RESEARCH

Abstract- The significance of Synthetic Aperture Radar (SAR) imaging in remote sensing is constrained by its limited spatial resolution, which results in limited clarity and detail. The super-resolution techniques used today face obstacles include a complicated network structure, limited sensing power, and trouble extracting characteristics that depend both locally and globally. In order to overcome these obstacles, DMSC-GAN, a SAR. This paper introduces the c-GAN framework-based picture super-resolution approach. Improving the model's flexibility and controllability is the design goal of DMSC-GAN by modulating the generated image features through the use of conditional inputs. The technique presents a feature extraction module that combines convolutional processes with Deformable Multi-Head Self-Attention (DMSA) and builds a generator using an encoder-decoder structure. This module is capable of effectively capturing the characteristics of objects with a wide range of shapes and extracting crucial background data required to reconstruct intricate image textures. Furthermore, a multi-scale feature extraction pyramid layer facilitates the acquisition of various sizes of image details. By utilising an upgraded dual-scale discriminator in conjunction with perceptual loss and feature matching loss, DMSC-GAN effectively pulls features from SAR pictures for superior super-resolution reconstruction. The remarkable performance of DMSC-GAN, which greatly enhances the spatial resolution and visual clarity of SAR images, has been confirmed by numerous experiments. This framework shows great potential and ability to advance SAR picture super-resolution techniques.

Keywords – Synthetic Aperture Radar (SAR) is a remote sensing technology based on phased array radar knowledge. It generates Earth surface images by coherently processing radar platform motion and multiple radar echoes.

DMSA-Deformable Multi-Head Self-Attention.

GAN-Generative Adversarial Network

c-GAN-conditional Generative Adversarial Network

SISR-Single Image Super resolution

Introduction-In the field of image SR, the task of Single-Image SR (SISR) [1] is regarded as complex and challenging to recover a high-resolution image from a single low-resolution image directly. The earliest methods applied to SR tasks were interpolation-based models. These models utilize simple techniques such as bilinear interpolation [2], bicubic interpolation [3], or nearest-neighbor interpolation [4] to generate images. These methods offer fast computation by computing the relationships between neighboring pixels to determine the target pixel values. However, they fail to recover high-frequency details, often resulting in blurry output effectively.In response to these SISR tasks have predominantly embraced deep learning models as a mainstream approach. Within the sphere of deep learning-based methods, three categories can be identified: convolutional neural network (CNN) [5], self-attention mechanism (transformer) [6], and Generative Adversarial Network (GAN) [7]. GAN is effectively applied for image generation by generating realistic data through a generative network and employing adversarial training of the discriminative network [8]. This enables the generator to learn the distribution of actual data progressively. By combining GAN with conditional learning, the development of a powerful deep learning model known as conditional Generative Adversarial Network (c-GAN) [9] was realized. The powerful generative method GAN has found successful application in SAR image SR [10]. NFGAN [11] combines SAR image denoising and super-resolution reconstruction, effectively eliminating noise from super-resolution reconstructed images. These methods incorporate a perceptual loss function comprising adversarial and content loss. Even so, more research is needed into the in-depth exploration of loss functions, as the currently introduced loss functions are relatively simplistic. Additionally, the feature extraction networks employed in these studies may need to adequately extract features, which limits the improvement in generated image quality and hinders the generation of high-resolution SAR images that meet the standards of human visual perception.

Inspired by the image super-resolution algorithms, this study proposes DMSC-GAN, a SAR image super-resolution framework based on the conditional Generative Adversarialm*Remote* Network. This framework is carefully designed to cater to the requirements of SAR image super-resolution tasks, representing a significant advancement over the conventional c GAN approaches.

GAN BASED METHODS-The demonstration of the pivotal role played by GAN in SR tasks is evident through the introduction of adversarial training mechanisms, which augment the fidelity of generated images through the competitive interaction between generators and discriminators. The foundational SRGAN [12] model, incorporating adversarial loss and perceptual loss, establishes a basis for generating high-quality super-resolution images.

