



Contents lists available at ScienceDirect

# International Journal of Applied Earth Observations and Geoinformation

journal homepage: [www.elsevier.com/locate/jag](http://www.elsevier.com/locate/jag)

## A review and meta-analysis of generative adversarial networks and their applications in remote sensing

Shahab Jozdani<sup>a</sup>, Dongmei Chen<sup>a,\*</sup>, Darren Pouliot<sup>b</sup>, Brian Alan Johnson<sup>c</sup><sup>a</sup> Department of Geography and Planning, Queen's University, Kingston, ON K7L 3N6, Canada<sup>b</sup> Landscape Science and Technology Division, Environment and Climate Change Canada, 1125 Colonel By Drive, Ottawa, Ontario K1A 0H3, Canada<sup>c</sup> Natural Resources and Ecosystem Services Area, Institute for Global Environmental Strategies, 2108-1 Kamiyamaguchi, Hayama, Kanagawa 240-0115, Japan

## ARTICLE INFO

## Keywords:

Remote sensing  
Generative adversarial networks  
GANs  
Deep learning

## ABSTRACT

Generative Adversarial Networks (GANs) are one of the most creative advances in Deep Learning (DL) in recent years. The Remote Sensing (RS) community has adopted GANs quickly, and reported successful use in a wide variety of applications. Given a sharp increase in research on GANs in the field of RS, there is a need for an in-depth review of the major technological/methodological advances and new applications. In this regard, we conducted a comprehensive review and meta-analysis of GAN-related RS papers, with the goals of familiarizing the RS community with the potential of GANs and helping researchers further explore RS applications of GANs by untangling challenges common in this field. Our review is based on 231 journal papers that were retrieved and selected through the Web of Science (WoS) database. We reviewed the theories, applications, and challenges of GANs, and highlighted the gaps to explore in future studies. Through the meta-analysis conducted in this study, we observed that image classification (especially urban mapping) has been the most popular application of GANs, potentially due to the wide availability of benchmark datasets. On the other hand, we found that relatively few studies have explored the potential of GANs for analyzing medium spatial-resolution multi-spectral images (e.g., Landsat or Sentinel-2), even though such images are often freely available and useful for a wide range of applications (e.g., urban expansion analysis, vegetation mapping, etc.). In spite of the applications of GANs for different RS processing tasks, there are still several gaps/questions in this field such as: 1) which GAN models/configurations are more suitable for different applications? 2) to what degree can GANs replace real RS data in different applications? Such gaps/questions can be appropriately addressed by, for example, conducting experimental studies on evaluating different GAN models for various RS applications to provide better insights into how/which GAN models can be best deployed. The meta-analysis results presented in this study could be helpful for RS researchers to know the opportunities of using GANs and understand how GANs contribute to the current challenges in different RS applications.

### 1. Introduction

One of the most important advances in machine learning (ML) in recent years is the resurrection of artificial neural networks (ANN) in the form of deep learning (DL). DL has led to the strengthening of ANNs by increasing the number of hidden layers (depth), mitigating outstanding problems (such as exploding/vanishing gradients), improving spatial characterization with convolution operations, and drastically reducing network training time (LeCun et al., 2015). These advancements have led to a new era in a wide variety of modeling tasks, including computer vision and remote sensing (RS) image analysis. DL models, especially

deep convolutional neural networks (DCNNs), have been widely used in various RS applications, largely mapping applications where traditional ML models (mainly working with hand-crafted features) sometimes struggle to produce accurate and generalizable results (Jozdani et al., 2019). In recent years, there have been different advances in the field of DL, ranging from new architectures and training approaches to new forms of data processing (such as Vision Transformers (Dosovitskiy et al., 2020), DL-based semi-supervised learning (Chen et al., 2020), and few/zero-shot learning (Wang et al., 2019). Among these advances, Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) are undoubtedly one of the most innovative. Although there were previous

\* Corresponding author.

E-mail address: [chendm@queensu.ca](mailto:chendm@queensu.ca) (D. Chen).<https://doi.org/10.1016/j.jag.2022.102734>

Received 6 December 2021; Received in revised form 12 February 2022; Accepted 24 February 2022

Available online 17 March 2022

0303-2434/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

approaches based on the generative model concepts (most notably, Variational Autoencoders (VAE)), GANs boast some unique properties that help generate data with unprecedented fidelity or improve network performance for different downstream tasks.

GANs are a class of DL models capable of generating fake data, including image/raster data (Goodfellow et al., 2014). As any other generative model, GANs are trained with some real data, which is used to learn important features that can then help synthesize data that looks like real data. The way GANs work can be considered analogous to counterfeiters and detectives; that is, counterfeiters aim to generate some fake product/information that is indistinguishable from real product/information. More technically, GANs employ two models (e.g., neural networks): a generator (responsible for generating fake data), and a discriminator (responsible for distinguishing fake data from real data) (Goodfellow et al., 2014; Isola et al., 2016). These two models are simultaneously trained in an adversarial manner. Intuitively, the goal of the generator is to generate fake data such that it is very similar to real data, causing the discriminator to fail to distinguish it from a real data. Because of the relatively unique characteristic of GANs in producing synthetic data (usually imagery in RS studies), they have been used for a wide variety of RS applications including image-to-image translation (e.g., converting SAR images to optical images) (Fuentes Reyes et al., 2019; Chen et al., 2021), image classification (i.e., semantic segmentation, object-detection, scene classification) (Han et al., 2020; Xiong et al., 2020; Zhu et al., 2020), and super-resolution (to synthetically upscale imagery to a higher resolution) (Jiang et al., 2019; Salgueiro Romero et al., 2020).

Although there have been several review papers on DL in the context of RS in recent years, to our knowledge, those studies have been on the broad topic of DL rather than explicitly focusing on GANs (Zhang et al., 2016; Yuan et al., 2020; Ma et al., 2019; Kattenborn et al., 2021). A comprehensive review study on GANs exclusively in the context of RS, however, would help the RS community to understand the potential and limitations of GANs in this field. This is important as GANs have shown to be powerful models in different computer vision tasks (Creswell et al., 2018; Gui et al., 2020; Shamsolmoali et al., 2021; Alqahtani et al., 2021), which are also common in RS. Thus, a review on this topic could also help RS researchers to identify RS data sources/applications/environmental contexts in which GANs should be explored sufficiently as well as other gaps and trends in the applications of GANs in RS. Systematic reviews are particularly helpful for elucidating this information, as they are backed by a transparent and replicable methodology (e.g., explicit literature search and appraisal approach) that allows for quantitative meta-analysis (i.e., techniques that statistically combine the results of multiple primary studies) (Grant and Booth, 2009). There are several seminal systematic review and meta-analysis papers in the field of RS focusing on different applications including image segmentation (Johnson and Jozdani, 2018; Kotaridis and Lazaridou, 2021), DL (Ma et al., 2019; Heydari and Mountrakis, 2019), and conventional image classification (Khatami et al., 2016). The most relevant review paper on GANs that also briefly reviews some applications of GANs in RS is the recent publication by Dash, et al. (Dash et al., 2021). However, there is a lack of a comprehensive review and meta-analysis study on the applications of GANs in RS, despite GANs wide use for RS analysis in recent years.

These are important information that are missing in the context of RS. Given the growing importance of GANs in the RS community, and the lack of prior systematic reviews on this topic, in this study we carried out a systematic review on the use of GANs in the context of RS. In addition to providing details on the papers reviewed, we also performed a meta-analysis to capture the status of trends in the applications of GANs in RS that is missing in the literature. The main objectives of this review were to:

(1) Identify the main applications of GANs in the field of RS,

- (2) Identify and categorize the various GAN-based algorithms/methods that have been used for different RS applications,
- (3) Conduct a comprehensive meta-analysis to capture the trend of GANs applications in the RS community,
- (4) Highlight the main challenges and gaps in training and applying GANs to RS data.

## 2. Review process

To retrieve the papers for our systematic review, we conducted a title/keyword/abstract search in the Web of Science (WoS) database using the terms: "remote sensing" AND "generative adversarial network". We downloaded the metadata for all journal papers identified through this search (assuming they were the most reliable type of documents due to their typically more rigorous peer-review process). The search was performed on July 17, 2021, resulting in the retrieval of 289 journal papers in total. Of these, 231 papers were deemed as relevant for our review, namely that papers that addressed RS image-related problems (Fig. 1). Papers only involving the use of non-RS data (i.e., non-overhead images) and/or non-image data were excluded from our reviews. From the full texts of the remaining 231 relevant papers, we collected several types of information deemed useful for a systematic review, including: the RS application(s) for which GANs were applied, the type(s) and spatial resolution(s) of the RS data used, the type of study area, the details of the GANs used in the study (loss function and evaluation metrics), the country of the corresponding author, and year of publication (Table 1).

## 3. Theoretical background of generative adversarial networks (GANs)

### 3.1. Vanilla GANs

Vanilla GANs were introduced by Goodfellow et al. (2014) based on a multi-layer perceptron (MLP) network. As a class of deep generative models, GANs consist of two components: a discriminator ( $D$ ), and a generator ( $G$ ). These two components can be considered as any differentiable system, but neural networks are commonly used as the generator and discriminator. The role of the generator is to approximate a distribution  $p_g(x)$  (which is a probability density function with a random variable  $x$  (e.g., pixel values)) that resembles that of real data  $p_{data}(x)$ . When such a distribution can be approximated, it is possible to draw

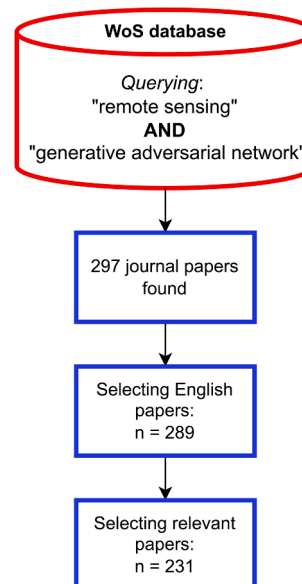


Fig. 1. Process of paper selection in this study.

**Table 1**  
Attributes recorded from the papers reviewed.

Attributes	Categories
Study application	<ul style="list-style-type: none"> <li>• Image classification:               <ul style="list-style-type: none"> <li>○ Semantic segmentation</li> <li>○ Object detection</li> <li>○ Scene classification</li> </ul> </li> <li>• Image reconstruction/restoration               <ul style="list-style-type: none"> <li>○ Super-resolution</li> <li>○ Image denoising</li> <li>○ Image inpainting (interpolation)</li> </ul> </li> <li>• Data translation:               <ul style="list-style-type: none"> <li>○ Image-to-image translation</li> <li>○ Text-to-image translation</li> </ul> </li> <li>• Domain Adaptation</li> <li>• Miscellaneous               <ul style="list-style-type: none"> <li>○ Image generation</li> <li>○ Image retrieval</li> <li>○ Image compression</li> <li>○ Image matching</li> <li>○ Image fusion</li> </ul> </li> </ul>
RS data used	Visible (RGB), multi-spectral, hyperspectral, SAR, LiDAR
Spatial resolution	–
Study area	Urban, forest, agriculture, other (e.g., mountains, water, airport, etc.)
Loss function	Table 2
Evaluation metrics	Table 3
Country of the corresponding author	–
Year of publication	–

samples from it that would look like real ones. To mathematically define the objective of a GAN, let  $G(z; \theta_g)$  be a neural network used as the generator where  $z$  is random noise taken from a prior distribution  $z \sim p_z$ , and  $\theta_g$  are network parameters. The goal of the generator is to perform the mapping from a given noise vector to the real data ( $G: z \rightarrow x$ ). This is equivalent to training a network that takes as input a random noise vector to simulate, for example, RS images. The discriminator network  $D(x; \theta_d)$ , parameterized by  $\theta_d$ , aims to distinguish fake data (generated by the generator) from the real data  $x \sim p_{data}$ , whereas the generator tries to maximize the probability of detecting fake samples as real ones. After each iteration of training, if the quality of fake data is insufficient (meaning that the discriminator can easily detect it as fake data), the generator receives some negative feedback/signal helping it improve the generating process. This process is repeated until the generator's fake images cannot be easily differentiated from the real ones by the discriminator. In essence, these two networks compete in an adversarial way through a two-player minimax game as presented in Equation (1).

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

The discriminator of the vanilla GANs uses a binary cross-entropy classifier ( $D(x) \in [0, 1]$ ). In the first paper on GANs, it was shown that the best discriminator for a fixed generator is the one converging to Equation (2).

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)}. \quad (2)$$

Plugging the unique optimal discriminator  $D_G^*(x)$  (Equation (2)) into the objective function of GANs (Equation (1)) results in the Jensen-Shannon (JS) divergence between  $p_g(x)$  and  $p_{data}(x)$  (Equation (3)). Since the goal of the generator is to approximate a distribution that matches the real distribution, the optimal generator minimizing the JSD ( $p_{data}(x) \parallel p_g(x)$ ) is  $p_g(x) = p_{data}(x)$ , leading to the optimal discriminator outputting a probability of  $\frac{1}{2}$ , implying that the discriminator will be uncertain if the sample is fake or real.

$$\begin{aligned} \max_D V(D, G) &= \mathbb{E}_{x \sim p_{data}(x)} [\log D_G^*(x)] + \mathbb{E}_{x \sim p_g(x)} [1 - \log D_G^*(G(x))] = \\ &= \mathbb{E}_{x \sim p_{data}} \left[ \underbrace{\log \frac{p_{data}(x)}{\frac{1}{2}(p_{data}(x) + p_g(x))}}_{\text{KL divergence}} \right] + \mathbb{E}_{x \sim p_g} \left[ \underbrace{\log \frac{p_g(x)}{\frac{1}{2}(p_{data}(x) + p_g(x))}}_{\text{KL divergence}} \right] - 2 \log 2 \\ &= 2 \text{JS}(p_{data} \parallel p_g) - 2 \log 2. \end{aligned} \quad (3)$$

### 3.2. GAN variants (conditional, semi-supervised, auxiliary classifier)

Vanilla GANs generate data based on random noise vectors without relying on any other information. One of the most important downsides of vanilla GANs is that we do not have control over the generator, and the samples generated accordingly. To address this problem, conditional GANs (cGANs) were introduced (Mirza and Osindero, 2014). Conditioning the generator as well as the discriminator using some additional information (such as class labels) can help improve/manipulate the fake data generated. In this regard, we can define the objective function of cGANs very much like that of vanilla GANs by including additional information  $y \sim p_y$  in both the discriminator and generator as presented in Equation (4).

$$\min_G \max_D V(G, D) = \min_G \max_D \mathbb{E}_{x \sim p_{data}(x)} [\log D(x|y)] + \mathbb{E}_{z \sim p_z(z)} [1 - \log D(G(z|y))]. \quad (4)$$

This aids the generator in developing fake images for a specific class. Rather than class labels, one can also condition a GAN using images. For example, in super-resolution tasks, a coarse-resolution image is input to the generator to produce the corresponding high-resolution image, and the discriminator determines if this is an enhanced (fake) image or a native high-resolution (real) image. Other applications in which GANs can be conditioned using images include image interpolation (i.e., inpainting), domain adaptation, and image-to-image translation. For these image translation types of applications, the generator is the main interest, and the discriminator network is used to improve the training of the generator. However, the opposite can be true for some classification tasks, where the discriminator network is of primary interest and the generator is used to improve the training of the discriminator network.

Different cGAN approaches have been proposed in recent years. Some of those approaches need paired data. One of the most well-known cGAN approaches is Pix2Pix (Isola et al., 2016) (and its improved version Pix2Pix HD (Wang et al., 2017)). The authors of Pix2Pix showed that without devising new loss functions or architectures, it would be possible to perform several tasks (such as map/image-to-image/map translation, image inpainting, image colorization). However, to improve training GANs for different applications, it is a common practice to include other loss functions to the objective function. A list of the most common loss functions used in GANs approaches is provided in Table 2. Although Pix2Pix was a breakthrough in image translation tasks, it was only applicable to paired data. In many tasks, collecting paired data is very difficult, if not impossible. To expand the applications of cGANs beyond paired data, Zhu et al. (2017) proposed CycleGAN, which is able to operate on unpaired data. In this regard, CycleGAN performs two mappings  $G: X \rightarrow Y$  and  $F: Y \rightarrow X$  and uses two corresponding discriminators  $D_X$  and  $D_Y$ . Along with the adversarial loss, the authors proposed a new loss function called cycle-consistency loss to make these two mappings possible (Equation (5)).

$$\begin{aligned} G^*, F^* &= \min_{G, F} \max_{D_X, D_Y} \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) + \lambda \mathcal{L}_{\text{cycle}}(G, F), \\ \mathcal{L}_{\text{cycle}}(G, F) &= \mathbb{E}_{x \sim p_{data}(x)} [\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{data}(y)} [\|G(F(y)) - y\|_1]. \end{aligned} \quad (5)$$

Another commonly used type of cGANs is the semi-supervised GAN

**Table 2**  
Most common loss functions found in GAN studies in RS.

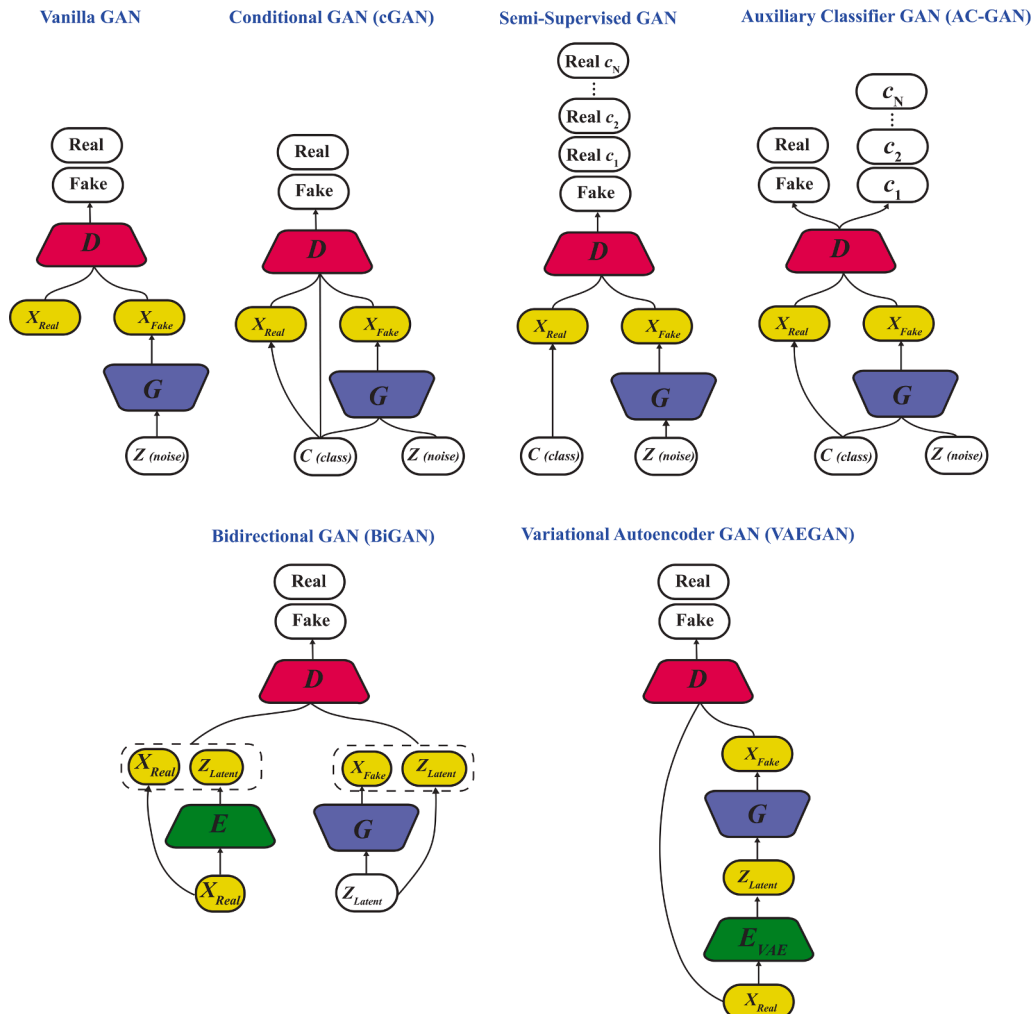
Loss	Description	Notes/references
Adversarial	This is imposed by the discriminator and most commonly is chosen to be (binary) cross-entropy loss or least-square loss in the context of RS.	Vanilla GAN (Goodfellow et al., 2014) WGAN (Arjovsky et al., 2017) WGAN-GP (Gulrajani et al., 2017) LSGAN (Mao et al., 2016)
Pixel-space	This is a measure of structural similarity that is commonly calculated based on L1 or L2 norm depending on the application at hand.	L1 norm <sup>1</sup> L2 norm (or squared L2 norm also known as mean-squared loss) <sup>2</sup>
Perceptual	This is a measure of visual perception that is calculated based on a well-trained model, most commonly VGG-16 or VGG-19.	L1/2 norm based on pretrained VGG-16/19 (Simonyan and Zisserman, 2014)
Classifier	This loss is used to include semantic information as well as other information for training GANs.	Cross-entropy (Good, 1952)
Total variation	This loss is popular in style-transfer tasks as it can help reduce noise in the generated images. However, it is also widely used for super-resolution tasks.	(Rudin et al., 1992)
Cycle-consistency	This is a measure of self-similarity used during unpaired image-to-image translation tasks.	(Zhu et al., 2017)
Saliency	This loss measures discrepancies of saliency maps of real and generated images.	(Zhai and Shah, 2006)
Edge	This loss calculates discrepancies between the edges of real and generated images.	Different edge detection approaches can be used to detect edges and then calculate this loss (Gonzalez and Woods, 2006).

$$^1 \sum_{i=1}^n |G(x) - x|.$$

$$^2 \sum_{i=1}^n (G(x) - x)^2 \text{ or } \sqrt{\sum_{i=1}^n (G(x) - x)^2}.$$

(Odena, 2016) which is an improved variant of vanilla GANs that aims to enhance supervised classification by involving unsupervised learning to learn from unlabeled data as well. As shown in Fig. 2, the discriminator of a semi-supervised GAN predicts N + 1 classes where “N” is the

number of classes in the data, and “1” is for an additional class “fake”; that is, if the discriminator detects a generated data as real, it not only labels it as “real”, but also assigns it a class label. Auxiliary classifier GAN (AC-GAN) (Odena et al., 2016) is another class of popular cGANs.



