
Singh P, Diwakar M, Shankar A, Shree R, Kumar M (2021) A review on SAR image and its despeckling. Arch Comput Methods Eng 28(7):4633–4653
Singh V, Jandial S, Chopra A, Ramesh S, Krishnamurthy B, Balasubramanian VN (2022) On conditioning the input noise for controlled image generation with diffusion models. arXiv:2205.03859
Smith JW, Alimam Y, Vedula G, Torlak M (2022) A vision transformer approach for efficient near-field SAR super-resolution under array perturbation. In: 2022 IEEE Texas symposium on wireless and microwave circuits and systems (WMCS), IEEE, pp 1–6
Sohl-Dickstein J, Weiss E, Maheswaranathan N, Ganguli S (2015) Deep unsupervised learning using nonequilibrium thermodynamics. In: International conference on machine learning, PMLR, pp 2256–2265
Solimene R, Catapano I, Gennarelli G, Cuccaro A, Dell'Aversano A, Soldovieri F (2014) SAR imaging algorithms and some unconventional applications: a unified mathematical overview. IEEE Signal Process Mag 31(4):90–98
Soloveitchik M, Diskin T, Morin E, Wiesel A (2021) Conditional frechet inception distance. arXiv:2103.11521
Song Q, Xu F (2017) Zero-shot learning of SAR target feature space with deep generative neural networks. IEEE Geosci Remote Sens Lett 14(12):2245–2249
Song Q, Xu F, Zhu XX, Jin YQ (2021) Learning to generate SAR images with adversarial autoencoder. IEEE Trans Geosci Remote Sens 60:1–15
Song Q, Xu F, Jin YQ (2019) SAR image representation learning with adversarial autoencoder networks. In: IGARSS 2019-2019 IEEE international geoscience and remote sensing symposium, IEEE, pp 9498–9501
Sun Y, Jiang W, Yang J, Li W (2022) SAR target recognition using cGAN-Based SAR-to-optical image translation. Remote Sens 14(8):1793
Sung F, Yang Y, Zhang L, Xiang T, Torr PH, Hospedales TM (2018) Learning to compare: Relation network for few-shot learning. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1199–1208
Tomiyasu K (1978) Tutorial review of synthetic-aperture radar (SAR) with applications to imaging of the ocean surface. Proc IEEE 66(5):563–583
Topouzelis KN (2008) Oil spill detection by SAR images: dark formation detection, feature extraction and classification algorithms. Sensors 8(10):6642–6659
Toriya H, Dewan A, Kitahara I (2019) SAR2OPT: Image alignment between multi-modal images using generative adversarial networks. In: IGARSS 2019-2019 IEEE international geoscience and remote sensing symposium, IEEE, pp 923–926
Vehmas R, Neuberger N (2021) Inverse synthetic aperture radar imaging: a historical perspective and state-of-the-art survey. IEEE Access
Wang K, Zhang G, Leng Y, Leung H (2018) Synthetic aperture radar image generation with deep generative models. IEEE Geosci Remote Sens Lett 16(6):912–916

Wang J, Li J, Sun B (2019) SAR image synthesis based on conditional generative adversarial networks. J Eng 21:8093–8097
Wang L, Xu X, Yu Y, Yang R, Gui R, Xu Z, Pu F (2019) SAR-to-optical image translation using supervised cycle-consistent adversarial networks. IEEE Access 7:129136–129149
Wang H, Li K, Lu X, Zhang Q, Luo Y, Kang L (2022) ISAR resolution enhancement method exploiting generative adversarial network. Remote Sens 14(5):1291
Wang R, Zhang H, Han B, Zhang Y, Guo J, Hong W, Sun W, Hu W (2022) Multiangle SAR dataset construction of aircraft targets based on angle interpolation simulation. J Radars 10:1–15
Wang C, Pei J, Liu X, Huang Y, Mao D, Zhang Y, Yang J (2022a) SAR target image generation method using azimuth-controllable generative adversarial network. IEEE J Selected Top Appl Earth Observ Remote Sens