C-GAN has demonstrated excellent image processing capabilities in various domains. The application of c-GAN to the semantic segmentation task by Wang et al. [13] resulted in enhanced performance through the collaborative action of the global generator and the local enhancement network, along with the integration of the PatchGAN discriminator and feature matching loss. Liu et al. [14] proposed CCWGAN, utilizing residual dense blocks to generate high-quality remote sensing images effectively. In the realm of MRI super-resolution, Nasseret al. [15] successfully elevated the performance of isotropic and anisotropic MRI super-resolution by incorporating perceptual loss and conditioning on low-resolution MRI images. Hanano et al. [16] achieved the improved generation of facial expression images by enhancing cGAN in combination with a self-supervised guided encoder. Drawing upon these methodologies, a DMSC-GAN is constructed in this study, utilizing an encoder–decoder structure and incorporating feature matching loss and perceptual loss. The framework includes two discriminators operating at distinct scales and conducting discrimination through 1×1 points. Through the integration of these components, the intended outcome is to proficiently steer the generation of high-resolution images, ultimately enhancing the accuracy and quality of the generated outcomes.

GENERAL FRAMEWORK- Existing SR algorithms primarily developed for optical images do not yield satisfactory results when directly applied to SAR image SR. In response to this limitation, a novel framework explicitly tailored for SAR image super-resolution based on the c-GAN architecture is introduced in this paper. Unlike non-GAN models that minimize the variance of mean square error for image reconstruction, resulting in smooth images with reduced high-frequency details, GAN-generated images have been shown to effectively enhance visual clarity [12]. The proposed model, DMSC-GAN, as depicted in Figure 1, comprises a generator with an encoder–decoder structure and two discriminators at different scales. In addition to generating the adversarial loss function, network training also uses perceptual loss and feature matching loss. The main steps of network training are as follows:

prepare the dataset, initialize the parameters, set the loss function, create the generator generate super-resolution images using adversarial training, update the network parameters using gradient descent, iteratively train the generator and the discriminators until a satisfactory result is obtained, and finally use the generator to generate images. The experimental process is divided into two stages: the training stage, where the complete network model is trained, and the testing stage, where only the trained generator is utilized to generate target images.

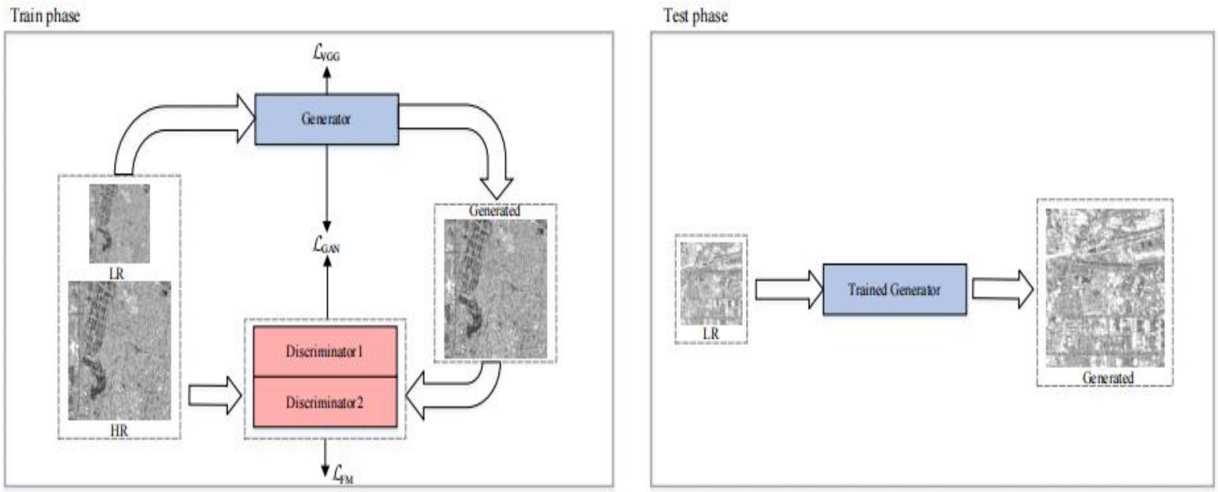


Fig.1 General framework of DMSC-GAN, including an enhanced feature extraction generator and discriminator

GENERATOR-The proposed generator structure in this paper, as depicted in Figure 2, is based on an encoder–decoder architecture. The encoder is on the left side, while the decoder is on the right. To facilitate feature reuse, skip connections are introduced between the encoder and decoder. This encoder–decoder design enables hierarchical feature extraction of the input data, capturing abstract and high-level semantic features that are subsequently recombined by the decoder. The generator effectively preserves rich details and texture features by incorporating multi-scale feature information, thereby improving image.