**Fig. 2.** Some popular unconditional and conditional GAN frameworks.

Similar to cGANs, AC-GAN has a class-conditional generator. However, unlike cGANs, AC-GAN's discriminator does not have access to the class labels. In fact, the way AC-GAN works is based on outputting two probabilities: 1) the probability of fake/real; and 2) the probability of classes. As with GANs, VAEs are generative models that can synthesize high-quality images. VAEs generally tend to generate blurry images because of the way the loss function of VAEs work; that is, although in many cases data have multi-modal distributions, the L2 loss function assumes a single Gaussian distribution for the data. However, Larsen et al. (2015) found that integrating a VAE and GAN (VAEGAN) can improve the quality of images generated. In this regard, the input is first encoded to a latent space (using the encoder of VAE). Then, the decoder (which is actually the generator) takes the encoded latent vector and generates fake images using it. Finally, just as with other types of GANs, a discriminator is used to distinguish the fake and real images. Bidirectional GAN (BiGAN) (Donahue et al., 2016) is another GAN model that includes latent-space features along with data-space features in the process of training. With this model, not only does the GAN map latent space to fake images, but it also inversely maps real data to latent space using an encoder. Due to the bi-directional design of BiGAN, in addition to distinguishing real from fake data (in the data space), the discriminator jointly distinguishes the tuple  $(x_{Real}, E(x_{Real}))$  from  $(G(z), z)$ . The authors of BiGAN reported the high efficacy of this unsupervised feature learning approach for different supervised and semi-supervised downstream tasks.

Because of the high-dimensional parameter space and use of a non-convex objective function, finding an optimal solution (i.e., Nash equilibrium) when training GANs can be difficult (Salimans et al., 2016). Due to this, GANs can suffer from instability (such as a phenomenon known as "mode collapse" (Salimans et al., 2016) that may not be easily remedied. To overcome this issue, several approaches have been proposed to date. These approaches primarily focus on improving the objective function to enhance stability, and/or on the use of specialized architectures (e.g., DCNNs). Mao et al. (2016) proposed a least-squares GAN (LSGAN) that replaces the sigmoid cross-entropy with the least-squares loss function, so as to mitigate the vanishing-gradient problem

during training. This was shown to result in generating higher-quality images in their experiments. Wasserstein GANs (WGANs) (Arjovsky et al., 2017) are another well-known approach for handling the vanishing-gradient problem in the training of GANs. To reduce training instability and mode collapse, WGAN uses an objective function which is based on an approximation of the Wasserstein distance (known also as Earth-mover) between  $p_{data}(x)$  and  $p_g(x)$ . Due to some limitations of training WGANs (like the K-Lipschitz constraint), WGAN-GP was proposed to account for such restrictions through including a gradient penalty term in the objective function of WGAN.

### 3.3. Evaluation metrics

Apart from the instability of GANs during training, it is difficult to measure the quality of generated images in an objective manner while training. This is mainly due to the fact that the loss values of GANs may not be indicative of the performance of the model during training. This problem in WGAN is less severe but even that alone could not be a reliable measure. To more objectively and efficiently measure the quality of generated images by GANs, several metrics have been proposed in the literature, which can be categorized into two main groups: 1) metrics that are evaluated against ground-truth data (supervised metrics), and 2) metrics that do not require ground-truth data (unsupervised metrics). The metrics requiring ground-truth data also include task-specific metrics that indirectly evaluate generated data by considering them in downstream tasks such as scene classification, object detection, or semantic segmentation. Such task-specific metrics include well-known metrics such as overall/average accuracy (OA/AA), F1-score (the harmonic mean of Precision and Recall metrics), intersection of union (IoU/mIoU), and Kappa coefficient, etc. In Table 3, the most common metrics that have been used in the context of RS for evaluating GANs can be seen.

## 4. Meta-analysis of GANs in remote sensing

In this section, we present the results of a meta-analysis conducted

**Table 3**  
Common evaluation metrics of GANs in RS studies.

Metrics	Supervised	Examples
RMSE/MSE/MAE	✓	(Zhang et al., 2018; Zheng et al., 2018; Leinonen et al., 2019; Ma et al., 2019; Zhang et al., 2020; Chen et al., 2021; Feng et al., 2021; Song et al., 2021; Zhang et al., 2021)
Structural similarity index (SSIM)	✓	(Salgueiro Romero et al., 2020; Zhang et al., 2018; Zheng et al., 2018; Leinonen et al., 2019; Ma et al., 2019; Zhang et al., 2020; Chen et al., 2021; Feng et al., 2021; Song et al., 2021; Zhang et al., 2021; Wang et al., 2018; Shi et al., 2019; Dou et al., 2020; Ma et al., 2020; Wang et al., 2020; Xiong et al., 2020; Fang et al., 2021; Zhang et al., 2021)
Peak signal-to-noise ratio (PSNR)	✓	(Chen et al., 2021; Zhang et al., 2018; Feng et al., 2021; Wang et al., 2018; Huang and Jing, 2020; Li et al., 2020; Yu et al., 2020; Yue et al., 2020; Zhang et al., 2020; Bashir and Wang, 2021)
Variance inflation factor (VIF)	✓	(Zhang et al., 2018; Zhao et al., 2021)
Spectral angle mapper (SAM)	✓	(Zhang et al., 2020; Zhang et al., 2021; Zhang, 2019; Gao et al., 2020; Li et al., 2020; Zhou et al., 2020; Ozcelik et al., 2021; Xie et al., 2021; Zhou et al., 2021)
Erreur relative globale adimensionnelle de synthèse (ERGAS)	✓	(Salgueiro Romero et al., 2020; Ma et al., 2019; Zhang et al., 2020; Zhang et al., 2021; Zhou et al., 2020; Zhou et al., 2021; Tang et al., 2020; Chen et al., 2021; Zhang et al., 2021)
Correlation coefficient (CC)	✓	(Zhang et al., 2020; Ozcelik et al., 2021; Zhang et al., 2021; Bittner et al., 2018; Ghamisi and Yokoya, 2018; Hu et al., 2020; Kim et al., 2020; Paoletti et al., 2021)
Fréchet inception distance (FID)	✓	(Zhu et al., 2020; Li et al., 2020; Adamiak et al., 2021; Ji et al., 2021; Shamsolmoali et al., 2021)
Inception score (IS)	×	(Zhu et al., 2020; Li et al., 2020; Wei et al., 2020; Chen et al., 2021)
Perceptual image (PI)	×	(Feng et al., 2021; Lei et al., 2020; Gong et al., 2021)
Learned perceptual image patch similarity (LPIPS)	✓	(Li et al., 2020; Lei et al., 2020; Gong et al., 2021)
Mutual information (MI)	✓	(Zhang et al., 2018; Wei et al., 2020)
Cosine similarity (CoS)	✓	(Ji et al., 2021; Wei et al., 2020; Xiong et al., 2020)
Quality with no reference (QNR)	×	(Zhang, 2019; Zhou et al., 2020; Zhou et al., 2021; Zhang et al., 2021)
Naturalness image quality evaluator (NIQE)	×	(Jiang et al., 2019; Wang et al., 2020; Burdziakowski, 2020; Ebel et al., 2021)
Perception based Image Quality Evaluator (PIQE)	×	(Burdziakowski, 2020; Tao and Muller, 2019; Tao and Muller, 2021)
Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)	×	(Burdziakowski, 2020; Tao and Muller, 2019; Tao and Muller, 2021)
R-squared	✓	(Zheng et al., 2018; Zhang et al., 2020; Xiong et al., 2021)
Task specific metrics used for different scene classification, object detection, and semantic segmentation image classification tasks	✓	(Hughes et al., 2018; Gao et al., 2019; He et al., 2019; Niu et al., 2019; Zhang et al., 2019; Yu et al., 2020; Liu et al., 2021; Peng et al., 2021; Saha et al., 2021; Zhang et al., 2021)

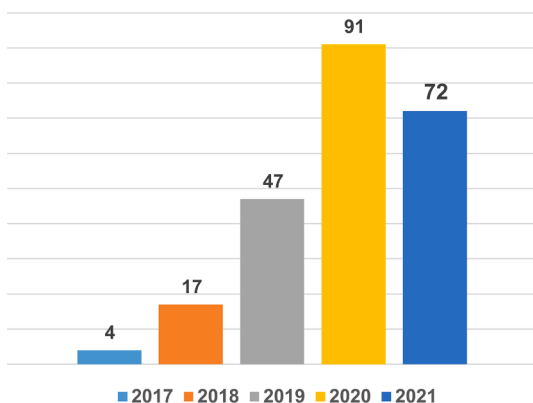


Fig. 3. Frequency of reviewed GAN papers published since 2017.

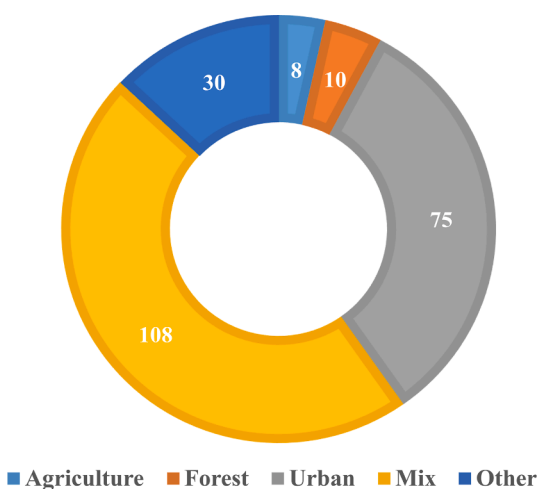


Fig. 4. Frequency of study areas in the papers reviewed.

based on the 231 papers reviewed in this study. Unsurprisingly, the number of papers on GANs in the context of RS has increased rapidly (Fig. 3). Among the papers reviewed in this study, the first journal paper on GANs in RS was published in 2017 (He et al., 2017), which is almost 3 years after GANs were first introduced. Although this 3-year delay was significant given the fast pace of advancements in the field of DL, we can

clearly see that the RS community very quickly adopted GANs for different applications since then. In this regard, in 2018, 4 times more papers on GANs were published. Compared to 2018, almost 3 times more papers were published in 2019. In 2020, 91 relevant papers were published that is almost twice the number of papers published in 2019. Finally, until June 17, 2021, we found that 72 RS-relevant papers on GANs have been published.

In the papers reviewed, we observed the applications of GANs over a wide variety of areas. As shown in Fig. 4, the majority of studies applied GANs in a mixture of different types of sites (n = 108), including, e.g. mixed urban, agricultural, forests, mountains, and water bodies. Of studies that considered a specific type of study site, the majority were studies on urban areas (n = 75), followed by studies on agricultural/rural areas (n = 30) and forests (n = 10), respectively. These results indicate that areas with high levels of human influence are the typical study sites where GANs have been applied, while natural areas (forests, wetlands, and other ecosystems) have received less research attention.

We also observed that the large majority (75%) of the papers published on GANs in the context of RS have been authored by researchers from China (Fig. 5). Authors from Germany, the United States, and a few other countries from Europe/Canada/Australia also had multiple papers published. This geographic imbalance in authorship may be a factor hindering further development of GANs for RS applications, which indicates that GANs are not being applied in many different geographic contexts. For example, there are almost no studies published by authors from tropical countries, so use of GANs for monitoring these areas is likely underdeveloped despite the importance of tropical ecosystems (e.g. tropical forests, mangroves, and coral/seagrass) and the ecosystem services they provide.

The type of RS data used for a given application is an important consideration. As shown in Fig. 6, most of the papers on GANs have used visible RGB images (e.g., UAV images, Google Earth images, etc.) followed by multi-spectral (i.e., more than three spectral bands) and hyperspectral images, respectively. The reason that RGB images were used more frequently than the other RS image sources was likely two-fold: 1) There is a wide availability of different popular benchmark datasets composed of RGB images (such as RESISC45 (Cheng et al., 2017), RSSCN7 (Zou et al., 2015), AID (Xia et al., 2017)); 2) greater availability of (free) RGB images than multi-spectral images. The choice of datasets used in the papers reviewed also affected the model developed/deployed. For example, the number of training data available in the benchmark datasets is generally more than non-benchmark datasets, so deeper networks can be applied more effectively in such cases. In addition, when using hyperspectral data, networks were designed to

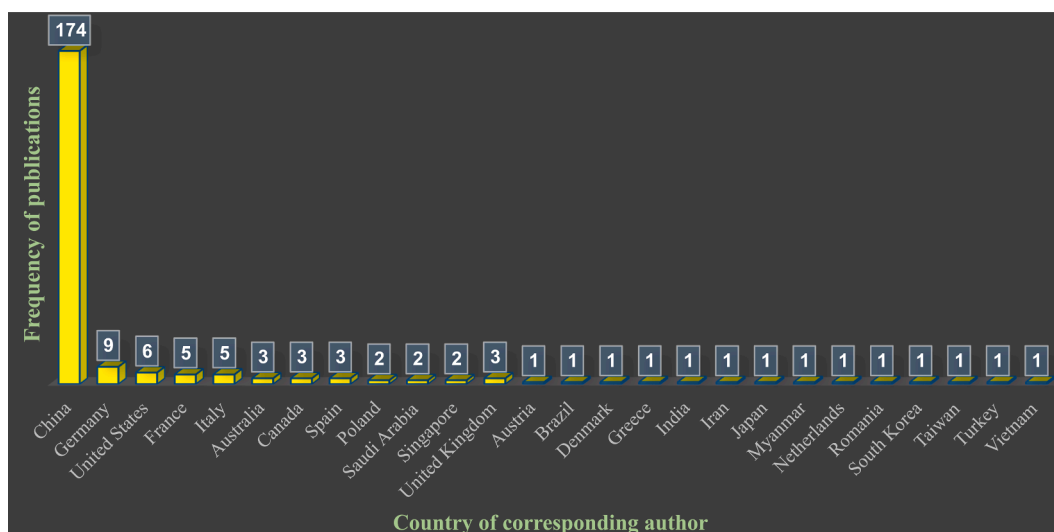


Fig. 5. Frequency of the nationalities of the corresponding authors of the papers reviewed in this study.

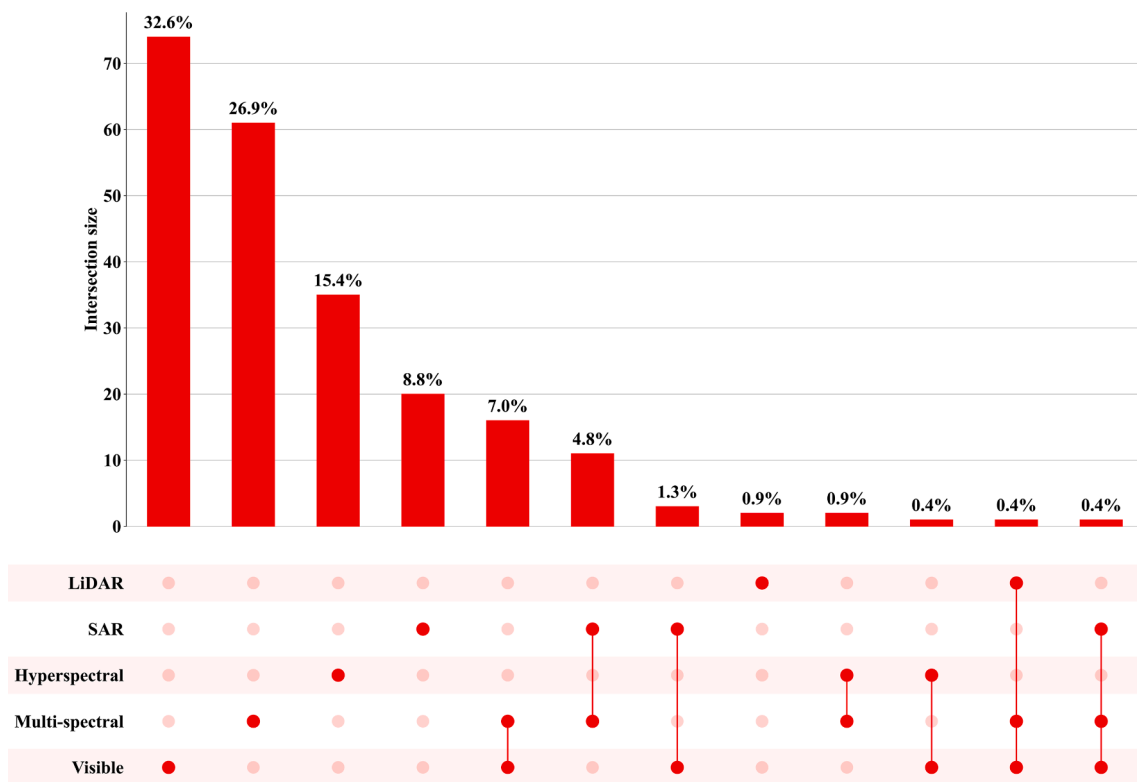


Fig. 6. Frequency of RS data used in the papers reviewed.

simultaneously handle the high-dimensionality of data (if not using any prior dimensionality reduction) and take advantage of the additional spectral information (e.g., through the use of channel-wise attention mechanisms). There were also several studies that used multiple sources of RS data for analysis, as shown in Fig. 6. Of those studies, the combination of visible and multi-spectral data was used more than the other RS data combinations. In Fig. 7, it is clear that most of the data used were very high-resolution data (spatial resolution  $\leq 1$  m). High-resolution (1–10 m) data was the second most frequently used type, followed by moderate resolution data (10–30 m) and low-resolution data ( $>30$  m), respectively. Only a small number of studies used moderate ( $n = 46$ ) and low-resolution images ( $n = 16$ ), while these types of images represent the majority of freely available RS datasets (e.g., Landsat and Sentinel data). Given this result, developing further applications of GANs for moderate/low resolution imagery could be an important area of future

research.