Wang X, Xie L, Dong C, Shan Y (2021) Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 1905–1914
Wang P, Zhang H, Patel VM (2017) Generative adversarial network-based restoration of speckled SAR images. In: 2017 IEEE 7th international workshop on computational advances in multi-sensor adaptive processing (CAMSAP), IEEE, pp 1–5
Wang L, Zheng M, Du W, Wei M, Li L (2018) Super-resolution SAR image reconstruction via generative adversarial network. 2018 12th international symposium on antennas. Propagation and EM Theory (ISAPE), IEEE, pp 1–4
Wei J, Zou H, Sun L, Cao X, Li M, He S, Liu S (2022) Generative Adversarial Network for SAR-to-Optical Image Translation with Feature Cross-Fusion Inference. In: IGARSS 2022-2022 IEEE international geoscience and remote sensing symposium, IEEE, pp 6025–6028
Wiley CA (1985) Synthetic aperture radars. IEEE Trans Aerospace Electronic Syst 3:440–443
Wu C (1976) A digital system to produce imagery from SAR data. In: Systems design driven by sensors, p 968
Xiang D, Tang T, Quan S, Guan D, Su Y (2019) Adaptive superpixel generation for SAR images with linear feature clustering and edge constraint. IEEE Trans Geosci Remote Sens 57(6):3873–3889
Xiao C, Gao X, Zhang C (2021) Deep convolution network with sparse prior for sparse ISAR image enhancement. In: 2021 2nd information communication technologies conference (ICTC), IEEE, pp 54–59
Xie D, Ma J, Li Y, Liu X (2021) Data Augmentation of Sar Sensor Image via Information Maximizing Generative Adversarial Net. In: 2021 IEEE 4th international conference on electronic information and communication technology (ICEICT), IEEE, pp 454–458
Yamaguchi Y (2012) Disaster monitoring by fully polarimetric SAR data acquired with ALOS-PALSAR. Proc IEEE 100(10):2851–2860
Yan J, Li G, Su J (2022) SAR aircraft data sets augmentation based on multi-scale generative adversarial network. Electron Opt Control 29(7):62–68
Yang R, Wang R, Deng Y, Jia X, Zhang H (2020) Rethinking the random cropping data augmentation method used in the training of CNN-based SAR image ship detector. Remote Sens 13(1):34
Yang X, Wang Z, Zhao J, Yang D (2022) FG-GAN: a fine-grained generative adversarial network for unsupervised SAR-to-optical image translation. IEEE Trans Geosci Remote Sens 60:1–11
Yang X, Zhao J, Wei Z, Wang N, Gao X (2022) SAR-to-optical image translation based on improved CGAN. Pattern Recogn 121:108208
Yanshan L, Li Z, Fan X, Shifu C (2022) OGSRN: optical-guided super-resolution network for SAR image. Chin J Aeron 35(5):204–219
Yates G, Horne A, Blake A, Middleton R (2006) Bistatic SAR image formation. IEE Proc-Radar Sonar Navig 153(3):208–213
Yuan H, Li H, Zhang Y, Wang Y, Liu Z, Wei C, Yao C (2022) High-resolution refocusing for defocused ISAR images by complex-valued Pix2pixHD network. IEEE Geosci Remote Sens Lett 19:1–5
Yuan Y, Luo Y, Ni J, Zhang Q (2022) Inverse synthetic aperture radar imaging using an attention generative adversarial network. Remote Sens 14(15):3509
Yu N, Ma A, Zhong Y, Gong X (2022) HFGAN: a heterogeneous fusion generative adversarial network for Sar-to-optical image translation. In: IGARSS 2022-2022 IEEE international geoscience and remote sensing symposium, IEEE, pp 2864–2867
Zenati H, Foo CS, Lecouat B, Manek G, Chandrasekhar VR (2018) Efficient GAN-based anomaly detection. arXiv:1802.06222