The generator takes in low-resolution SAR images, denoted as $I_{LR} \in R^{3\times H\times W}$, as input. Upon reading, the original SAR image, a single-channel grayscale image, is converted to RGB mode

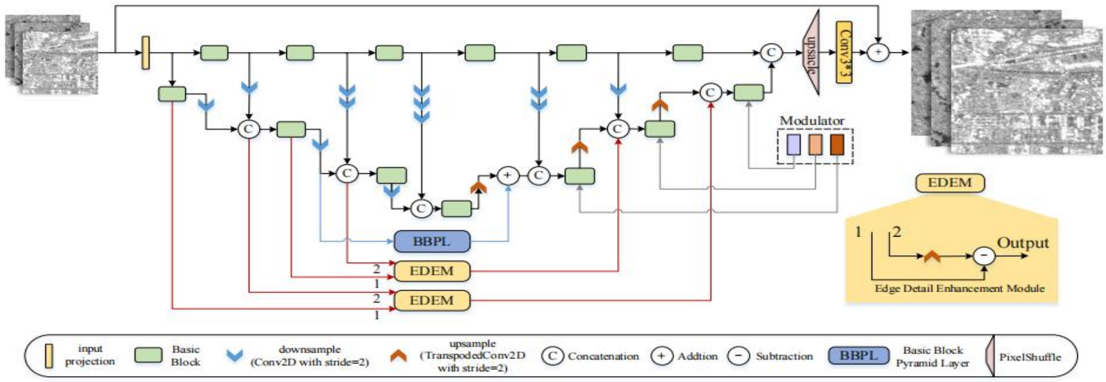
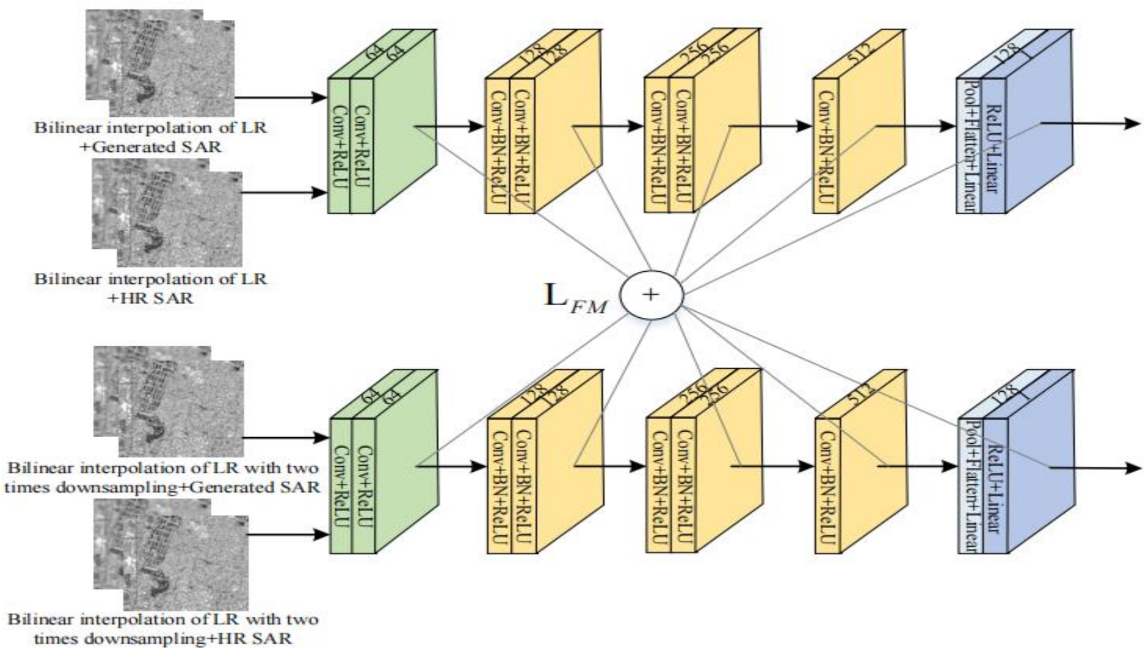