Among the different RS image processing tasks for which GANs were used, the most common was image classification, specifically semantic segmentation (Fig. 8). After that, data translation (especially image-to-image translation) and image reconstruction/restoration (especially super-resolution) were the second and third commonly used applications of GANs, respectively. Such results are generally in line with the applications of RS where image classification is one of the most common applications. Data translation (specifically image-to-image translation) is also popular in the context of GANs in RS as it has several applications in different fields including domain adaptation (DA) that is used in different image classification tasks, or converting different sources of RS data to each other to improve the performance of models in various downstream tasks or to improve the interpretability of images (e.g., translating SAR images to optical images).

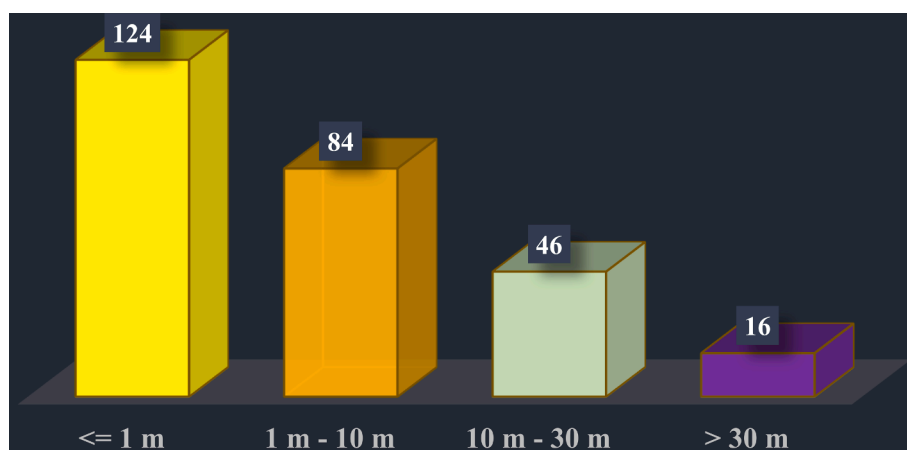


Fig. 7. Distribution of spatial resolutions of the images used in the papers reviewed. Note: Since many papers used multiple data sets, this figure is based on all the data sets used in each paper provided that their spatial resolutions were stated.

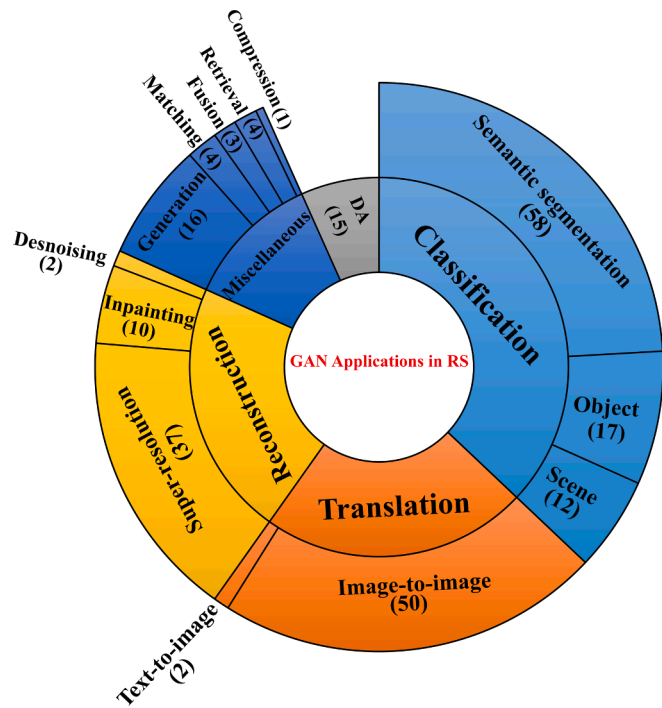


Fig. 8. Frequency of applications of GANs in RS.

We also analyzed the loss functions used for training GANs applied in the papers reviewed. According to Fig. 9, the pixel-based loss has been the most common loss term used in conjunction with the adversarial loss. It was also observed that the combination of the pixel-space loss and perceptual loss was very popular in the GANs trained in the reviewed papers. As shown in Fig. 9, it is also obvious that binary/multi-class classifier loss functions (i.e., for image classification tasks) were also popular in the papers reviewed.

In terms of the frequency of evaluation metrics, we observed that the overall accuracy (OA) metric (i.e., classification accuracy) has been used more than any other metric in the papers reviewed (Fig. 10). Specific

image-quality evaluation metrics like SSIM and PSNR have been the second and third most commonly used metrics. The common use of classification-based evaluation metrics in GAN papers reviewed is not surprising as image classification was found to be the most popular application of GANs in the RS community (Fig. 8). In addition, it is generally easier to interpret the quality of results based on such metrics when the goal is image classification.

### 5. Applications of GANs in remote sensing

GANs have been widely used in various RS applications (Fig. 11). In this section, we review the applications of GANs in RS based on the 231 journal papers considered in this study. To better organize the applications, we also propose a taxonomy of different applications of GANs in the field of RS. It should be, however, noted that some applications have overlap with each other, but we attempted to categorize the papers (and the subsequent reviews) based on the central problem targeted to be solved.

#### 5.1. Image reconstruction/restoration

##### 5.1.1. Super-resolution

Affording high-resolution multi-spectral RS imagery can be difficult and expensive. This problem is exacerbated if the aim is to conduct time-series analysis using high-resolution RS data. Super-resolution methods are one of the ways of synthetically improving the resolution of RS imagery. Advances in developing DL algorithms (especially DCNNs) have brought about unprecedented opportunities for developing more versatile super-resolution methods. Along with CNN-based super-resolution methods, GANs have proven to generate state-of-the-art super-resolution results. In recent years, GANs have also been used for different super-resolution tasks in RS (Zhang et al., 2020; Shi et al., 2019; Huang and Jing, 2020; Bashir and Wang, 2021; Ge et al., 2021). Methods used for this application can be categorized into single-image super-resolution (SISRR) and multi-image/frame super-resolution (MISRR). In RS applications, most of super-resolution applications are based on SISRR. However, since temporal analysis in RS is a common application, there are also few studies employing MISRR.

In a study conducted by Ma et al. (2019), dense residual WGAN-GP

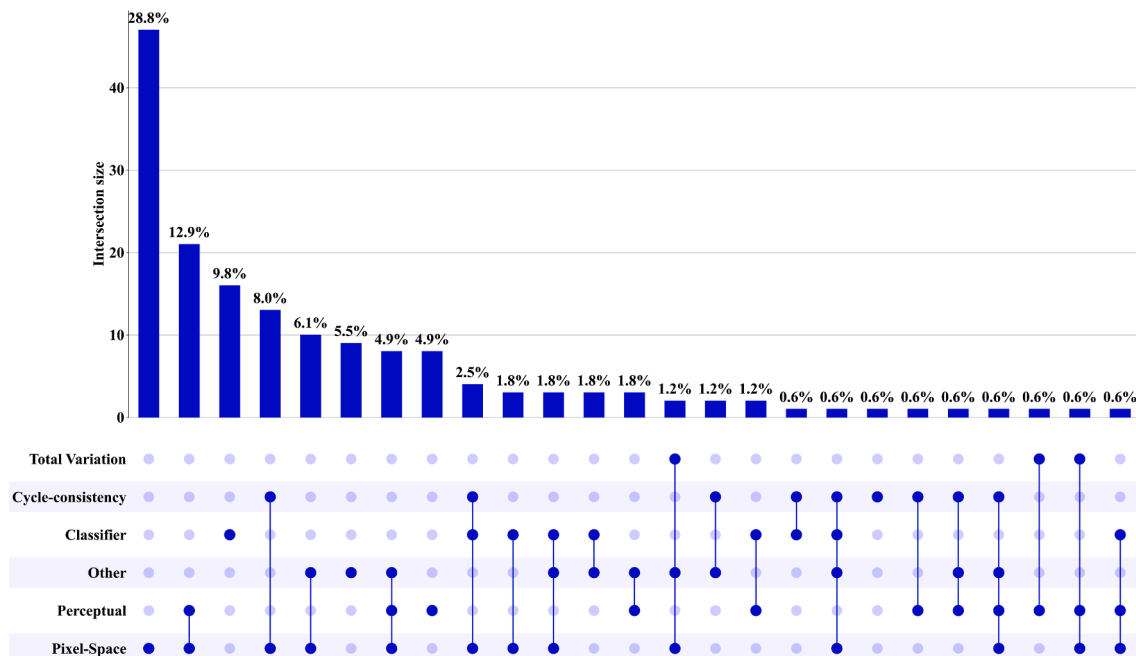


Fig. 9. Frequency of loss functions used in conjunction with adversarial loss in the papers reviewed.



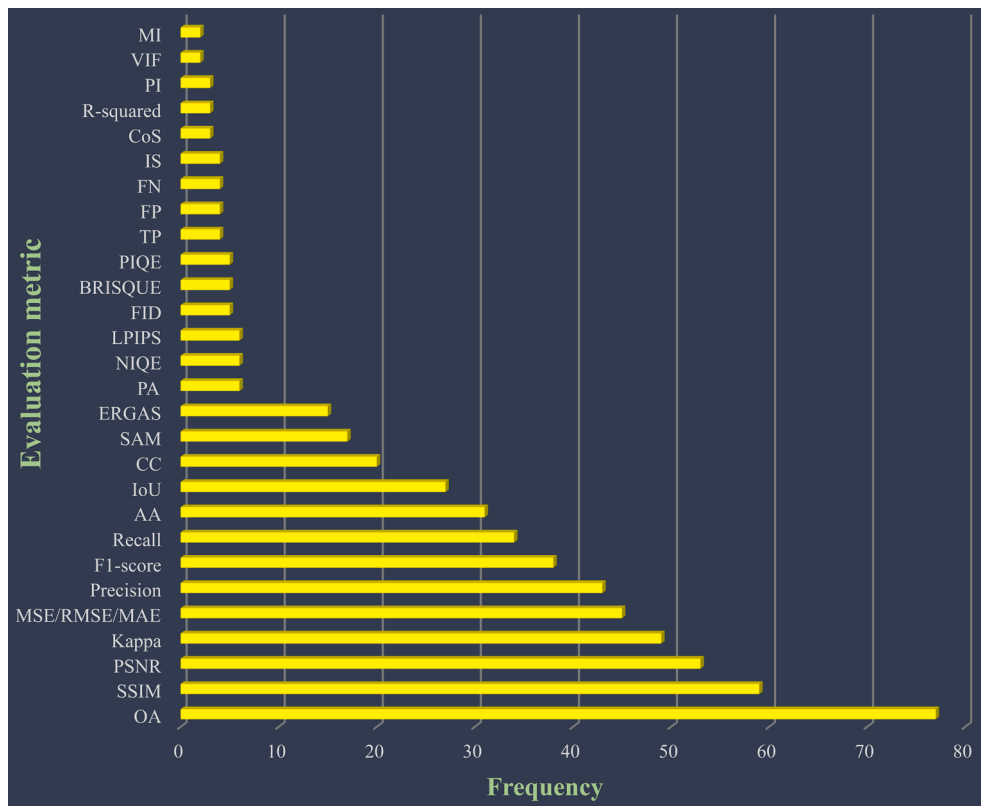


Fig. 10. Frequency of the metrics commonly used to evaluate GANs in the papers reviewed.

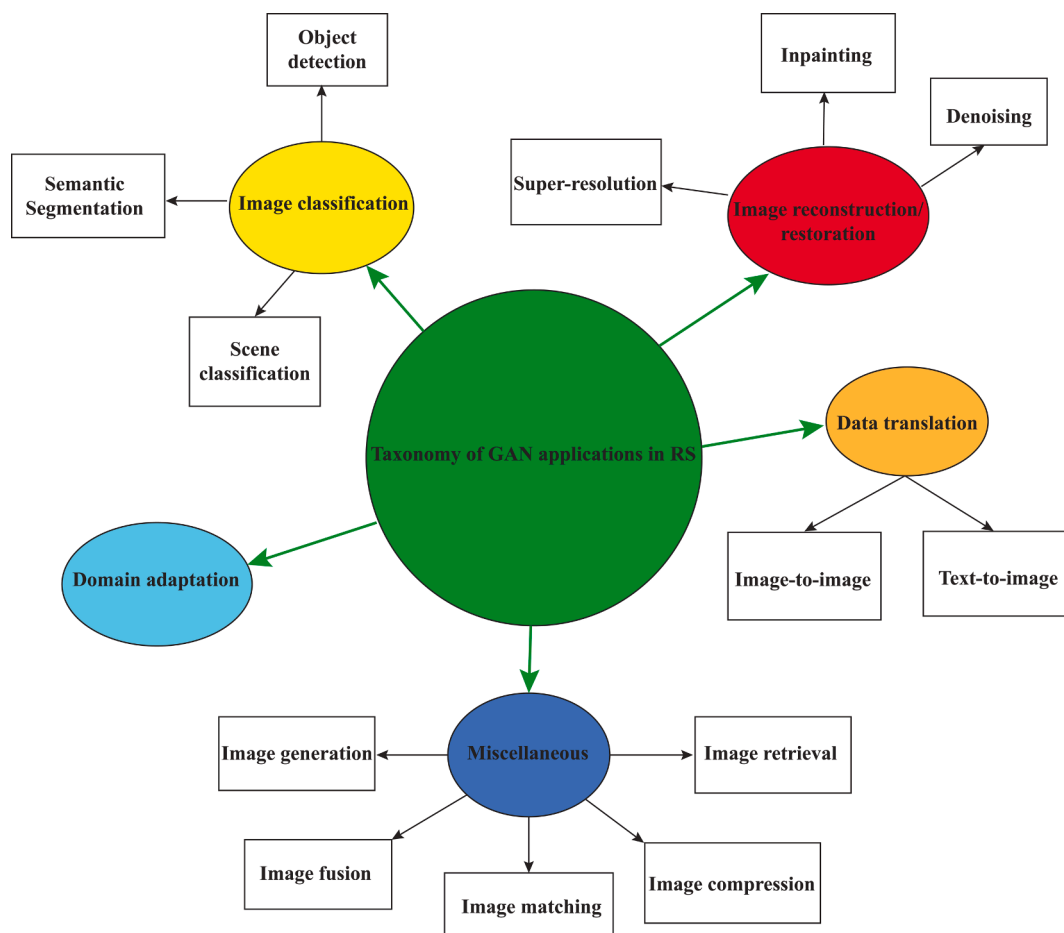


Fig. 11. Proposed taxonomy of GAN applications in RS.

(DRGAN) was proposed to perform super-resolution. In order to exert the potential of the residual blocks used in DRGAN, the authors used contiguous memory mechanisms in the network. Their experiments showed that DRGAN was superior to other approaches like SRGAN. In some studies, super-resolution has been performed to enhance images for a task-specific application, especially those in which targets are small, and thus discriminative features may not be extracted properly (Shi et al., 2019; Rabbi et al., 2020). In Shi et al. (2019), SRGAN was used to enhance low-resolution RADAR images for target recognition tasks. In Rabbi et al. (2020), an end-to-end edge-enhanced GAN-based super-resolution and object detection was proposed to detect vehicles in RS imagery. Yue et al. (2020) proposed a super-resolution method based on classifier-based GANs (called CSGAN). The authors hypothesized that considering the confidence scores of the classification when optimizing GANs for super-resolution can improve the results. Their experiments showed that the proposed GAN-based super-resolution framework resulted in higher quality results compared to SRGAN (Ledig et al., 2016) and enhanced SRGAN (ESRGAN) frameworks (Wang et al., 2018). A potentially significant application of super-resolution in the context of RS is to improve the resolution of publicly available RS imagery based on commercial high-resolution images. For example, in Salgueiro Romero et al. (2020), the authors used a modified version (removing upsampling layers) of ESRGAN (called RS-ESRGAN) to improve the resolution of Sentinel-2 imagery from 10 m to 2 m based on reference WorldView (WV) images as the high resolution reference. The proposed framework was first trained with synthetic low-resolution and high-resolution (LR-HR) WV images and then with real LR-HR pairs of Sentinel-2 and WV images. In a study by Zhang et al. (2021), the authors proposed a GAN-based super-resolution approach (called MS-SRGAN) to enhance the resolution of 16-m Gaofen wide-field-of-view imagery to 4 m based on 4-m Gaofen-2 reference images. The generator of the proposed MS-SRGAN comprised a residual squeeze-excitation (RSE)

block to improve feature extraction and super-resolved images accordingly. The authors reported that MS-SRGAN overall performed better than well-known super-resolution approaches including ESRGAN. Jiang et al. (2019) proposed a GAN-based super-resolution technique featuring an edge-enhancement network (called EEGAN) to generate sharper, more realistic super-resolution results. Along with using a sub-network that generates sharp images like other approaches, EEGAN used an edge-enhancement subnetwork to improve the quality of the first sub-network outputs, resulting in higher quality images than SRGAN and SRCNN. Variants of SRGAN were also applied to super-resolution of hyperspectral imagery. Dou et al. (2020) proposed a 3D attention-SRGAN network (3DASRGAN) to improve the resolution of hyperspectral imagery. The motivation behind the proposed network was that many super-resolution methods mainly consider spatial information during the super-resolution process. Since hyperspectral images are spectrally rich, not properly making use of spectral information can degrade super-resolution. Given this, along with spatial and adversarial losses, 3DASRGAN used the spectral angle mapper (SAM) as a spectral loss in its objective function to improve the learning of high-fidelity spectral information during the super-resolution process. In a study conducted by Yu et al. (2020), the authors proposed a GAN-based super-resolution approach whose generator was based on deep back-projection network (DBPN (Haris et al., 2018), which was earlier proposed as a super-resolution network (but not in a GAN configuration). The authors enhanced DBPN (E-DBPN) by adding a residual channel attention module and replacing the concatenation operation with a proposed feature fusion module. One important factor in their approach was that they first trained the generator with an MSE loss (like the original DBPN), and then the pre-trained generator was re-trained based on an objective function containing content (based on L2 norm of VGG-19) and adversarial losses. In their experiments, E-DBPN achieved higher quantitative evaluation results than other methods (including RCAN



Fig. 12. Super-resolution results presented in Yu et al. (2020): (a) HR, (b) bicubic resampling, (c) SRFeat (Park et al., 2018), (d) E-DBPN (Yu et al., 2020).

(Zhang et al., 2018) (Fig. 12). Considering the fact that detailed information and edges may be lost during super-resolution, Gao et al. (2021) used a residual channel attention module (aiming to improve loss of detailed information and blurred edges) in the generator of their WGAN to improve feature extraction while training for super-resolution. In fact, this approach was shown to provide more attention on high-frequency information (e.g., edges) responsible for enhancing the sharpness of the super-resolved images generated.