Zhai J, Dang X, Chen F, Xie X, Zhu Y, Yin H (2019) SAR image generation using structural bayesian deep generative adversarial network. In: 2019 Photonics & electromagnetics research symposium-fall (PIERS-Fall), IEEE, pp 1386–1392

Zhang S (2021) Research and implementation of sample amplification technology based on deep learning. Master's thesis, University of Electronic Science and Technology of China
Zhang L, Zhang L, Mou X, Zhang D (2011) FSIM: a feature similarity index for image quality assessment. IEEE Trans Image Process 20(8):2378–2386
Zhang J, Zhou J, Lu X (2020) Feature-guided SAR-to-optical image translation. IEEE Access 8:70925–70937
Zhang L, Lu S, Hu C, Xiang D, Liu T, Su Y (2021) Superpixel generation for SAR imagery based on fast DBSCAN clustering with edge penalty. IEEE J Select Top Appl Earth Observ Remote Sens 15:804–819
Zhang Q, Liu X, Liu M, Zou X, Zhu L, Ruan X (2021) Comparative analysis of edge information and polarization on sar-to-optical translation based on conditional generative adversarial networks. Remote Sens 13(1):128
Zhang J, Ma P, Jiang T, Zhao X, Tan W, Zhang J, Zou S, Huang X, Grzegorzek M, Li C (2022) SEM-RCNN: a squeeze-and-excitation-based mask region convolutional neural network for multi-class environmental microorganism detection. Appl Sci 12(19):9902
Zhang J, Zhao X, Jiang T, Rahaman MM, Yao Y, Lin YH, Zhang J, Pan A, Grzegorzek M, Li C (2022) An application of Pixel Interval Down-sampling (PID) for dense tiny microorganism counting on environmental microorganism images. Appl Sci 12(14):7314
Zhang M, Cui Z, Wang X, Cao Z (2018) Data augmentation method of SAR image dataset. In: IGARSS 2018-2018 IEEE international geoscience and remote sensing symposium, IEEE, pp 5292–5295
Zhang J, Li C, Rahaman MM, Yao Y, Ma P, Zhang J, Zhao X, Jiang T, Grzegorzek M (2022a) A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches. Artif Intell Rev pp 1–70
Zhang J, Li C, Rahaman MM, Yao Y, Ma P, Zhang J, Zhao X, Jiang T, Grzegorzek M (2022b) A Comprehensive survey with quantitative comparison of image analysis methods for microorganism biovolume measurements. Arch Comput Methods Eng pp 1–35
Zhang J, Li C, Yin Y, Zhang J, Grzegorzek M (2022c) Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer. Artif Intell Rev pp 1–58
Zhao K, Zhou Y, Chen X, Zhang H (2020) A generative adversarial network based image augmentation method for ship segmentation in SAR images. In: 2020 IEEE 9th joint international information technology and artificial intelligence conference (ITAIC), IEEE, vol 9, pp 1285–1289
Zheng C, Jiang X, Zhang Y, Liu X, Yuan B, Li Z (2019) Self-normalizing generative adversarial network for super-resolution reconstruction of SAR images. In: IGARSS 2019-2019 IEEE international geoscience and remote sensing symposium, IEEE, pp 1911–1914
Zhou RY, Yang ZL, Wang F (2021) ISAR images generation via generative adversarial networks. In: 2021 IEEE international geoscience and remote sensing symposium IGARSS, IEEE, pp 5267–5270
Zhou Y (2020) The research of SAR image generation and data expansion based on generative adversarial network. Master's thesis, Xidian University
Zhu JY, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision, pp 2223–2232
Zou L, Zhang H, Wang C, Wu F, Gu F (2020) MW-ACGAN: generating multiscale high-resolution SAR images for ship detection. Sensors 20(22):6673
Zuo Z, Li Y (2021) A SAR-to-optical image translation method based on PIX2PIX. In: 2021 IEEE international geoscience and remote sensing symposium IGARSS, IEEE, pp 3026–3029

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.