Fig.2 Generator4

DISCRIMINATOR- Designing a suitable discriminator for SAR image generation tasks presents a challenge due to the distinct characteristics of SAR images compared to other optical remote sensing images. SAR images are typically grayscale and encompass intense information and advanced features like polarization and phase, enabling sophisticated analysis and interpretation. Moreover, SAR images exhibit intricate texture details. Thus, to generate SAR images that are more realistic, it is essential to develop a discriminator that is carefully tailored to these unique attributes. We propose a multi-scale discriminator design to address the limitations of relying solely on a single 1×1 point for discrimination in SAR image SR tasks. The multi-scale discriminator consists of two discriminators with the same structure but different input image scales. This design allows for more reliable discrimination information. The structure of the multi-scale discriminator is depicted in Figure 3



Our discriminator operates in two input modes to enable effective discrimination between the generated SAR images and authentic high-resolution SAR images. In one mode, the upsampled result of the low-resolution SAR image is concatenated with the generated image. In contrast, in the other mode, it is concatenated with the actual image. To handle these two input scales, we employ two discriminators. The first discriminator discriminates the 256 × 256 images corresponding to the generated images. The second discriminator focuses on the down sampled results: the 128 × 128 images.

GENERATE ADVERSARIAL LOSS- Two discriminators are utilized in this study to assess the images at both the original scale and the down sampled scale. As a result, the adversarial loss comprises two components: the adversarial loss between the generator and the first discriminator, which discriminates images at the original scale, and the adversarial loss between the generator and the second discriminator, which operates at the down sampled scale. The low-resolution SAR image is denoted as x, the up sampled version of x using bilinear interpolation is represented as $x\uparrow$, and the generator's output, which is the SR image, is denoted as $G(x)$. The discriminators at the two scales are denoted as $D1$ and $D2$, corresponding to the original and down sampled scales, respectively. The low-resolution SAR image contains comprehensive information, and the up sampled low-resolution SAR image concatenation with the image under discrimination is performed. This combined input is provided to the discriminators, allowing them to leverage the valuable reference in formation embedded in the low-resolution SAR image, thus facilitating accurate judgments by the discriminators.

DATASETS- The SEN1-2 dataset [17], which consists of SAR and optical images acquired from the Sentinel-1 and Sentinel-2 satellites, was utilized in this study. The dataset is accessible through the https://mediatum.ub.tum.de/1436631. With a spatial resolution of 5 m, the dataset is commonly employed in image fusion applications. The SAR images from this dataset were specifically selected for our experimental objectives.

The initial step of the data preprocessing involves cropping the SAR images to dimensions of 256 × 256 pixels. We manually excluded regions exhibiting evident duplication or significant issues to ensure the dataset's quality. Data augmentation techniques, including rotation and mirroring, are applied to augment the dataset. To generate the required low resolution SAR images for input to DMSC-GAN, we downsample the high-resolution SAR images using bilinear interpolation

RESULT-GAN uses various alternative metrics such as MSE , FID , and LPIPS have been introduced, which enhance the generation of visually pleasing and realistic SR images during GAN training. The selection of objective metrics aims to mitigate potential inaccuracies arising from subjective evaluation when assessing the quality of the generated SR images. MSE quantifies the mean squared difference between the generated and target images, ensuring proximity in terms of pixel values. Perceived similarity, evaluated by LPIPS, considers higher-level features such as edges, texture, and overall appearance, yielding a more meaningful metric from a perceptual standpoint. On the other hand, FID employs the Inception network to assess the feature representations of the generated and authentic images, encompassing both quality and diversity aspects and offering a metric applicable to generators.