Many of the GAN-based approaches used for super-resolution tasks in RS have been based on approaches proposed earlier for super-resolution reconstruction of natural images. Directly applying those models without adapting them to RS imagery might not result in desirable results. As found by Lei et al. (2020), RS imagery may have more low-frequency components than natural images, possibly affecting the performance of the discriminator to correctly distinguish real images from generated ones when facing low-frequency parts of the image. To remedy this problem, the authors proposed a GAN-based super-resolution method called coupled GANs (CDGANs). Instead of considering one image at a time, the discriminator of CDGANs takes as input a pair of generated images and their corresponding ground-truth or reference. To further improve the quality of super-resolution, they also utilized some architectural modifications (e.g., the use of a dual-pathway network in the discriminator) as well as a specialized loss function (including coupled adversarial loss and pixel-space MSE loss). In a related study on considering the characteristics of RS images when designing super-resolution methods, Zhang et al. (2020) emphasized the importance of paying attention to the varying complexity of texture information in RS imagery when employing super-resolution. In this regard, the authors argued that applying a uniform super-resolution strategy for different land cover types may not be suitable for RS imagery. To account for this characteristic of RS imagery in super-resolution, the authors proposed a saliency-driven unequal super-resolution approach to adaptively consider the spatial characteristics of different regions. Their results showed the significance of their approach to considering the characteristics of different regions in RS imagery when utilizing super-resolution. In a relevant study, Ma et al. (2020) proposed a different saliency-driven GAN-based super-resolution making use of a pair of

discriminators to better identify salient regions in the image, resulting in the reconstruction of sharper details with less textural artifacts. Gong et al. (2021) also echoed the importance of the need for designing RS-specific super-resolution methods rather than directly using the ones tested on natural images. In their study, the well-known ESRGAN model was modified to be applicable to mid-resolution RS imagery (in this case, Sentinel-2 imagery with a resolution of 10 m). In this regard, in their WGAN-based super-resolution approach, the authors designed a convolutional block (called enlighten block) to improve the convergence of the network by providing an easier task (2-times up-sampling) as well as the target/harder task (4-times up-sampling), eventually resulting in better convergence and learning high-frequency details. Another important modification was the addition of a self-supervised hierarchical perceptual loss along with the image-space L2 and adversarial losses.

### 5.1.2. Image inpainting (interpolation)

Image inpainting (also commonly known as image interpolation in the context of RS) can also be done using GANs. Traditional approaches (especially non-ML approaches) used for recovering missing regions (caused, for example, by sensor defects, clouds, etc.) in a given image may result in unfavorable results, especially if missing areas are large or located in high-frequency regions of the image. GANs can be very advantageous for recovering missing regions in RS images. In a study conducted by Dong et al. (2019), the authors used a GAN-based inpainting approach to recover cloud-contaminated regions in sea-surface temperature (SST) products. Rather than relying solely on spatial information, that method also considered temporal information (i.e., historical SST data) in the process of image inpainting. Image inpainting also has applications in improving the density/resolution of rasterized sparse elevation data. In this regard, Zhu et al. (2020) used a GAN-based image inpainting approach to interpolate missing areas in digital elevation models (DEMs). Yan et al. (2021) also proposed a GAN-based interpolation approach to predict unobserved elevation data. Their proposed GAN was composed of a gated and symmetric-dilated U-net GAN (inspired by DeepFill v2 (Yu et al., 2018) to learn better features for the inpainting task. In Dong et al. (2020), the authors proposed a GAN framework to interpolate missing elevation information caused

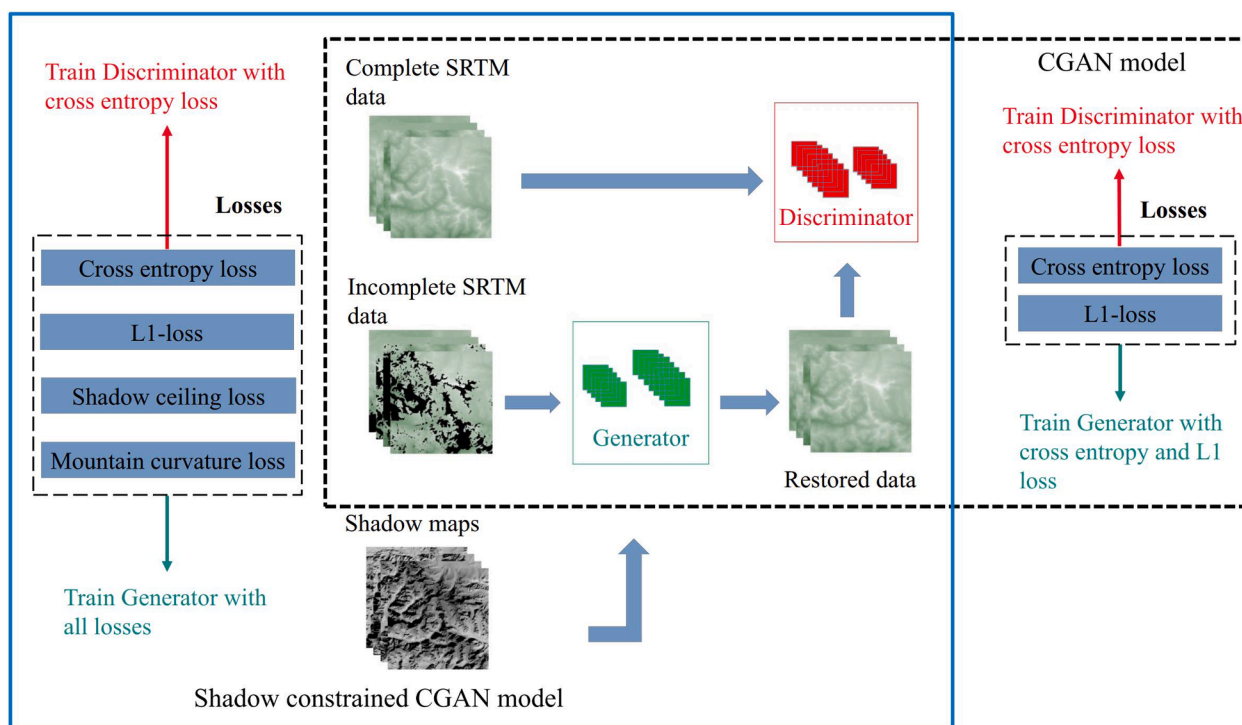


Fig. 13. GAN-based SRTM inpainting approach presented in Dong et al. (2020).

by shadow over mountainous areas in SRTM data. In this regard, to train the GAN, the authors used two additional loss terms (shadow loss and mountain-curvature loss) in the objective function to improve the inpainting of shadow-driven missing areas. More details on this approach can be seen in Fig. 13.

### 5.1.3. Image denoising

Noise is an integral part of RS data. However, the amount of noise varies depending on the sensor used for imaging. GANs have shown remarkable capabilities for reducing noise from RS imagery while preserving the overall sharpness and detailed contents of the image. Feng et al. (2021) proposed a unified GAN framework for simultaneous denoising and super-resolution reconstruction of RS imagery. Due to the limitations of the spatial domain for simultaneous image denoising and super-resolution reconstruction, the authors trained the GAN in the wavelet transform domain, helping them to process different frequency components separately. The effectiveness of using GAN for image denoising was also presented in Wang et al. (2018). The authors reported the efficacy of the GAN-based denoised images used for image matching and classification (trucks and cars) compared to the use of other denoising approaches like DCNNs (Mao et al., 2016). As shown in Fig. 14, images denoised by this approach maintained image details while keeping noise level at a very low level.

## 5.2. Image classification

Image classification is one of the most common applications of RS imagery. There are numerous ML-based approaches proposed for different image classification tasks. In this study, we categorize them as semantic segmentation, scene classification, and object detection. GANs have been widely used for various image classification tasks to improve the generalization power of DL models, especially in cases where labeled/annotated data are scarce. In the following three sub-sections, we review some of the applications of GANs for image classification of RS imagery.

### 5.2.1. Semantic segmentation

Semantic segmentation (or (dense) pixel-wise/level classification) is

one of the most common applications of RS imagery. Due to their exceptional performance, CNNs are quite popular for semantic segmentation in the RS community. In many semantic segmentation tasks in RS, the main problem is the lack of diverse/representative ground-truth data (also called labeled data in the literature), because collecting ground-truth data is an expensive, cumbersome task. Due to this issue, ML models (in the context of this review paper, DL models) fail to generalize well to test data. To mitigate this issue, there are several approaches that can be used such as data augmentation techniques, regularizers (dropout, batch/group/instance normalization, and regularized weights), multi-task training, etc. GANs can be useful for improving semantic segmentation results in different ways ranging from data augmentation to domain adaptation. Building on a Bayesian framework (having a prior network and a likelihood network), He et al. (2019) proposed an end-to-end cGAN integrated with conditional random fields (CRF) for semantic segmentation of RS images. The integration of the skip-connected encoder-decoder generator with a CRF layer helped the extraction of better local and global information from the image, and thus improving segmentation results compared to DeepLab (Chen et al., 2018) (which is a powerful DL semantic segmentation architecture). In a similar study, Xiong et al. (2020) proposed a different end-to-end GAN-based Bayesian segmentation framework in which a prior network produced preliminary segmentations, and the likelihood network (i.e., GAN operating on the segmentation map for exploring spatial relationships among labels) refined the segmentation outputs of the prior network. He et al. (2019) integrated object-based image analysis (OBIA) with a semi-supervised GAN framework for the classification of wetlands. To improve the trade-off between accuracy and speed, the authors in that study incorporated ShuffleNet units (Ma et al., 2018) into their proposed semi-supervised GAN (called ShuffleGAN). The experiments in that study showed that ShuffleGAN was superior to the original semi-supervised GAN for the classification of wetlands. In Guo et al. (2021), the authors modified Pix2Pix by utilizing an FC-DenseNet as the generator to perform semantic segmentation on RS images. Their experiments showed the effectiveness of this modification over the vanilla Pix2Pix for semantic segmentation. Sui et al. (2021) used a cGAN framework conditioned on labels, intra-class edge features, and inter-class boundary features to augment data for semantic

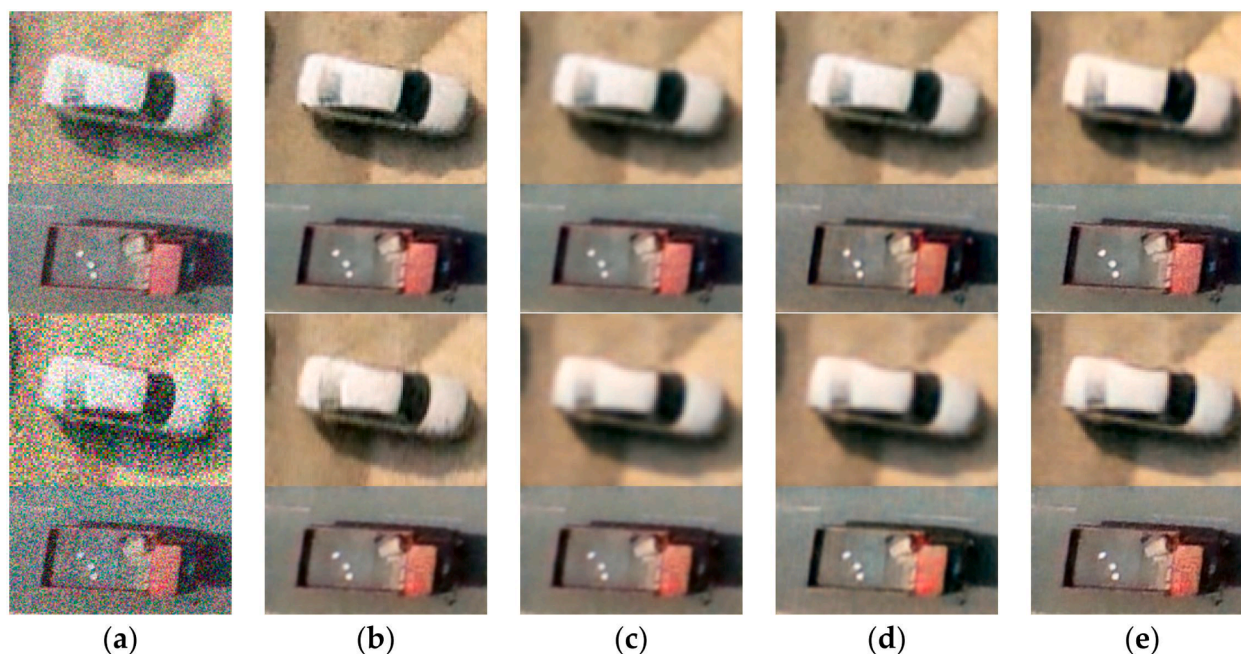


Fig. 14. Denoising results presented in Wang et al. (2018) based on: (a) noisy images, (b) denoised images by Chang et al. (2014), (c) denoised images by Xu et al. (2017), (d) denoised images by Mao et al. (2016), (e) denoised images by Wang et al. (2018).

segmentation tasks. It was shown that the proposed cGAN-based augmentation in that study resulted in more accurate segmentation maps compared to a plain cGAN conditioned only on labels. To address the vanishing-gradient problem, common in the training of DL models, Zhang and Hu (2017) proposed a conditional LSGAN approach for semantic segmentation of RS images. The authors reported the superiority of their approach compared to a cGAN and plain FCN framework. In order to improve the learning of high-level discriminative features from raw SAR images, Ren et al. (2020) proposed an enhanced AC-GAN framework (distribution and structure match AC-GAN (DSM-ACGAN)). Due to its dual-adversarial learning nature and due to considering the corresponding class-specific distribution and structure characteristics of real SAR images, DSM-ACGAN was able to improve the learning of discriminative, high-level features helping produce more accurate SAR classification maps compared to some other methods like plain AC-GAN (Odena et al., 2016) and ResNet18 models.

Building extraction is one of the most common applications of semantic segmentation in RS. This is a challenging task as buildings can have intra-class variability and inter-class similarity to some other impervious features such roads. To address high intra-class variations of buildings, Sun et al. (2021) proposed the generation of a background map and building map separately through two inter-connected orthogonal GANs (O-GAN (Su, 2019)). In a study conducted by Shi et al. (2019), it was shown that building footprint maps generated using cWGAN(-GP) frameworks were more accurate than those generated based on cGAN, and the plain U-Net model. In a study conducted by Pan et al. (2019), a GAN-based building extraction approach with spatial and channel attention (SCA) mechanisms was proposed. An important advantage of adopting SCA mechanisms in the proposed GAN-based building extraction framework was to selectively concentrate more on some specific important features (both spatially and spectrally), leading to more accurate building extraction results. In order to improve building footprint extraction using GANs over complex regions, Abdollahi et al. (2020) employed bi-directional convolutional LSTM (BConvLSTM) layers in the generator (SegNet (Badrinarayanan et al., 2017)). The authors showed that their proposed GAN framework was superior to some other approaches (such as the use of the same generator trained in a non-adversarial way).

Road extraction is another important application in semantic segmentation of RS imagery. Several studies have adopted GANs to improve road extraction from RS images (Shamsolmoali et al., 2021; Zhang et al., 2019a, 2019b; Shi et al., 2018; Li et al., 2019). Shi et al. (2018) designed an end-to-end cGAN-based segmentation framework (with a SegNet as the generator) to extract roads from RS imagery (a mapping task from image to road map), which was shown to be overall more accurate than applying a plain SegNet. In another study on road segmentation, Abdollahi et al. (2021) used a cGAN, with a modified U-Net generator inspired by Enokiya et al. (2018), to extract roads in high-resolution aerial imagery. In Li et al. (2019), the authors used a cGAN-based segmentation framework integrated with a multi-scale feature aggregation module to account for the uneven distribution of roads (depending on the UAV altitude) in UAV images. Their experiments showed the importance of using the multi-scale feature aggregation in their cGAN-based segmentation approach. Yang and Wang (2020) adopted an ensemble cWGAN-GP segmentation approach (based on Pix2Pix) by training two GANs and then intersecting their outputs to improve road segmentation accuracies. One of the challenges in road extraction is to preserve the coherency of the road networks. However, different types of occlusions (such as shadows) prevent this. To address this issue, Zhang et al. (2019) proposed a topology-aware road network extraction approach based on cGANs. Their proposed approach (called multi-supervised GAN (MsGAN)) was able to consider both spectral and topological features to improve road extraction. In another study on improving the topology of road networks extracted from RS imagery, Zhang et al. (2019) proposed a multi-conditional GAN framework with two discriminators, one of which was aimed for topology reconstruction

of road networks, and the other one for topology refinement of road networks. Improving the efficiency of DL models is important as they may be overparameterized and inefficient for the application at hand and lead to an overfitting. By improving its efficiency, a lightweight version of Pix2Pix was proposed by Cira et al. (2021) to improve segmented roads from aerial imagery. In this regard, the proposed framework took as input Gaussian noise and an initial segmented road map (as the condition), and generated a refined road map based on the corresponding ground-truth. To improve multi-scale feature extraction (which is an important factor for road segmentation, as mentioned earlier), Shamsolmoali et al. (2021) proposed a GAN-based approach (called adversarial spatial pyramid network (ASPN)). The key advantage of their proposed approach was the use of an efficient spatial Laplacian pyramid network as the generator that helped extract pyramid features at multiple scales.

GANs have also been widely used for semantic segmentation of hyperspectral data. Zhu et al. (2018) experimented with a 1D-CNN-based and a 3D-CNN-based AC-GAN approach to classify hyperspectral imagery. After training the AC-GANs on PCA-transformed data (for the sake of efficiency), the authors further fine-tuned the discriminators on the real and fake data. According to the experiments, the 3D-CNN-based approach resulted in better accuracies, as it considered both spectral and spatial information. Wang et al. (2021) also proposed an AC-GAN approach for classification of hyperspectral imagery using an adaptive Dropblock (Ghiasi et al., 2018) in both the generator and discriminator to mitigate the mode-collapse issue. The Dropblock adaptively generated variable drop shapes rather than a fixed-size one. Semi-supervised learning (involving unlabeled data as well as labeled data in the training process) is one of the approaches to improving the performance of the model when sufficient labeled data are not available. This can be effectively accomplished with GANs. He et al. (2017) employed a semi-supervised GAN approach for classifying hyperspectral images. Rather than inputting hyperspectral images/pixels directly into the GAN, the authors used spatial-spectral features extracted with a 3D bilateral filter. In a relevant study, Zhan et al. (2018) proposed a semi-supervised 1D-GAN that improved the performance of the model when few training samples were available. Taking into account the lack of training data and noise in hyperspectral images, Gao et al. (2019) proposed a semi-supervised multi-discriminator GAN-based segmentation approach to mitigate some of the common issues of GANs (such as mode collapse) for classifying hyperspectral imagery. To classify hyperspectral images, Hang et al. (2021) proposed a multi-task GAN framework. The generator in the proposed framework aimed to perform the reconstruction and classification tasks simultaneously, resulting in more accurate hyperspectral classifications than some other approaches (like WGAN, CNN). For this purpose, the generator was composed of an encoder-decoder and a classifier, in which the structure of the encoder was shared to improve the classification task as well as the reconstruction one. Ali-pour-Fard and Arefi (2020) proposed a structure-aware GAN framework for generating fake hyperspectral image patches to be used as a form of data augmentation for a subsequent classification. In their proposed unconditional GAN, the discriminator was tasked with distinguishing fake data from real data and with distinguishing structurally corrupted data from non-corrupted data. Applying these two tasks, compared to the conventional GANs, helped generate more realistic and diverse samples to be used in conjunction with real ones to improve the final classification. Feng et al. (2020) proposed a GAN-based hyperspectral classification approach based on collaborative learning and a spatial-spectral attention mechanism. To further improve the performance of the GAN approach for image segmentation, the authors also used a convolutional LSTM layer in the discriminator to be able to capture long-term spectral dependencies as well as contextual and spatial features. The experiments showed that the proposed approach overall led to more accurate results than some other approaches like 3D-GAN. One of the powerful variants of GANs is Variational Autoencoder GANs (VAE-GANs (Larsen et al., 2015)), which have been reported to be more accurate than

VAEs and conventional GANs (Larsen et al., 2015). The structure of VAE-GANs is basically similar to conventional GANs except that a VAE is integrated with the generator, resulting in three components: encoder, decoder/generator, and discriminator. Tao et al. (2020) proposed an end-to-end hyperspectral-classification approach based on a semi-supervised VAEGAN (SSVGAN) with ensemble predictions, capable of taking advantage of both labeled and unlabeled data during model training. In order to generate realistic, meaningful samples for training the classifier in the SSVGAN approach, the authors also introduced a new collaborative optimization mechanism. Overall, the tests conducted in that study showed the effectiveness of SSVGAN for hyperspectral image classification.

### 5.2.2. Object detection

As with semantic segmentation and DL methods in general, one of the most challenging issues in object detection is the lack of sufficient training data, especially if the object of interest is already rare or is too small (Bashir and Wang, 2021; Courtrai et al., 2020). This, as a result, causes the object-detection model to fail to perform properly. Data augmentation techniques and transfer learning are the two approaches that may improve object detection with few training samples. GANs have also been reported to ameliorate object detection tasks as they can assist with data augmentation and can improve the robustness of the model accordingly. In a study by Zhu et al. (2020), a multi-branch cGAN (called MCGAN) was proposed to generate diverse samples for object detection tasks. To improve the diversity of the images generated, MCGAN used three branches for distinguishing fake from real data, and one classification branch for predicting the classes of input objects. The authors also used an adaptive sample-selection strategy to filter out generated samples that had distributions different from that of the real data. The sampled generated images were then used as augmented data for training a subsequent supervised CNN object-detection model (Faster R-CNN (Ren et al., 2015), and found to improve the accuracy of object detection compared to the use of traditional data augmentation techniques (i.e., rotation, flipping, and contrast transformation). In another study, Zheng et al. (2019) utilized a GAN framework for synthesizing vehicles in RS images for vehicle detection tasks. Their proposed framework had one generator and two discriminators that were tasked with simultaneously learning vehicle generation and background context. The experiments in that study showed that vehicle detection models trained on the combined real and synthesized vehicle samples (generated by the proposed framework) resulted in higher detection accuracies than the models trained only on real data. Zhang et al. (2018) used a cGAN framework to more accurately recognize aircraft types in RS images. Basically, the main goal of the proposed cGAN framework was to extract more discriminative, representative features to be fed into a subsequent aircraft classifier (in this case, an SVM). The proposed GAN took as input aircraft masks (generated using key points) and generated corresponding synthetic aircraft images. In order to learn multi-scale features (to account for the resolution and scale of aircrafts appearing in RS images), three discriminators were used for operating on three different image sizes. According to the experiments conducted in that study, it was shown that the proposed aircraft recognition framework was more accurate than CNN approaches like ResNet-18. GANs can also be used to improve the robustness of the object detector in cases where objects are small compared to the spatial resolution of the data. Such approaches are mainly based on super-resolution reconstruction. In fact, to improve the extraction of small objects in RS images, such studies have typically adopted super-resolution to enhance the resolution of images that can then facilitate the detection of objects of interest (Bashir and Wang, 2021; Rabbi et al., 2020; Courtrai et al., 2020).

### 5.2.3. Scene classification

As seen in Fig. 8, although scene classification (also known as image-level classification) in RS is less common than semantic segmentation and object detection, the advent of DL has provided an unprecedented

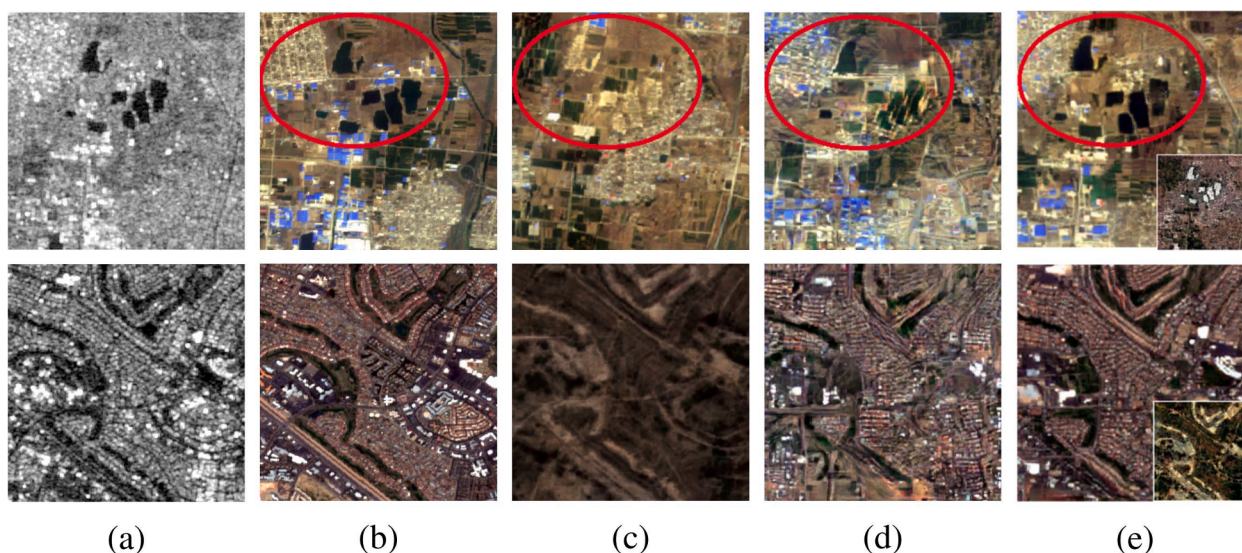
opportunity to further explore RS scene classification in recent years. Unsurprisingly, GANs can also be used to improve scene classification by improving the robustness of the classifier (Teng et al., 2020; Xu et al., 2018). As with any other classification task in RS, DL-based scene classification requires a large number of annotated data, which is difficult to afford due to many data- and labor-related challenges. Taking this issue into account, Han et al. (2020) utilized an AC-WGAN framework to simulate annotated RS samples for scene classification tasks. The experiments in that study showed that not only the use of AC-WGAN improved classification accuracies in the presence of limited training data, but also it was superior to some other data augmentation approaches (including CycleGAN). In Yan et al. (2020), the authors used a semi-supervised GAN framework to improve scene classification. In addition to the semi-supervised nature of the framework in that study, inspired by MARTA-GAN (Lin et al., 2017), another key factor in that method was the use of a feature matching loss component, calculated based on the multi-feature layer of the discriminator. The importance of the use of multi-feature layer for training GANs in scene classification was also demonstrated in Wei et al. (2020). In this respect, the authors made use of a multi-feature layer module to train WGAN-GP (called MF-WGAN). After training the GAN model, the discriminator (along with the multi-feature layer) was unplugged and used to train an MLP classifier. The authors reported that their MF-WGAN was superior to MARTA-GAN. Inspired by the symmetrical and incremental GAN training strategy (Karras et al., 2017; Pan et al., 2020) proposed a GAN framework capable of generating more diverse, controllable samples for scene classification purposes. In this regard, the proposed approach used a progressive technique to generate fake samples from coarse to fine resolutions, which eventually helped learn more diverse structure features at different scales that led to generating more diverse samples accordingly. The experiments in that study showed that training a CNN by combining the real and fake images generated by the proposed approach resulted in higher classification accuracies than using only real images, and using real, augmented (rotation and flipping), and MARTA-GAN. As discussed earlier, the use of attention mechanisms can help narrow down the focus on more useful/relevant features and can help model long-range dependencies, and thus improve generating more realistic images. In a study by Guo et al. (2021), the authors proposed a scene classification framework based on gated self-attention GAN (SGSAGAN) with a novel similarity loss. One of the key features in their proposed GAN-based framework was the use of two networks as the discriminator, namely an online and a target into which two augmented views of images were fed and the similarity loss was calculated based on the outputs of the two networks. As the authors reported, the experiments on two benchmark datasets showed that the SGSAGAN led to higher scene-classification accuracies than the other approaches including Attention GANs (Yu et al., 2020).

### 5.3. Data translation

One of the most popular applications of GANs in RS is data translation including image-to-image translation and text-to-image translation. In this section, we review some of data translation applications performed with GANs in the context of RS.

#### 5.3.1. Image-to-image translation

Image-to-image translation has various applications ranging from image colorization to cross-sensor style transferring. Such applications have also been explored in several RS studies (Zheng et al., 2018; Song et al., 2021; Kim et al., 2020; Hayatbini et al., 2019). One of the most practical applications of image-to-image translation is to generate cloud-free images from cloud-contaminated ones. In a study conducted by Li et al. (2020), the authors proposed a semi-supervised cloud removal framework based on cGANs and a physical model of cloud distortion to recover thin-cloud-contaminated areas using unpaired RS images (in this case, 10-m bands of Sentinel-2 imagery). Their proposed cloud-



**Fig. 15.** SAR to optical image translation results presented in Ji et al. (2021): (a) SAR image, (b) real optical image, (c) generated optical image by CycleGAN, (d) generated optical image by Pix2Pix, (e) generated optical image by Ji et al. (2021).

removal framework was composed of two components, namely “removal network” (cGAN; responsible for translating cloud-contaminated image to cloud-free image) and “extraction network” (responsible for decomposing a cloud-contaminated image into three cloud distortion layers). According to the evaluations performed in the study, it was shown that the unpaired GAN-based cloud-removal approach was overall comparable to paired approaches (like U-Net), which may not be efficient as collecting paired data may not be possible in some cases. In a recent study on simulating cloud-free Sentinel-2 images using Sentinel-1 images, Xiong et al. (2021) proposed a multi-temporal paired GAN-based approach (called Multi-channels Conditional Generative Adversarial Network (MCcGAN)), which was designed based on the supervised CycleGAN (S-CycleGAN) approach proposed earlier (Wang et al., 2019). According to their experiments, it was reported that MCcGAN overall performed better than the other mono- and multi-temporal approaches including S-CycleGAN. In another study on the use of GANs for thin-cloud removal from RS imagery, Chen et al. (2021) utilized a cGAN with a spatial Attention mechanism in its generator to improve the restoration of thin-cloud-contaminated regions. The generator of the proposed approach had two sub-networks, namely an attention network (to better identify cloud-contaminated regions) and a contextual autoencoder (to generate cloud-free images). Their experiments on Google Earth images showed that their approach performed better than plain cGANs. Thin-cloud removal using GANs was also conducted in Wen et al. (2021). In that study, rather than the RGB color space, the authors used the YUV color space to reduce the number of unrestorable bright and dark pixels. In the training process, the authors first pretrained the GAN based on simulated image pairs (simulated cloud-contaminated and cloud-free), and then fine-tuned it on real image pairs, which, based on comparisons with other GAN and non-adversarial approaches, overall led to higher-quality simulated cloud-free images.

Image-to-image translation can also be used for synthesizing elevation data from single-view RS images. For example, in a study carried out by Ghamisi and Yokoya (2018), the authors used a GAN framework (called IMG2DSM) to simulate DSM data from high-resolution single-view RS images. As well as measuring the reconstruction quality, the authors reported that the use of synthesized data together with high-resolution optical data improved the overall accuracy of semantic segmentations by a large margin. In a recent study, Paoletti et al. (2021) introduced an improved version of IMG2DSM (U-IMG2DSM) for optical-to-DSM translation tasks. One of the improvements of U-IMG2DSM was

that it was able to operate on unpaired data. In order to make the approach work with unpaired images, the authors integrated weight-sharing constraint of coGANs (Liu and Tuzel, 2016) with the latent-encoding of VAE-GANs (Larsen et al., 2015). The experiments in that study showed that although U-IMG2DSM performed slightly worse than IMG2DSM (but better than CycleGAN), it was overall more efficient than IMG2DSM which was not able to perform on unpaired data.

SAR images are much less affected by atmospheric variations than optical images. However, they may be difficult to interpret compared to optical images. In this regard, a few studies aimed to translate SAR images to optical images and used them for different downstream tasks such as semantic segmentation tasks. Fuentes Reyes et al. (2019) evaluated the potential of CycleGAN for unsupervised translation of SAR images (Sentinel-1 and TerraSAR-X) to optical images (Sentinel-2 and ALOS PRISM). To better evaluate the quality of the translation task, the authors trained a road extraction model (DeepLab V3+) on the real SAR images and synthesized optical images separately. They found that although the model trained on real SAR images achieved higher accuracies than the one trained on synthesized optical data, the approach was overall able to preserve important features of roads during the translation. In a more recent study, Zhang et al. (2021) adopted Pix2Pix for a SAR-to-optical translation task (Sentinel-1 to Landsat-8 (bands 1–7)) and reported that including more SAR-related information (such as different polarizations and edge information) for such a translation task improved the quality of the generated optical images. In another attempt to generate optical images from SAR images, Ji et al. (2021) proposed an unpaired GAN framework that accounted for the various types of terrains when performing the translation task. The authors argued that since different terrains in SAR images can have different color domains, not considering this may cause color confusion in the translated images. Their evaluations showed the superiority of the proposed approach for simulating visible RGB images (Sentinel-2 RGB bands) from SAR images (Sentinel-1 VV) compared to CycleGAN and Pix2Pix (Fig. 15).

Another interesting application of image-to-image translation is to create styled maps (like Google Maps) from RS images. For example, Song et al. (2021) proposed an efficient image-to-map framework using cGANs (called MapGen-GAN). One of the primary factors when translating images to styled maps is to meet geometrical consistency requirements. In order to account for semantic distortion and unpaired data for this translation task, MapGen-GAN used a geometrical-consistency loss as well as a cycle-consistency loss in the objective

function of the proposed GAN approach. The evaluations based on different metrics (such as SSIM and PSNR) showed the superiority of MapGen-GAN compared to other approaches like CycleGAN.

### 5.3.2. Text-to-image translation

In the previous section, we reviewed GAN-based image-to-image translation methods used in the field of RS. In addition to simulating RS images from other RS image (sources), it is also possible to simulate RS images based on text data. This unusual yet interesting task can also be performed with GANs. In a study conducted by [Bejiga et al. \(2019\)](#), the researchers investigated the potential of a GAN-based framework for synthesizing RS imagery based on geographically described ancient texts. In the proposed framework, the synthetization was done by feeding encoded text descriptions (along with a noise vector) into a cWGAN-GP responsible for simulating RS imagery based on the encoded texts. In a later study ([Bejiga et al., 2021](#)), the above-mentioned approach was improved to be able to generate more realistic images from ancient text descriptions. In the original approach, text encoding was performed by a simple binary approach that may not result in semantically meaningful synthetization of RS images corresponding to the text descriptions. To address this issue in their enhanced approach, the authors adopted a Doc2Vec text encoder in both the generator and discriminator, resulting in synthesized RS images more consistent with the text descriptions.

### 5.4. Domain adaptation (DA)

Domain adaptation (DA) is a sub-field of transfer learning, where the goal is to adapt an ML model to a target domain so it can hopefully perform as good as it does in the source domain. This way of improving the generalization power of ML models has been proven to be effective, especially in cases where collecting training samples in the target domain is time consuming. GANs have been reported to be advantageous for different DA tasks. In this section, we review the applications of GANs for DA in the context of RS. It should be noted that some of the applications in this section may have overlap with some of the previous applications (e.g., image classification), but we decided to separate them to better focus on the DA aspects of the approaches proposed.

Regardless of segmentation models used, it is apparent that the performance of an ML model is degraded if the test data are from a different domain ([Benjdira et al., 2019](#)). GANs can be used to alleviate cross-domain shifts in RS data, and thus improving the performance of the segmentation model to generalize well on a target domain. Intuitively, one workaround is to make target-domain images look like the source image(s) on which, for example, a segmentation model was trained, which is technically equivalent to correcting for the domain shift. There are several papers on GAN-based DA in the field of RS. [Ji et al. \(2021\)](#) proposed an unsupervised DA (UDA) approach based on an end-to-end unpaired GAN framework for LULC classification. One of the key features of their GAN framework was to apply three DA modules (i.e. image domain adaptation, feature domain adaptation, and output domain adaptation) to improve the quality of the style translation and segmentation accordingly. The proposed GAN framework was composed of two learning stages. In the first stage, the goal was to transfer style from the source domain images to the target domain images based on image domain adaptation and target domain adaptation modules. In the second stage, the target image, target-stylized source image, source image, and reconstructed source image (from the first stage) were used to train a segmentation network with an output space adaptation module (to align the output segmentation maps). Their evaluations on two datasets showed that their approach was superior to the commonly used AdaptSegNet ([Tsai et al., 2018](#)) and CycleGAN. [Tasar et al. \(2020\)](#) presented another unpaired GAN-based UDA framework (called ColorMapGAN) for LULC classification tasks. One of the key features of ColorMapGAN was its high efficiency resulting from the fact that it did not use any convolution or pooling operation in the generator, but rather

only an element-wise matrix multiplication and a matrix addition. The experiments in that study showed that ColorMapGAN overall resulted in more accurate LULC classification maps than the other approaches like CycleGAN. In another study on UDA for semantic segmentation, [Liu and Su \(2020\)](#) proposed a framework consisting of three trainable components: a feature extractor network, a cGAN, and an MLP classifier. Extracted curve features (from feature curves derived from a pretrained DeepLab V3+) along with a random noise vector were fed into the generator. In their proposed approach, the feature extractor network, GAN, and MLP classifier were trained simultaneously to optimize the framework, but during the test, only the feature extractor and classifier were used. According to the evaluations reported in that study, their proposed UDA approach achieved higher accuracies than other advanced approaches including CycleGAN and CyCADA ([Hoffman et al., 2017](#)). [Liu et al. \(2021\)](#) proposed a GAN-based UDA approach for hyperspectral classification by emphasizing the importance of aligning class-conditional distributions rather than only marginal distributions of the source and target domains. For this purpose, the proposed framework deployed an adaptation network consisting of class-wise adversarial adaptation and probability maximum mean discrepancy (PMMD, which was calculated based on the predicted probability outputs of the target data), resulting in higher classification accuracies than other state-of-the-art approaches including MADA ([Pei et al., 2018](#)). Although adapting a given source-domain RS image to a target-domain RS image has been shown to improve the accuracy of the semantic segmentation of target-domain RS images, it is also interesting to investigate how good translated non-RS images can contribute to the accuracy improvement of semantic segmentation. [Zou et al. \(2020\)](#) considered such a special UDA problem where the authors translated a set of in-game rendered aerial images of the popular "GTA V" videogame to aerial RS images to study how well such translated images would generalize to building detection tasks in RS images. For this purpose, the authors utilized CycleGAN (trained in an end-to-end fashion with an FCN for the building detection task). Not only did the authors report this approach was superior to applying a plain, non-adapted building detection model, but also they reported that jointly training CycleGAN and the building detection model resulted in higher accuracies than training and using them separately.

GAN-based DA methods have also been used for bi-temporal change detection tasks. For example, [Fang et al. \(2021\)](#) showed that the use of a GAN-based Siamese model improved accuracy mapping of landslide inventory. In the proposed approach, the GAN was responsible for performing DA between a pre- and post-landslide images. In the landslide detection part of the proposed framework, the adapted/translated pre-landslide image along with the original post-landslide image were then fed into a Siamese network optimized based on the contrastive loss calculated between the feature maps extracted from the two image inputs. In a more recent study, [Kou et al. \(2020\)](#) integrated a ConvLSTM network with a cGAN to spectrally align multi-temporal RS images to be used for change detection tasks, which can increase the likelihood of detecting real changes. Their proposed GAN-based DA approach consisted of two cascaded modules: 1) progressive translation (i.e., progressively mitigating seasonal-driven domain discrepancies between bi-temporal images), and 2) group discrimination (i.e., evaluating if unpaired generated and post-event images are real/fake).

GAN-based DA can also be carried out for RS scene classification tasks to improve the generalization of scene classifiers. In a study conducted by [Teng et al. \(2020\)](#), rather than using a single fake/real discriminator, the authors used two different scene land-cover classifiers in the discriminator. In fact, as the authors argued, the use of a dual-classifier setting for the discriminator would help prevent ambiguity in the vicinity of land cover decision boundaries, resulting in higher accuracies on target images. [Liu and Su \(2020\)](#) utilized a UDA approach by integrating an AC-GAN and a domain confusion network (trained in the target domain with the pseudo-labels generated by a source-trained classifier) to improve scene classification in the presence of domain



shifts. To strengthen the capability of the GAN-based UDA framework for decreasing feature discrepancies between domains, the authors also incorporated a feature extractor (jointly trained with the GAN) whose output features were fed into the AC-GAN, resulting in higher accuracies than CycleGAN.