SEN1–2 dataset. In order to assess the superior performance of DMSC-GAN in SAR image SR, a comparative study was conducted, evaluating it against traditional convolutional neural network methods, including HSENET, SWINIR with transformer architecture, supervised SRGAN, unsupervised CycleGAN [49] and Pix2pixHD [34]. The experimental results for these six methods, considering upscale factor 4.

	Number of Parameters	$PSNR_{\uparrow}$	$SSIM_{\uparrow}$	$MSE_{\downarrow}$	$FID_{\downarrow}$	$LPIPS_{\downarrow}$
HSENET [23]	4.79 M	24.334956	0.626694	0.1233	42.56807	0.365132
SWINIR [28]	0.89 M	**24.865050**	**0.645047**	0.1135	94.86809	0.427522
SRGAN [30]	25.12 M	20.959648	0.444068	0.1660	38.37431	0.369272
CycleGAN [49]	89.2 M	21.052766	0.418084	0.1674	71.22200	0.401133
pix2pixHD [34]	64.13 M	20.974240	0.271080	0.1345	158.3319	0.446123
DMSC-GAN	50.08 M	24.379220	0.632996	**0.0905**	**24.52293**	**0.304919**

Note: Bold and red font in the table indicate optimal values for each indicator.

Fig.4

Based on the compared methods, DMSC-GAN achieves the second-highest scores in terms of PSNR and SSIM, slightly below SWINIR. However, it surpasses all other methods in the remaining three metrics. The FID score reached 24.52293, and the LPIPS score was 0.304919. These results indicate a significant improvement of 13.85138 and 0.064353, respectively, compared to the second-best SRGAN. Additionally, DMSC-GAN outperforms both supervised SRGAN and unsupervised CycleGAN in all five metrics. Compared with pix2pixHD, DMSC-GAN exhibits significant improvements, including a substantial PSNR increase of 3.40498 dB, an SSIM improvement of 0.361916, a reduced MSE of 0.044, a substantial FID improvement of 133.80897, and a decreased LPIPS by 0.141204. Notably, DMSC-GAN excels in improving the FID and LPIPS scores compared to non-GAN models while demonstrating overall enhancements across all five metrics compared to GAN models.

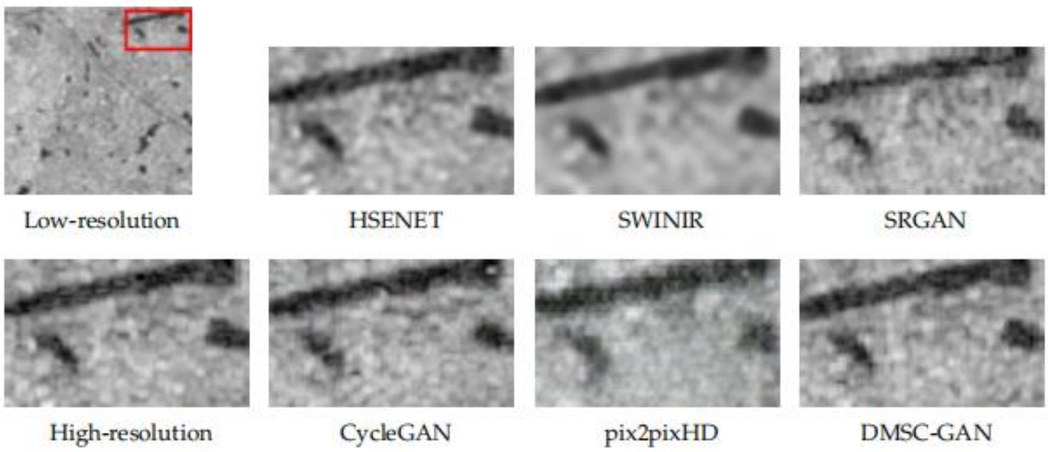


Fig.5

Figure 5 displays the generated images from various models. The images demonstrate that the results produced by DMSC-GAN closely resemble the original high-resolution SAR images, significantly improving the quality of the generated images.