### 5.5. Miscellaneous applications

Apart from the applications of GANs that were reviewed above, there are some other applications that may not be considered as a sub-category of the above applications or may overlap more than one of the above applications.

One such application is image pansharpening, which can also be considered as a special type of image fusion and super-resolution. The main goal of performing pansharpening is to enhance the resolution of multi-spectral bands using the panchromatic band, resulting in a single image containing both rich spectral and spatial information (Zhang et al., 2018). GANs have been reported to be advantageous for pansharpening (Zhang, 2019; Zhou et al., 2020; Zhou et al., 2021; Zhang et al., 2021). For example, Xie et al. (2021) proposed a 3D GAN (HPGAN) for pansharpening hyperspectral imagery to take advantage of both high-spatial and high-spectral information in a single image cube. Unlike SRGAN and ESRGAN, HPGAN used a least-squares loss function in the discriminator. One of the key points in HPGAN was that its generator operated in the high-frequency domain rather than the image domain to improve the generalization power of the network. In Shao et al. (2020), a residual encoder-decoder cGAN (RED-cGAN) was proposed to perform pansharpening. In order to preserve both spatial and spectral details as much as possible, the generator used a two-branch sub-network (followed by a Residual encoder-decoder network) capable of extracting hierarchical features (Shao and Cai, 2018). The experiments on WorldView-2 and -3 images showed that RED-cGAN overall led to higher fidelity pansharpened images than other methods like PSGAN (Liu et al., 2018).

Image retrieval is another miscellaneous application of GANs where the goal is to retrieve data of interest from a database based on image contents. The RS community has reported the potential of GANs for image retrieval in recent years (Xiong et al., 2020; Cao et al., 2020; Zhang et al., 2018; Zhang et al., 2019). In Cao et al. (2020), the authors proposed a deep metric learning method with GANs regularization (DML-GANR) for retrieval of high spatial resolution imagery (Fig. 16). In addition to the use of a multi-layer DML and a high-level feature extractor, the proposed retrieval framework used an unconditional GAN to reduce overfitting while training (in an end-to-end fashion), which was shown to improve retrieval accuracies in cases where training data are not sufficient. As discussed earlier, after training an unconditional GAN, the discriminator can be unplugged and used as a powerful

unsupervised feature extractor for different downstream tasks including image retrieval. For example, in Zhang et al. (2018) and Zhang et al. (2019), the authors proposed GAN-based approaches to improve retrieving hyperspectral images. In their proposed approaches, an unconditional DCGAN was used to extract more representative, discriminative spatial/spectral features helping improve the retrieval of queries.

GANs have also been used for image matching in the field of RS (Merkle et al., 2018; Du et al., 2021; Ma et al., 2021). Image matching has many important applications in computer vision. The key factor in image matching is to identify common features in the image pairs to be matched. Due to different limitations (e.g., illumination variations, sensor view angle variations, temporal variations, etc.), identifying common features can be difficult. Most of GAN-based image matching approaches are based on image translation concepts that aim to make a pair of images similar to each other (e.g., in terms of lighting variations) and then perform image matching techniques to the translated image and the other image. For instance, Ma et al. (2021) trained a cGAN (regularized with a pixel-space L1 norm and gradient L1 norm) to first conduct image translation on the target image to reduce nonlinear variations. The tests in that study based on applying SIFT and SURF feature detectors/descriptors showed the significant efficacy of using GANs for improving RS image matching. In a study by Merkle et al. (2018), the authors considered matching visible images with SAR images based on a cGAN, cLSGAN, and cWGAN. In this regard, the GANs were used to transfer visible images to SAR-like images, and feature detectors/descriptors were applied on the generated SAR images and real ones. The authors reported that the use of images generated by the cLSGAN led to more accurate image matching results than those generated by cGAN and cWGAN.

One of the challenges of RS data (especially if they are multispectral/hyperspectral) can be their size. Given this, it is favorable to reduce their size by utilizing, if possible, a lossless compression technique. GANs have been reported to be powerful approaches to compress RS images while preserving the quality and spectral/spatial information as much as possible. In one of the few studies in this field, Zhao et al. (2021) proposed a symmetrical lattice GAN (SLGAN) to compress RS images. Since the quality of compressed images can be degraded, the authors designed an enhanced Laplacian of Gaussian (ELoG) loss to improve edges, textures, and counters of the compressed images. The evaluations of that study (on Gaofen-2 images with a spatial resolution of 1 m) showed that SLGAN was overall superior to other classic (e.g., JPEG 2000) and advanced approaches (e.g., ComGAN (Santurkar et al., 2017)).

## 6. Challenges and future directions

As discussed in the previous sections, GANs can be advantageous for many types of RS applications. For example, in image classifications,

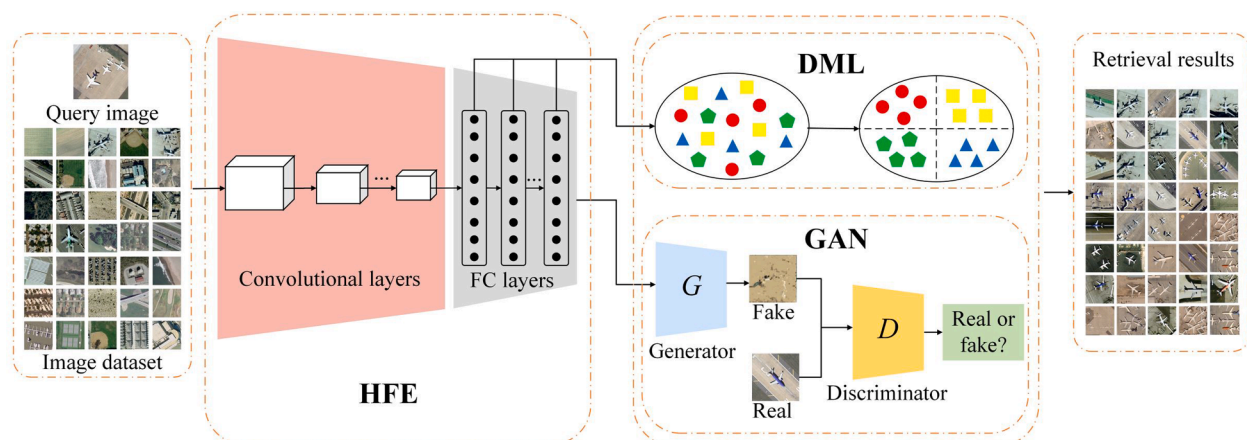


Fig. 16. GAN-based image retrieval approach proposed and presented in Cao et al. (2020).

GANs are valuable for data augmentation and improving the robustness and generalization of classifiers. This is crucial in cases where sufficient labeled data are not available, which is a common problem in most, if not all, practical applications in the context of RS. For image reconstruction/restoration applications, GANs have shown to be more accurate than traditional approaches, especially for super-resolution tasks; GANs can generate detailed, realistically super-resolved images that may not be possible to replicate using non-GAN approaches. For data-translation tasks, particularly image-to-image translation, GANs have shown top performance. In fact, along with the sharpness of translated images, past research has shown that color-transferring accuracy in translated images derived from GANs is higher than plain CNNs. GANs have also proven to be beneficial for other RS applications including image matching, image fusion, and image retrieval for which traditional processing approaches often struggle to handle due to their high level of complexity.

Although GANs have provided unprecedented opportunities for synthesizing RS images and for improving model performance in different applications, particularly image classification tasks, there are still several critical challenges that have not yet been addressed properly. One of the main challenges is the difficulty in optimizing GANs. Although there have been new ways (e.g., new architectures, objective functions, training strategies, etc.) for mitigating this issue, it is not guaranteed by any means to achieve optimally converged GANs. Another important challenge is the way through which generated images are evaluated. Supervised image-level element-wise metrics (e.g., based on L1 and L2 norms) are not generally reliable to judge the quality of generated images against the real ones. Feature-level metrics based on pretrained models (most notably, VGG-16 and -19) are generally more reliable because small shifts negligible to human eyes can result in different image-level losses although the contents of the generated and real images perceptually are very similar. However, such pre-trained models for evaluating generated images operate on RGB images, as they were trained on millions of natural RGB images. Therefore, applying such models to multi-spectral images is not possible by default. Task specific metrics are also useful to quantify the quality of generated images based on the application at hand. However, this indirect way of evaluation does not take into account the uncertainty in the classification accuracy. Given such challenges in evaluating generated images by GANs, one of the future directions in this field can be to conduct a comprehensive empirical study to analyze which metric(s) can objectively and reliably best evaluate GANs performance in different RS applications.

As the meta-analysis results showed, GANs need to be tested in more diverse study areas. This is, however, understandable as collecting labeled data in RS is typically a significantly challenging task. Given this, a number of studies tested the potential of their approaches mainly on benchmark datasets. However, several popular benchmark datasets are composed of RGB images. Given the fact that many RS data have more than three bands, such benchmarks may not be sufficiently representative of RS data commonly applied for different practical applications. We also observed that GANs have not been widely applied to commonly used publicly available data (such as Sentinel and Landsat images) compared to high-resolution data. Since the constellation and continuation of such images have been one of the most important advances in the field of RS, there needs to be more research on the potential of GANs on such publicly available medium-resolution images that are widely used in different academic fields.

Moreover, the literature is scarce of comprehensive empirical studies on the performance of GANs compared to other advanced and classic approaches in various applications. In fact, an open question is that for which applications GANs make a significant improvement compared to other approaches. For example, in a comprehensive study by Li et al. (2020), the potential of a CycleGAN and AGGAN (Tang et al., 2019) DA approach for segmentation and object-detection tasks was analyzed. The authors found that although the two GAN approaches improved

semantic segmentation compared to the direct prediction, they did not perform better than classic, simpler approaches. For object-detection, the authors reported that the GAN approaches performed the worst. Such a study is also worth conducting for other applications and with different RS datasets to better inform the RS community about how well GANs can perform against other approaches. Another relevant gap in this field is the lack of the evaluation of GANs performance against other well-known generative models like VAEs, which tend to generate less crisp images than GANs do. Given this, there is a need for a solid empirical study that shows if the use of GANs has any meaningful advantage over VAEs for different downstream RS tasks, such as image classification.

One other direction in need of future research is the effectiveness and significance of different loss terms in training GANs. It is still unclear which loss functions and for what applications can tangibly improve the performance of GANs in the field of RS. In this regard, an empirical study on analyzing the effectiveness of various loss functions for training GANs in different applications can substantially help the community to more reliably configure objective functions for training GANs.

## 7. Conclusions

Generative adversarial networks (GANs) have been one of the most creative advances in the field of deep learning (DL) in recent years. Although there have been several review papers on DL related to RS in the last 4 years, to our knowledge, there have not been any journal review paper explicitly on the applications of GANs in remote sensing (RS). In this review, 231 journal papers published from 2017 to 2021 (July 17th) were considered. We conducted a comprehensive study on GAN approaches, applications, challenges, and trends in the field of RS.

Our meta-analysis results showed a significant increase in the number of papers on GANs in the field of RS since 2017 (from 4 to 91 in 2020, and to 72 until July 17th, 2021). We observed that GANs have been applied to various RS applications and using a wide variety of datasets. Some of the most important results of the meta-analysis conducted in this study were as follows:

- **Study application:** Image classification (38%), data translation (23%), image reconstruction (21%) have been the top three most common applications of GANs in the papers reviewed, respectively.
- **RS data source:** Among the different RS data sources, RGB images were used more than multi-spectral, hyperspectral, SAR, and LiDAR, respectively.
- **Spatial resolution of images:** Sub-meter resolution images were the most common types of images used in the papers reviewed in this study, while popular medium-resolution image sources (such as Landsat) were not used as widely.
- **Study area:** A large number of papers considered a mixture of different types of sites (47%). Of these different types of study sites, the majority of papers considered urban areas (32%) as their study areas, possibly due to the high popularity of urban mapping using RS data in recent years.
- **Loss function:** Among different loss functions, the pixel-space (based on L1 or L2 norm) and perceptual loss functions (based on VGG-16/19) have been the most common loss functions, respectively, used along with the adversarial loss.
- **Evaluation metrics:** Image-classification accuracy measures (specifically OA) were the most commonly used type of evaluation metrics used. This is rather unsurprising as we found that most papers focused on image classification. Among direct evaluation metrics, SSIM and PSNR were used most frequently (to compare the images generated by GANs with corresponding reference images).

The meta-analysis also indicated that many papers reviewed in this study properly evaluated the performance of GANs either using direct evaluation metrics (i.e., metrics that evaluate the quality of the images

generated by a GAN or indirect evaluation metrics (i.e., metrics that evaluate the quality/accuracy of GAN-generated images used for different downstream tasks such as image classification). Although there have been seminal papers on GANs in the context of RS, there are still several challenges that should be further studied and addressed to improve the performance and applicability of GANs in various RS applications. As elaborated, these challenges are not limited to the difficulty in training GANs, but also the way we evaluate their results. For future studies, we suggest conducting some comprehensive empirical studies on best practices for training GANs, such as the choice of objective functions, evaluation metrics, and network architectures. Such studies would be very worthwhile to develop more versatile GANs for different RS applications. It is also crucial to apply GANs to more diverse areas, especially those areas where non-adversarial approaches generally struggle to produce satisfactory results.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

This research was financially supported by National Science and Engineering Research Council (NSERC) Discovery grant and Queen's University Graduate Award.

### References

- Abdollahi, A., Pradhan, B., Gite, S., Alamri, A., 2020. Building footprint extraction from high resolution aerial images using generative adversarial network (GAN) architecture. *IEEE Access* 8, 209517–209527. <https://doi.org/10.1109/ACCESS.2020.3038225>.
- Abdollahi, A., Pradhan, B., Sharma, G., Maulud, K.N.A., Alamri, A., 2021. Improving road semantic segmentation using generative adversarial network. *IEEE Access* 9, 64381–64392. <https://doi.org/10.1109/ACCESS.2021.3075951>.
- Adamiak, M., Będkowski, K., Majchrowska, A., 2021. Aerial imagery feature engineering using bidirectional generative adversarial networks: a case study of the Pilica River Region, Poland. *Remote Sens.* 13 (2), 306. <https://doi.org/10.3390/rs13020306>.
- Alipour-Fard, T., Arefi, H., 2020. Structure aware generative adversarial networks for hyperspectral image classification. *IEEE J. Selected Top. Appl. Earth Observat. Remote Sens.* 13, 5424–5438. <https://doi.org/10.1109/JSTARS.2020.3022781>.
- Alqahtani, H., Kavakli-Thorne, M., Kumar, G., 2021. Applications of generative adversarial networks (GANs): an updated review. *Arch. Comput. Methods Eng.* 28 (2), 525–552. <https://doi.org/10.1007/s11831-019-09388-y>.
- M. Arjovsky, S. Chintala, L. Bottou, Wasserstein GAN, 2017. [Online]. Available: <https://arxiv.org/abs/1701.07875>.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (12), 2481–2495. <https://doi.org/10.1109/TPAMI.2016.2644615>.
- Bashir, S.M.A., Wang, Y., 2021. Small Object detection in remote sensing images with residual feature aggregation-based super-resolution and object detector network. *Remote Sens.* 13 (9) <https://doi.org/10.3390/rs13091854>.
- Bejiga, M.B., Melgani, F., Vascotto, A., 2019. Retro-remote sensing: generating images from ancient texts. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing* 12 (3), 950–960.
- Bejiga, M.B., Hoxha, G., Melgani, F., 2021. Improving text encoding for retro-remote sensing. *IEEE Geosci. Remote Sensing Lett.* 18 (4), 622–626.
- Benjdira, B., Bazi, Y., Koubaa, A., Ouni, K., 2019. Unsupervised domain adaptation using generative adversarial networks for semantic segmentation of aerial images. *Remote Sensing* 11 (11), 1369. <https://doi.org/10.3390/rs11111369>.
- Bittner, K., d'Angelo, P., Körner, M., Reinartz, P., 2018. DSM-to-LoD2: spaceborne stereo digital surface model refinement. *Remote Sensing* 10 (12), 1926. <https://doi.org/10.3390/rs10121926>.
- Burdziakowski, P., 2020. A novel method for the deblurring of photogrammetric images using conditional generative adversarial networks. *Remote Sens.* 12 (16), 2586. <https://doi.org/10.3390/rs12162586>.
- Cao, Y., Wang, Y., Peng, J., Zhang, L., Xu, L., Yan, K., et al., 2020. DML-GANR: deep metric learning with generative adversarial network regularization for high spatial resolution remote sensing image retrieval. *IEEE Trans. Geosci. Remote Sensing* 58 (12), 8888–8904.
- Chang, Y., Yan, L., Fang, H., Liu, H., 2014. Simultaneous destriping and denoising for remote sensing images with unidirectional total variation and sparse representation. *IEEE Geosci. Remote Sens. Lett.* 11 (6), 1051–1055. <https://doi.org/10.1109/LGRS.2013.2285124>.
- Chen, H., Chen, R., Li, N., 2021. Attentive generative adversarial network for removing thin cloud from a single remote sensing image. *IET Image Process* 15 (4), 856–867.
- Chen, X.u., Chen, S., Xu, T., Yin, B., Peng, J., Mei, X., et al., 2021. SMAPGAN: generative adversarial network-based semisupervised styled map tile generation method. *IEEE Trans. Geosci. Remote Sensing* 59 (5), 4388–4406.
- Chen, C., Ma, H., Yao, G., Lv, N., Yang, H., Li, C., et al., 2021. Remote sensing image augmentation based on text description for waterline change detection. *Remote Sens.* 13 (10), 1894. <https://doi.org/10.3390/rs13101894>.
- Chen, L., Papanandreu, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4), 834–848. <https://doi.org/10.1109/TPAMI.2017.2699184>.
- Chen, J., Wang, L., Feng, R., Liu, P., Han, W., Chen, X., 2021. CycleGAN-STF: spatiotemporal fusion via CycleGAN-based image generation. *IEEE Trans. Geosci. Remote Sensing* 59 (7), 5851–5865.
- L.-C. Chen et al., Naive-Student: Leveraging Semi-Supervised Learning in Video Sequences for Urban Scene Segmentation, 2020. [Online]. Available: <https://arxiv.org/abs/2005.10266>.
- Cheng, G., Han, J., Lu, X., 2017. Remote sensing image scene classification: benchmark and state of the art. *Proc. IEEE* 105 (10), 1865–1883. <https://doi.org/10.1109/JPROC.2017.2675998>.
- Cira, C.-I., Manso-Callejo, M.-Á., Alcarria, R., Fernández Pareja, T., Bordel Sánchez, B., Serradilla, F., 2021. Generative learning for postprocessing semantic segmentation predictions: a lightweight conditional generative adversarial network based on Pix2pix to improve the extraction of road surface areas. *Land* 10 (1), 79. <https://doi.org/10.3390/land10010079>.
- Courtrai, L., Pham, M.-T., Lefevre, S., 2020. Small object detection in remote sensing images based on super-resolution with auxiliary generative adversarial networks. *Remote Sensing* 12 (19), 3152. <https://doi.org/10.3390/rs12193152>.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A., 2018. Generative adversarial networks: an overview. *IEEE Signal Process Mag.* 35 (1), 53–65. <https://doi.org/10.1109/MSP.2017.2765202>.
- A. Dash, J. Ye, G. Wang, A review of Generative Adversarial Networks (GANs) and its applications in a wide variety of disciplines - From Medical to Remote Sensing, *ArXiv*, vol. abs/2110.01442, 2021.
- J. Donahue, P. Krähenbühl, T. Darrell, Adversarial Feature Learning, 2016. [Online]. Available: <https://arxiv.org/abs/1605.09782>.
- Dong, G.S., Huang, W.M., Smith, W.A.P., Ren, P., 2020. A shadow constrained conditional generative adversarial net for SRTM data restoration. *Remote Sens. Environ.* 237, FEB. <https://doi.org/10.1016/j.rse.2019.111602>.
- Dong, J., Yin, R., Sun, X., Li, Q., Yang, Y., Qin, X., 2019. inpainting of remote sensing SST images with deep convolutional generative adversarial network. *IEEE Geosci. Remote Sensing Lett.* 16 (2), 173–177.
- A. Dosovitskiy et al., An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, 2020. [Online]. Available: <https://arxiv.org/abs/2010.11929>.
- Dou, X.Y., Li, C.Y., Shi, Q., Liu, M.X., 2020. Super-resolution for hyperspectral remote sensing images based on the 3D attention-SRGAN network. *Remote Sens.* 12 (7) <https://doi.org/10.3390/rs12071204>.
- Du, W.L., Zhou, Y., Zhao, J.Q., Tian, X.L., Yang, Z., Bian, F.Q., 2021. Exploring the potential of unsupervised image synthesis for SAR-optical image matching. *IEEE Access* 9, 71022–71033. <https://doi.org/10.1109/ACCESS.2021.3079327>.
- Ebel, P., Meraner, A., Schmitt, M., Zhu, X.X., 2021. Multisensor data fusion for cloud removal in global and all-season sentinel-2 imagery. *IEEE Trans. Geosci. Remote Sensing* 59 (7), 5866–5878.
- Enokiyama, Y., Iwamoto, Y., Chen, Y.-W., Han, X.-H., 2018. Automatic liver segmentation using U-net with Wasserstein GANs. *J. Image Graph.* 6 (2), 152–159.
- Fang, J., Cao, X., Wang, D., Xu, S., 2021. Multitask learning mechanism for remote sensing image motion deblurring. *IEEE J. Selected Topics Applied Earth Observ. Remote Sens.* 14, 2184–2193. <https://doi.org/10.1109/JSTARS.2020.3047636>.
- Fang, B.o., Chen, G., Pan, L.i., Kou, R., Wang, L., 2021. GAN-based siamese framework for landslide inventory mapping using bi-temporal optical remote sensing images. *IEEE Geosci. Remote Sensing Lett.* 18 (3), 391–395.
- Feng, J., Feng, X., Chen, J., Cao, X., Zhang, X., Jiao, L., et al., 2020. Generative adversarial networks based on collaborative learning and attention mechanism for hyperspectral image classification. *Remote Sens.* 12 (7), 1149. <https://doi.org/10.3390/rs12071149>.
- Feng, X.B., Zhang, W.X., Su, X.Q., Xu, Z.P., 2021. Optical remote sensing image denoising and super-resolution reconstructing using optimized generative network in wavelet transform domain. *Remote Sens.* 13 (9) <https://doi.org/10.3390/rs13091858>.
- 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. <https://doi.org/10.3390/rs11172067>.
- Gao, L.i., Sun, H.-M., Cui, Z., Du, Y.-B., Sun, H.-B., Jia, R.-S., 2021. Super-resolution reconstruction of single remote sensing images based on residual channel attention. *J. Appl. Rem. Sens.* 15 (01) <https://doi.org/10.1117/1.JRS.15.016153>.
- Gao, H., Yao, D., Wang, M., Li, C., Liu, H., Hua, Z., et al., 2019. A hyperspectral image classification method based on multi-discriminator generative adversarial networks. *Sensors* 19 (15), 3269. <https://doi.org/10.3390/s19153269>.
- 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. <https://doi.org/10.3390/rs12010191>.
- Ge, W.Y., Wang, Z.T., Wang, G.G., Tan, S.H., Zhang, J.W., 2021. Remote sensing image super-resolution for the visual system of a flight simulator: dataset and baseline. *Aerospace* 8 (3). <https://doi.org/10.3390/aerospace8030076>.

- Ghamisi, P., Yokoya, N., 2018. IMG2DSM: height simulation from single imagery using conditional generative adversarial net. *IEEE Geosci. Remote Sensing Lett.* 15 (5), 794–798.
- G. Ghiasi, T.-Y. Lin, Q. V. Le, DropBlock: A regularization method for convolutional networks, 2018. [Online]. Available: <https://arxiv.org/abs/1810.12890>.
- Gong, Y., Liao, P., Zhang, X., Zhang, L., Chen, G., Zhu, K., et al., 2021. Enlighten-GAN for super resolution reconstruction in mid-resolution remote sensing images. *REMOTE Sens.* 13 (6), 1104. <https://doi.org/10.3390/rs13061104>.
- Gonzalez, R.C., Woods, R.E., 2006. *Digital Image Processing*, 3rd ed. Prentice-Hall Inc.
- Good, I.J., 1952. Rational decisions. *J. Roy. Stat. Soc.: Ser. B (Methodol.)* 14 (1), 107–114.
- I. J. Goodfellow et al., “Generative Adversarial Networks,” 2014. [Online]. Available: <https://arxiv.org/abs/1406.2661>.
- Grant, M.J., Booth, A., 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Inform. Libraries J.* 26 (2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>.
- J. Gui, Z. Sun, Y. Wen, D. Tao, J. Ye, A review on generative adversarial networks: algorithms, theory, and applications, 2020. [Online]. Available: <http://arXiv.org/abs/>
- I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville, Improved Training of Wasserstein GANs, 2017. [Online]. Available: <https://arxiv.org/abs/1704.00028>.
- Guo, X., Chen, Z., Wang, Q., 2021. Fully convolutional DenseNet with adversarial training for semantic segmentation of high-resolution remote sensing images. *J. Appl. Rem. Sens.* 15 (01) <https://doi.org/10.1117/1.JRS.15.016520>.
- Guo, D., Xia, Y., Luo, X., 2021. Self-supervised GANs with similarity loss for remote sensing image scene classification. *IEEE J. Selected Top. Appl. Earth Observat. Remote Sens.* 14, 2508–2521. <https://doi.org/10.1109/JSTARS.2021.3056883>.
- Han, W., Wang, L., Feng, R., Gao, L., Chen, X., Deng, Z.e., et al., 2020. Sample generation based on a supervised Wasserstein Generative Adversarial Network for high-resolution remote-sensing scene classification. *Inf. Sci.* 539, 177–194.
- Hang, R., Zhou, F., Liu, Q., Ghamisi, P., 2021. Classification of hyperspectral images via multitask generative adversarial networks. *IEEE Trans. Geosci. Remote Sensing* 59 (2), 1424–1436.
- M. Haris, G. Shakhnarovich, N. Ukita, Deep Back-Projection Networks For Super-Resolution, 2018. [Online]. Available: <https://arxiv.org/abs/1803.02735>.
- Hayatbini, N., Kong, B., Hsu, K.-L., Nguyen, P., Sorooshian, S., Stephens, G., et al., 2019. Conditional generative adversarial networks (cGANs) for near real-time precipitation estimation from multispectral GOES-16 satellite imagery—PERSIANN-cGAN. *Remote Sens.* 11 (19), 2193. <https://doi.org/10.3390/rs11192193>.
- He, C., Fang, P., Zhang, Z., Xiong, D., Liao, M., 2019. An end-to-end conditional random fields and skip-connected generative adversarial segmentation network for remote sensing images. *Remote Sensing* 11 (13), 1604. <https://doi.org/10.3390/rs11131604>.
- He, Z., Liu, H., Wang, Y., Hu, J., 2017. Generative adversarial networks-based semi-supervised learning for hyperspectral image classification. *Remote Sensing* 9 (10), 1042. <https://doi.org/10.3390/rs9101042>.
- He, Z., He, D., Mei, X., Hu, S., 2019. Wetland classification based on a new efficient generative adversarial network and jilin-1 satellite image. *Remote Sensing* 11 (20), 2455. <https://doi.org/10.3390/rs11202455>.
- Heydari, S.S., Mountrakis, G., 2019. Meta-analysis of deep neural networks in remote sensing: a comparative study of mono-temporal classification to support vector machines. *ISPRS J. Photogramm. Remote Sens.* 152, 192–210. <https://doi.org/10.1016/j.isprsjprs.2019.04.016>.
- J. Hoffman et al., CyCADA: Cycle-Consistent Adversarial Domain Adaptation,” 2017, doi: abs/.
- Hu, A., Xie, Z., Xu, Y., Xie, M., Wu, L., Qiu, Q., 2020. Unsupervised haze removal for high-resolution optical remote-sensing images based on improved generative adversarial networks. *Remote Sensing* 12 (24), 4162. <https://doi.org/10.3390/rs12244162>.
- Huang, Z.-X., Jing, C.-W., 2020. Super-resolution reconstruction method of remote sensing image based on multi-feature fusion. *IEEE Access* 8, 18764–18771. <https://doi.org/10.1109/ACCESS.2020.2967804>.
- Hughes, L., Schmitt, M., Zhu, X., 2018. Mining hard negative samples for SAR-optical image matching using generative adversarial networks. *Remote Sensing* 10 (10), 1552. <https://doi.org/10.3390/rs10101552>.
- P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks, 2016. [Online]. Available: <https://arxiv.org/abs/1611.07004>.
- Ji, G., Wang, Z., Zhou, L., Xia, Y.u., Zhong, S., Gong, S., 2021. SAR image colorization using multidomain cycle-consistency generative adversarial network. *IEEE Geosci. Remote Sensing Lett.* 18 (2), 296–300.
- Ji, S., Wang, D., Luo, M., 2021. Generative adversarial network-based full-space domain adaptation for land cover classification from multiple-source remote sensing images. *IEEE Trans. Geosci. Remote Sensing* 59 (5), 3816–3828.
- Jiang, K., Wang, Z., Yi, P., Wang, G., Lu, T., Jiang, J., 2019. Edge-enhanced GAN for remote sensing image superresolution. *IEEE Trans. Geosci. Remote Sensing* 57 (8), 5799–5812.
- Johnson, B., Jozdani, S., 2018. Identifying generalizable image segmentation parameters for urban land cover mapping through meta-analysis and regression tree modeling. *Remote Sensing* 10 (2), 73. <https://doi.org/10.3390/rs10010073>.
- Jozdani, S.E., Johnson, B.A., Chen, D., 2019. Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. *Remote Sensing* 11 (14), 1713. <https://doi.org/10.3390/rs11141713>.
- T. Karras, T. Aila, S. Laine, J. Lehtinen, Progressive Growing of GANs for Improved Quality, Stability, and Variation, 2017. [Online]. Available: <https://arxiv.org/abs/1710.10196>.
- Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on convolutional neural networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* 173, 24–49. <https://doi.org/10.1016/j.isprsjprs.2020.12.010>.
- Khatami, R., Mountrakis, G., Stehman, S.V., 2016. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: general guidelines for practitioners and future research. *Remote Sens. Environ.* 177, 89–100. <https://doi.org/10.1016/j.rse.2016.02.028>.
- Kim, J.-H., Ryu, S., Jeong, J., So, D., Ban, H.-J., Hong, S., 2020. Impact of satellite sounding data on virtual visible imagery generation using conditional generative adversarial network. *IEEE J. Selected Top. Appl. Earth Observat. Remote Sens.* 13, 4532–4541. <https://doi.org/10.1109/JSTARS.2020.3013598>.
- Kotaridis, I., Lazaridou, M., 2021. Remote sensing image segmentation advances: a meta-analysis. *ISPRS J. Photogramm. Remote Sens.* 173, 309–322. <https://doi.org/10.1016/j.isprsjprs.2021.01.020>.
- Kou, R., Fang, B.o., Chen, G., Wang, L., 2020. Progressive domain adaptation for change detection using season-varying remote sensing images. *Remote Sensing* 12 (22), 3815. <https://doi.org/10.3390/rs12223815>.
- A. B. L. Larsen, S. K. Sønderby, H. Larochelle, O. Winther, Autoencoding beyond pixels using a learned similarity metric, 2015. [Online]. Available: <https://arxiv.org/abs/1512.09300>.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521 (7553), 436–444. <https://doi.org/10.1038/nature14539>.
- C. Ledig et al., Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, 2016. [Online]. Available: <https://arxiv.org/abs/1609.04802>.
- Lei, S., Shi, Z., Zou, Z., 2020. Coupled adversarial training for remote sensing image super-resolution. *IEEE Trans. Geosci. Remote Sens.* 58 (5), 3633–3643.
- Leinonen, J., Guillaume, A., Yuan, T., 2019. Reconstruction of cloud vertical structure with a generative adversarial network. *Geophys. Res. Lett.* 46 (12), 7035–7044.
- Li, J., Chen, Z., Zhao, X., Shao, L., 2020. MapGAN: an intelligent generation model for network tile maps. *Sensors* 20 (11), 3119. <https://doi.org/10.3390/s20113119>.
- Li, H., Gao, S., Liu, G., Guo, D., Grecos, C., Ren, P., 2020. Visual prediction of typhoon clouds with hierarchical generative adversarial networks. *IEEE Geosci. Remote Sens. Lett.* 17 (9), 1478–1482.
- Li, G., Li, J., Fan, H., 2020. Edge-guided multispectral image fusion algorithm. *J. Appl. Rem. Sens.* 14 (04) <https://doi.org/10.1117/1.JRS.14.046515>.
- Li, X., Luo, M., Ji, S., Zhang, L.i., Lu, M., 2020. Evaluating generative adversarial network based image-level domain transfer for multi-source remote sensing image segmentation and object detection. *Int. J. Remote Sens.* 41 (19), 7343–7367.
- Li, Y., Peng, B.o., He, L., Fan, K., Tong, L., 2019. Road segmentation of unmanned aerial vehicle remote sensing images using adversarial network with multiscale context aggregation. *IEEE J. Sel. Top. Appl. Earth Observat. Remote Sens.* 12 (7), 2279–2287.
- Li, J., Wu, Z., Hu, Z., Zhang, J., Li, M., Mo, L.u., et al., 2020. Thin cloud removal in optical remote sensing images based on generative adversarial networks and physical model of cloud distortion. *ISPRS J. Photogramm. Remote Sens.* 166, 373–389.
- Lin, D., Fu, K., Wang, Y., Xu, G., Sun, X., 2017. MARTA GANs: unsupervised representation learning for remote sensing image classification. *IEEE Geosci. Remote Sensing Lett.* 14 (11), 2092–2096.
- M.-Y. Liu, O. Tuzel, Coupled Generative Adversarial Networks, 2016. [Online]. Available: <https://arxiv.org/abs/1606.07536>.
- X. Liu, Y. Wang, Q. Liu, Psgan: A generative adversarial network for remote sensing image pan-sharpening, in: *2018 25th IEEE International Conference on Image Processing (ICIP)*, 7–10 Oct. 2018 2018, pp. 873–877, doi: 10.1109/ICIP.2018.8451049.
- Liu, Z., Ma, L.i., Du, Q., 2021. Class-wise distribution adaptation for unsupervised classification of hyperspectral remote sensing images. *IEEE Trans. Geosci. Remote Sensing* 59 (1), 508–521.
- Liu, W., Su, F., 2020. A novel unsupervised adversarial domain adaptation network for remotely sensed scene classification. *Int. J. Remote Sens.* 41 (16), 6099–6116.
- Liu, W., Su, F., 2020. Unsupervised adversarial domain adaptation network for semantic segmentation. *IEEE Geosci. Remote Sensing Lett.* 17 (11), 1978–1982.
- Liu, Y., Wang, W., Fang, F., Zhou, L., Sun, C., Zheng, Y., et al., 2021. CscGAN: conditional scale-consistent generation network for multi-level remote sensing image to map translation. *Remote Sens.* 13 (10), 1936. <https://doi.org/10.3390/rs13101936>.
- N. Ma, X. Zhang, H.-T. Zheng, J. Sun, ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design, 2018. [Online]. Available: <https://arxiv.org/abs/1807.11164>.
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: a meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* 152, 166–177. <https://doi.org/10.1016/j.isprsjprs.2019.04.015>.
- Ma, T., Ma, J., Yu, K., Zhang, J., Fu, W., 2021. Multispectral remote sensing image matching via image transfer by regularized conditional generative adversarial networks and local feature. *IEEE Geosci. Remote Sensing Lett.* 18 (2), 351–355.
- Ma, W., Pan, Z., Yuan, F., Lei, B., 2019. Super-resolution of remote sensing images via a dense residual generative adversarial network. *Remote Sensing* 11 (21), 2578. <https://doi.org/10.3390/rs11212578>.
- Ma, J., Zhang, L., Zhang, J., 2020. SD-GAN: saliency-discriminated GAN for remote sensing image superresolution. *IEEE Geosci. Remote Sensing Lett.* 17 (11), 1973–1977.
- X. J. Mao, C. Shen, Y. B. Yang, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections, 2016.