CONCLUSION- In this study, DMSC-GAN is proposed for SAR image super-resolution. The generator module follows an encoder–decoder architecture, incorporating convolutional operations and DMSA to extract informative features. A multi-scale feature extraction pyramid is created using varying window sizes to capture features at different scales. The discriminator is enhanced by employing two discriminators to assess inputs at different scales to enhance discriminative capability. Perceptual loss and feature matching loss are introduced to provide more comprehensive feedback to the generator from the discriminator. The superiority of DMSC-GAN in terms of performance metrics compared to the SRGAN method is demonstrated through the analysis of research experiment results. Specifically, a significant improvement of 3.4196 dB in PSNR and 0.1889 in SSIM on the SEN1-2 dataset was observed with DMSC-GAN. Additionally, there was a substantial increase of 13.85 in the FID score.

REFERENCES-

1. Yang, W.; Zhang, X.; Tian, Y.; Wang, W.; Xue, J.H.; Liao, Q. Deep Learning for Single Image Supe Resolution: A Brief Review. *Trans. Multimed.* **2019**, *21*, 3106–3121.

2. Tong, C.; Leung, K. Super-resolution reconstruction based on linear interpolation of wavelet coefficients. *Multidimens. Syst. Signal Process.* **2007**, *18*, 153–171.

3. Liu, J.; Gan, Z.; Zhu, X. Directional bicubic interpolation—A new method of image super-resolution. In Proceedings of the 3rd International Conference on Multimedia Technology (ICMT-13), Guangzhou, China, 29 November–1 December 2013 ; Atlantis Press: Amsterdam, The Netherlands, 2013; pp. 463–470

4. Chang, H.; Yeung, D.Y.; Xiong, Y. Super-resolution through neighbor embedding. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; Volume 1, p. I.

5. Tong, C.; Leung, K. Super-resolution reconstruction based on linear interpolation of wavelet coefficients. *Multidimens. Syst. Signal Process.* **2007**, *18*, 153–171

6. Liu, J.; Gan, Z.; Zhu, X. Directional bicubic interpolation—A new method of image super-resolution. In Proceedings of the 3rd International Conference on Multimedia Technology (ICMT-13), Guangzhou, China, 29 November–1 December 2013 ; Atlantis Press: Amsterdam, The Netherlands, 2013; pp. 463–470

7. Chang, H.; Yeung, D.Y.; Xiong, Y. Super-resolution through neighbor embedding. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; Volume 1, p. I.

8. Conchello, J.A. Superresolution and convergence properties of the expectation-maximization algorithm for maximum-likelihood deconvolution of incoherent images. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **1998**, *15*, 2609–2619. [

9. Zheng, W.; Deng, F.; Mo, S.; Jin, X.; Qu, Y.; Zhou, J.; Zou, R.; Shuai, J.; Xie, Z.; Long, S.; et al. Image super-resolution reconstruction algorithm based on Bayesian theory. In Proceedings of the 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, China, 31 May–2 June 2018; pp. 1934–1938.

10. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]

11. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.

12. G Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. PhotoRealistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 105–114.

13. Wang, T.C.; Liu, M.Y.; Zhu, J.Y.; Tao, A.; Kautz, J.; Catanzaro, B. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8798–8807.

14. Liu, B.; Li, H.; Zhou, Y.; Peng, Y.; Elazab, A.; Wang, C. A Super Resolution Method for Remote Sensing Images Based on Cascaded Conditional Wasserstein GANs. In Proceedings of the 2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP), Virtual, 12–15 September 2020; pp. 284–289.

15. Nasser, S.A.; Shamsi, S.; Bundele, V.; Garg, B.; Sethi, A. Perceptual cGAN for MRI Super-resolution. In Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Glasgow, UK, 11–15 July 2022; pp. 3035–3038.

16. Hanano, T.; Seo, M.; Chen, Y.W. An Improved cGAN with Self-Supervised Guidance Encoder for Generation of High-Resolution Facial Expression Images. In Proceedings of the 2023 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 6–8 January 2023; pp. 1–4.

17. Schmitt, M.; Hughes, L.H.; Zhu, X.X. The SEN1-2 Dataset for Deep Learning in SAR-Optical Data Fusion. *arXiv* **2018**, arXiv:1807.01569