- X.-J. Mao, C. Shen, and Y.-B. Yang, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections, 2016. [Online]. Available: <https://arxiv.org/abs/1603.09056>.
- X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, S.P. Smolley, Least Squares Generative Adversarial Networks, 2016. [Online]. Available: <https://arxiv.org/abs/1611.04076>.
- Merkle, N., Auer, S., Muller, R., Reinartz, P., 2018. Exploring the potential of conditional adversarial networks for optical and SAR image matching. *IEEE J. Sel. Top. Appl. Earth Observat. Remote Sens.* 11 (6), 1811–1820.
- M. Mirza, S. Osindero, Conditional Generative Adversarial Nets, 2014. [Online]. Available: <https://arxiv.org/abs/1411.1784>.
- Niu, X., Gong, M., Zhan, T., Yang, Y., 2019. A conditional adversarial network for change detection in heterogeneous images. *IEEE Geosci. Remote Sensing Lett.* 16 (1), 45–49.
- A. Odena, C. Olah, J. Shlens, Conditional Image Synthesis With Auxiliary Classifier GANs, 2016. [Online]. Available: <https://arxiv.org/abs/1610.09585>.
- A. Odena, Semi-Supervised Learning with Generative Adversarial Networks, 2016. [Online]. Available: <https://arxiv.org/abs/1606.01583>.
- Ozcelik, F., Alganci, U., Sertel, E., Unal, G., 2021. Rethinking CNN-based pansharpening: guided colorization of panchromatic images via GANs. *IEEE Trans. Geosci. Remote Sens.* 59 (4), 3486–3501.
- Pan, X., Yang, F., Gao, L., Chen, Z., Zhang, B., Fan, H., et al., 2019. Building extraction from high-resolution aerial imagery using a generative adversarial network with spatial and channel attention mechanisms. *Remote Sensing* 11 (8), 917. <https://doi.org/10.3390/rs11080917>.
- Pan, X., Zhao, J., Xu, J., 2020. A Scene images diversity improvement generative adversarial network for remote sensing image scene classification. *IEEE Geosci. Remote Sensing Lett.* 17 (10), 1692–1696.
- Paoletti, M.E., Haut, J.M., Ghamisi, P., Yokoya, N., Plaza, J., Plaza, A., 2021. U-IMG2DSM: unpaired simulation of digital surface models with generative adversarial networks. *IEEE Geosci. Remote Sensing Lett.* 18 (7), 1288–1292.
- S.J. Park, H. Son, S. Cho, K. S. Hong, S. Lee, "SRFeat: Single Image Super-Resolution with Feature Discrimination, 2018, doi: 10.1007/978-3-030-01270-0\_27.
- Z. Pei, Z. Cao, M. Long, J. Wang, Multi-Adversarial Domain Adaptation, 2018, doi: abs/. Peng, D., Bruzzone, L., Zhang, Y., Guan, H., Ding, H., Huang, X.u., 2021. SemiCDNet: A semisupervised convolutional neural network for change detection in high resolution remote-sensing images. *IEEE Trans. Geosci. Remote Sensing* 59 (7), 5891–5906.
- Rabbi, J., Ray, N., Schubert, M., Chowdhury, S., Chao, D., 2020. Small-object detection in remote sensing images with end-to-end edge-enhanced GAN and object detector network. *Remote Sens.* 12 (9) <https://doi.org/10.3390/rs12091432>.
- S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, 2015. [Online]. Available: <https://arxiv.org/abs/1506.01497>.
- Ren, Z., Hou, B., Wu, Q., Wen, Z., Jiao, L., 2020. A distribution and structure match generative adversarial network for SAR image classification. *IEEE Trans. Geosci. Remote Sensing* 58 (6), 3864–3880.
- Rudin, L.I., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. *Physica D* 60 (1), 259–268. [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F).
- Saha, S., Bovolo, F., Bruzzone, L., 2021. Building change detection in VHR SAR images via unsupervised deep transcoding. *IEEE Trans. Geosci. Remote Sensing* 59 (3), 1917–1929.
- Salgueiro Romero, L., Marcello, J., Vilaplana, V., 2020. Super-resolution of sentinel-2 imagery using generative adversarial networks. *Remote Sensing* 12 (15), 2424. <https://doi.org/10.3390/rs12152424>.
- T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, X. Chen, Improved Techniques for Training GANs, 2016. [Online]. Available: <https://arxiv.org/abs/1606.03498>.
- S. Santurkar, D. Budden, N. Shavit, Generative Compression, 2017. [Online]. Available: <https://arxiv.org/abs/1703.01467>.
- Shamsolmoali, P., Zareapoor, M., Zhou, H., Wang, R., Yang, J., 2021. Road segmentation for remote sensing images using adversarial spatial pyramid networks. *IEEE Trans. Geosci. Remote Sensing* 59 (6), 4673–4688.
- Shamsolmoali, P., Zareapoor, M., Granjer, E., Zhou, H., Wang, R., Celebi, M.E., et al., 2021. Image synthesis with adversarial networks: a comprehensive survey and case studies. *Information Fusion* 72, 126–146.
- Shao, Z., Cai, J., 2018. Remote sensing image fusion with deep convolutional neural network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 11 (5), 1656–1669. <https://doi.org/10.1109/JSTARS.2018.2805923>.
- Shao, Z., Lu, Z., Ran, M., Fang, L., Zhou, J., Zhang, Y.i., 2020. Residual encoder-decoder conditional generative adversarial network for pansharpening. *IEEE Geosci. Remote Sens. Lett.* 17 (9), 1573–1577.
- Shi, Y., Li, Q., Zhu, X.X., 2019. Building footprint generation using improved generative adversarial networks. *IEEE Geosci. Remote Sensing Lett.* 16 (4), 603–607.
- Shi, Q., Liu, X.P., Li, X., 2018. Road detection from remote sensing images by generative adversarial networks. *IEEE Access* 6, 25486–25494. <https://doi.org/10.1109/ACCESS.2017.2773142>.
- Shi, X., Zhou, F., Yang, S., Zhang, Z., Su, T., 2019. Automatic target recognition for synthetic aperture radar images based on super-resolution generative adversarial network and deep convolutional neural network. *Remote Sens.* 11 (2), 135. <https://doi.org/10.3390/rs11020135>.
- K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014. [Online]. Available: <https://arxiv.org/abs/1409.1556>.
- Song, J., Li, J., Chen, H., Wu, J., 2021. MapGen-GAN: a fast translator for remote sensing image to map via unsupervised adversarial learning. *IEEE J. Selected Top. Appl. Earth Observat. Remote Sens.* 14, 2341–2357. <https://doi.org/10.1109/JSTARS.2021.3049905>.
- J. Su, O-GAN: Extremely Concise Approach for Auto-Encoding Generative Adversarial Networks, 2019. [Online]. Available: <https://arxiv.org/abs/1903.01931>.
- Sui, B., Jiang, T., Zhang, Z., Pan, X., 2021. ECGAN: an improved conditional generative adversarial network with edge detection to augment limited training data for the classification of remote sensing images with high spatial resolution. *IEEE J. Selected Top. Appl. Earth Observat. Remote Sens.* 14, 1311–1325. <https://doi.org/10.1109/JSTARS.2020.3033529>.
- Sun, S., Mu, L., Wang, L., Liu, P., Liu, X., Zhang, Y., 2021. Semantic segmentation for buildings of large intra-class variation in remote sensing images with O-GAN. *Remote Sensing* 13 (3), 475. <https://doi.org/10.3390/rs13030475>.
- H. Tang, D. Xu, N. Sebe, Y. Yan, Attention-guided generative adversarial networks for unsupervised image-to-image translation, 2019. [Online]. Available: <https://arxiv.org/abs/1903.12296>.
- Tang, R., Liu, H., Wei, J., 2020. Visualizing near infrared hyperspectral images with generative adversarial networks. *Remote Sensing* 12 (23), 3848. <https://doi.org/10.3390/rs12233848>.
- Tao, Y.u., Muller, J.-P., 2019. Super-resolution restoration of MISR images using the UCL MAGiGAN system. *Remote Sens.* 11 (1), 52. <https://doi.org/10.3390/rs11010052>.
- Tao, Y.u., Muller, J.-P., 2021. Super-resolution restoration of spaceborne ultra-high-resolution images using the UCL OpTiGAN system. *Remote Sens.* 13 (12), 2269. <https://doi.org/10.3390/rs13122269>.
- Tao, C., Wang, H., Qi, J.i., Li, H., 2020. Semisupervised variational generative adversarial networks for hyperspectral image classification. *IEEE J. Selected Top. Appl. Earth Observat. Remote Sens.* 13, 914–927. <https://doi.org/10.1109/JSTARS.2020.2974577>.
- Tasar, O., Happy, S.L., Tarabalka, Y., Alliez, P., 2020. ColorMapGAN: unsupervised domain adaptation for semantic segmentation using color mapping generative adversarial networks. *IEEE Trans. Geosci. Remote Sens.* 58 (10), 7178–7193.
- Teng, W., Wang, N.i., Shi, H., Liu, Y., Wang, J., 2020. Classifier-constrained deep adversarial domain adaptation for cross-domain semisupervised classification in remote sensing images. *IEEE Geosci. Remote Sensing Lett.* 17 (5), 789–793.
- Y.-H. Tsai, W.-C. Hung, S. Schultzer, K. Sohn, M.-H. Yang, M. Chandraker, Learning to Adapt Structured Output Space for Semantic Segmentation, 2018. [Online]. Available: <https://arxiv.org/abs/1802.10349>.
- T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs, 2017. [Online]. Available: <https://arxiv.org/abs/1711.11585>.
- Y. Wang, Q. Yao, J. Kwok, and L. M. Ni, "Generalizing from a Few Examples: A Survey on Few-Shot Learning," 2019. [Online]. Available: <https://arxiv.org/abs/1904.05046>.
- Wang, J., Gao, F., Dong, J., Du, Q., 2021. Adaptive DropBlock-enhanced generative adversarial networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* 59 (6), 5040–5053.
- Wang, Z., Jiang, K., Yi, P., Han, Z., He, Z., 2020. Ultra-dense GAN for satellite imagery super-resolution. *Neurocomputing* 398, 328–337.
- Wang, R., Xiao, X., Guo, B., Qin, Q., Chen, R., 2018. An Effective image denoising method for UAV images via improved generative adversarial networks. *Sensors* 18 (7), 1985. <https://doi.org/10.3390/s18071985>.
- Wang, L., Xu, X., Yu, Y., Yang, R., Gui, R., Xu, Z., et al., 2019. SAR-to-optical image translation using supervised cycle-consistent adversarial networks. *IEEE Access* 7, 129136–129149. <https://doi.org/10.1109/ACCESS.2019.2939649>.
- X. Wang et al., ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks, 2018. [Online]. Available: <https://arxiv.org/abs/1809.00219>.
- Wei, Y., Luo, X., Hu, L., Peng, Y., Feng, J., 2020. An improved unsupervised representation learning generative adversarial network for remote sensing image scene classification. *Remote Sensing Letters* 11 (6), 598–607.
- Wen, X., Pan, Z.X., Hu, Y.X., Liu, J.Y., 2021. Generative adversarial learning in YUV color space for thin cloud removal on satellite imagery. *REMOTE SENSING* 13 (6). <https://doi.org/10.3390/rs13061079>.
- Xia, G.-S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., et al., 2017. AID: a benchmark data set for performance evaluation of aerial scene classification. *IEEE Trans. Geosci. Remote Sens.* 55 (7), 3965–3981. <https://doi.org/10.1109/TGRS.2017.2685945>.
- Xie, W., Cui, Y., Li, Y., Lei, J., Du, Q., Li, J., 2021. HPGAN: hyperspectral pansharpening using 3-D generative adversarial networks. *IEEE Trans. Geosci. Remote Sens.* 59 (1), 463–477.
- Xiong, Q., Di, L., Feng, Q., Liu, D., Liu, W., Zan, X., et al., 2021. Deriving non-cloud contaminated sentinel-2 images with RGB and near-infrared bands from sentinel-1 images based on a conditional generative adversarial network. *Remote Sens.* 13 (8), 1512. <https://doi.org/10.3390/rs13081512>.
- Xiong, Y., Guo, S., Chen, J., Deng, X., Sun, L., Zheng, X., et al., 2020. Improved SRGAN for remote sensing image super-resolution across locations and sensors. *Remote Sens.* 12 (8), 1263. <https://doi.org/10.3390/rs12081263>.
- Xiong, D., He, C., Liu, X., Liao, M., 2020. An end-to-end bayesian segmentation network based on a generative adversarial network for remote sensing images. *Remote Sensing* 12 (2), 216. <https://doi.org/10.3390/rs12020216>.
- Xiong, W., Lv, Y., Zhang, X., Cui, Y., 2020. Learning to translate for cross-source remote sensing image retrieval. *IEEE Trans. Geosci. Remote Sensing* 58 (7), 4860–4874.
- Xu, S.H., Mu, X.D., Chai, D., Zhang, X.M., 2018. Remote sensing image scene classification based on generative adversarial networks. *Remote Sens. Lett.* 9 (7), 617–626. <https://doi.org/10.1080/2150704X.2018.1453173>.
- Xu, S., Zhou, Y., Xiang, H., Li, S., 2017. Remote sensing image denoising using patch grouping-based nonlocal means algorithm. *IEEE Geosci. Remote Sens. Lett.* 14 (12), 2275–2279. <https://doi.org/10.1109/LGRS.2017.2761812>.
- Yan, P., He, F., Yang, Y., Hu, F., 2020. Semi-supervised representation learning for remote sensing image classification based on generative adversarial networks. *IEEE Access* 8, 54135–54144. <https://doi.org/10.1109/ACCESS.2020.2981358>.

- Yan, L., Tang, X.F., Zhang, Y., 2021. High accuracy interpolation of DEM using generative adversarial network. *Remote Sens.* 13 (4) <https://doi.org/10.3390/rs13040676>.
- Yang, C., Wang, Z., 2020. An ensemble Wasserstein generative adversarial network method for road extraction from high resolution remote sensing images in rural areas. *IEEE Access* 8, 174317–174324. <https://doi.org/10.1109/ACCESS.2020.3026084>.
- Yu, W., Bai, J., Jiao, L., 2020. Background subtraction based on GAN and domain adaptation for VHR optical remote sensing videos. *IEEE Access* 8, 119144–119157. <https://doi.org/10.1109/ACCESS.2020.3004495>.
- J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, T. Huang, Free-form image inpainting with gated convolution, 2018. [Online]. Available: <https://arxiv.org/abs/1806.03589>.
- Yu, Y., Li, X., Liu, F., 2020. E-DBPN: enhanced deep back-projection networks for remote sensing scene image superresolution. *IEEE Trans. Geosci. Remote Sens.* 58 (8), 5503–5515.
- Yu, Y., Li, X., Liu, F., 2020. Attention GANs: unsupervised deep feature learning for aerial scene classification. *IEEE Trans. Geosci. Remote Sens.* 58 (1), 519–531.
- Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., et al., 2020. Deep learning in environmental remote sensing: achievements and challenges. *Remote Sens. Environ.* 241, 111716. <https://doi.org/10.1016/j.rse.2020.111716>.
- Yue, H., Cheng, J., Liu, Z., Chen, W., 2020. Remote-sensing image super-resolution using classifier-based generative adversarial networks. *J. Appl. Rem. Sens.* 14 (04) <https://doi.org/10.1117/1.JRS.14.046514>.
- Y. Zhai, M. Shah, Visual attention detection in video sequences using spatiotemporal cues, presented at the Proceedings of the 14th ACM international conference on Multimedia, Santa Barbara, CA, USA, 2006. [Online]. Available: <https://doi.org/10.1145/1180639.1180824>.
- Zhan, Y., Hu, D., Wang, Y., Yu, X., 2018. Semisupervised hyperspectral image classification based on generative adversarial networks. *IEEE Geosci. Remote Sensing Lett.* 15 (2), 212–216.
- Zhang, Y., 2019. SFTGAN: a generative adversarial network for pan-sharpening equipped with spatial feature transform layers. *J. Appl. Rem. Sens.* 13 (02), 1. <https://doi.org/10.1117/1.JRS.13.026507>.
- Zhang, J., Chen, L.u., Zhuo, L.i., Liang, X.i., Li, J., 2018. An efficient hyperspectral image retrieval method: deep spectral-spatial feature extraction with DCGAN and dimensionality reduction using t-SNE-based NM hashing. *Remote Sens.* 10 (2), 271. <https://doi.org/10.3390/rs10020271>.
- Zhang, J., Chen, L.u., Liang, X.i., Zhuo, L.i., Tian, Q.i., 2019. Hyperspectral image secure retrieval based on encrypted deep spectral-spatial features. *J. Appl. Rem. Sens.* 13 (01), 1. <https://doi.org/10.1117/1.JRS.13.018501>.
- Zhang, L., Chen, D., Ma, J., Zhang, J., 2020. Remote-sensing image superresolution based on visual saliency analysis and unequal reconstruction networks. *IEEE Trans. Geosci. Remote Sens.* 58 (6), 4099–4115.
- Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, Y. Fu, Image super-resolution using very deep residual channel attention networks, 2018. [Online]. Available: <https://arxiv.org/abs/1807.02758>.
- Zhang, X., Han, X., Li, C., Tang, X.u., Zhou, H., Jiao, L., 2019. Aerial image road extraction based on an improved generative adversarial network. *Remote Sensing* 11 (8), 930. <https://doi.org/10.3390/rs11080930>.
- Zhang, M.i., Hu, X., 2017. Translation-aware semantic segmentation via conditional least-square generative adversarial networks. *J. Appl. Rem. Sens.* 11 (04), 1. <https://doi.org/10.1117/1.JRS.11.042622>.
- Zhang, L., Li, W., Huang, H., Lei, D., 2021. A pansharpening generative adversarial network with multilevel structure enhancement and a multistream fusion architecture. *Remote Sensing* 13 (12), 2423. <https://doi.org/10.3390/rs13122423>.
- 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 Sensing* 13 (1), 128. <https://doi.org/10.3390/rs13010128>.
- Zhang, J., Shamsolmoali, P., Zhang, P., Feng, D., Yang, J., 2018. Multispectral image fusion using super-resolution conditional generative adversarial networks. *J. Appl. Rem. Sens.* 13 (02), 1. <https://doi.org/10.1117/1.JRS.13.022002>.
- Zhang, C., Shi, S., Ge, Y., Liu, H., Cui, W., 2020. DEM void filling based on context attention generation model. *IJGI* 9 (12), 734. <https://doi.org/10.3390/ijgi9120734>.
- Zhang, H., Song, Y., Han, C., Zhang, L., 2021. Remote sensing image spatiotemporal fusion using a generative adversarial network. *IEEE Trans. Geosci. Remote Sensing* 59 (5), 4273–4286.
- Zhang, X., Su, H., Zhang, C.e., Gu, X., Tan, X., Atkinson, P.M., 2021. Robust unsupervised small area change detection from SAR imagery using deep learning. *ISPRS J. Photogramm. Remote Sens.* 173, 79–94.
- Zhang, Y., Sun, H., Zuo, J., Wang, H., Xu, G., Sun, X., 2018. Aircraft type recognition in remote sensing images based on feature learning with conditional generative adversarial networks. *Remote Sens.* 10 (7), 1123. <https://doi.org/10.3390/rs10071123>.
- Zhang, Y., Li, X., Zhang, Q., 2019. Road topology refinement via a multi-conditional generative adversarial network. *Sensors* 19 (5), 1162. <https://doi.org/10.3390/s19051162>.
- Zhang, N., Wang, Y., Zhang, X., Xu, D., Wang, X., 2020. An unsupervised remote sensing single-image super-resolution method based on generative adversarial network. *IEEE Access* 8, 29027–29039. <https://doi.org/10.1109/ACCESS.2020.2972300>.
- Zhang, Y., Xiong, Z., Zang, Y.u., Wang, C., Li, J., Li, X., 2019. Topology-aware road network extraction via multi-supervised generative adversarial networks. *Remote Sensing* 11 (9), 1017. <https://doi.org/10.3390/rs11091017>.
- Zhang, L., Zhang, L., Du, B., 2016. Deep learning for remote sensing data: a technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* 4 (2), 22–40. <https://doi.org/10.1109/MGRS.2016.2540798>.
- Zhang, Z., Zhang, C., Wu, M., Han, Y., Yin, H., Kong, A., et al., 2021. Super-resolution method using generative adversarial network for Gaofen wide-field-view images. *J. Appl. Rem. Sens.* 15 (02) <https://doi.org/10.1117/1.JRS.15.028506>.
- Zhao, S., Yang, S., Gu, J., Liu, Z., Feng, Z., 2021. Symmetrical lattice generative adversarial network for remote sensing images compression. *ISPRS J. Photogramm. Remote Sens.* 176, 169–181.
- Zheng, K., Wei, M., Sun, G., Anas, B., Li, Y.u., 2019. Using vehicle synthesis generative adversarial networks to improve vehicle detection in remote sensing images. *IJGI* 8 (9), 390. <https://doi.org/10.3390/ijgi8090390>.
- Zheng, R., Wu, G., Yan, C., Zhang, R., Luo, Z.e., Yan, B., 2018. Exploration in mapping kernel-based home range models from remote sensing imagery with conditional adversarial networks. *Remote Sensing* 10 (11), 1722. <https://doi.org/10.3390/rs10111722>.
- Zhou, H.Y., Liu, Q.J., Wang, Y.H., 2021. PGMAN: an unsupervised generative multiadversarial network for pansharpening. *IEEE J. Selected Top. Appl. Earth Observat. Remote Sens.* 14, 6316–6327. <https://doi.org/10.1109/JSTARS.2021.3090252>.
- Zhou, C., Zhang, J., Liu, J., Zhang, C., Fei, R., Xu, S., 2020. PercepPan: towards unsupervised pan-sharpening based on perceptual loss. *Remote Sens.* 12 (14), 2318. <https://doi.org/10.3390/rs12142318>.
- Zhu, L., Chen, Y., Ghamisi, P., Benediktsson, J.A., 2018. Generative adversarial networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sensing* 56 (9), 5046–5063.
- Zhu, D.i., Cheng, X., Zhang, F., Yao, X., Gao, Y., Liu, Y.u., 2020. Spatial interpolation using conditional generative adversarial neural networks. *Int. J. Geogr. Inform. Sci.* 34 (4), 735–758.
- J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, 2017. [Online]. Available: <https://arxiv.org/abs/1703.10593>.
- Zhu, D., Xia, S., Zhao, J., Zhou, Y., Jian, M., Niu, Q., et al., 2020. Diverse sample generation with multi-branch conditional generative adversarial network for remote sensing objects detection. *Neurocomputing* 381, 40–51.
- Zou, Q., Ni, L., Zhang, T., Wang, Q., 2015. Deep learning based feature selection for remote sensing scene classification. *IEEE Geosci. Remote Sens. Lett.* 12 (11), 2321–2325. <https://doi.org/10.1109/LGRS.2015.2475299>.
- Zou, Z., Shi, T., Li, W., Zhang, Z., Shi, Z., 2020. Do game data generalize well for remote sensing image segmentation? *Remote Sensing* 12 (2), 275. <https://doi.org/10.3390/rs12020275>.