Journal of Scientific & Industrial Research
Vol. 83, June 2024, pp. 627-638
DOI: 10.56042/jsir.v83i6.7320

A Review and Analysis of GAN-Based Super-Resolution Approaches for INSAT 3D/3DR Satellite Imagery using Artificial Intelligence

S P Rajamohana[1], S Thamaraiselvi[2], Bibraj R[3] & Samir Mitha[4]
[1]Department of Computer Science, Pondicherry University, Karaikal 609 605, Puducherry, India
[2]EGS Pillay Engineering College, Nagapattinam 611 002, Tamil Nadu, India
[3]India Meteorological Department, MoES, Chennai 600 006, Tamil Nadu, India
[4]Electrical, Computer and Biomedical Engineering, Toronto Metropolitan University, Toronto, Ontario, Canada

Received 20 December 2023; revised 14 May 2024; accepted 24 May 2024

The Indian National Satellite System (INSAT)-3D/3DR is a geostationary satellite that is used for meteorological applications in the Indian region. Geostationary satellites have significant spatial coverage and good temporal resolution that help to monitor the evolution and propagation of meteorological systems. Meteorologists use satellite images to observe the locations of severe weather and understand the physical processes involved in the system. Image Super-Resolution (SR) aims to convert low-resolution images into high-resolution images while maintaining image quality. The SR techniques will improve the visualization of convective systems and tropical cyclones, facilitating accurate location-based warnings. This paper presents a comparative comparison of computer models for converting Low-Resolution (LR)(INSAT)-3D/3DR images into super-resolution images. This study also discusses and investigates the various Generative Adversarial Network (GAN)-based models, including the Super Resolution Generative Adversarial Network (SRGAN), Enhanced Super Resolution Generative Adversarial Network (ESRGAN), and Real Enhanced Super Resolution Generative Adversarial Network (Real-ESRGAN). The findings are compared to established approaches such as Bicubic Interpolation and Super-Resolution Convolution Neural Network (SRCNN). This study demonstrates that Real-ESRGAN performs better on weather satellite images than other cutting-edge approaches.

Keywords: Deep learning, Generative adversarial network, Meteorology, Remote sensing, Weather monitoring

Introduction

Super-Resolution (SR) has gained significant attention in the fields of computer vision and image processing due to its capability of transforming lower-resolution images into high-resolution ones.[1] Its versatility extends to various domains including satellite imagery, remote sensing, aerial photo evaluation, compressed imagery, security, medical imaging, and video surveillance.[1–3] During the last decade, various image processing techniques have been developed for image restoration, filtering, compression, segmentation, and interpolation.[3]

Meteorological agencies frequently utilize satellite imagery from polar orbiting and geostationary satellites for weather observation. Weather forecasts for India and neighboring nations are provided by the INSAT-3D and -3DR satellites, which are in geostationary orbit and continuously scan the Indian Ocean region from an altitude of 36,000 km. Effective launch dates for these satellites were July 26, 2013, and September 8, 2016, respectively.[4] The visible to thermal infrared region of the electromagnetic spectrum is covered by the 19-channel atmospheric sounders and 6-channel imagers that are featured in the INSAT-3D and INSAT-3DR series. Both satellites sensors operate at a spatial resolution of 1 km in the Visible (VIS) and Short-Wave Infrared (SWIR) bands, 4km for Mid-Wave Infrared (MWIR), Thermal Infrared 1 (TIR1), Thermal Infrared 2 (TIR2), and 8 km for Water Vapour (WV) bands.[4–6]

Weather satellites have observed meteorological phenomena such as monsoons, fog, thunderstorms, and tropical cyclones and have played an important role in weather monitoring and forecasting.[7] They have been widely used to understand tropical convection, from isolated cumulus systems to mesoscale convective systems and mesoscale convective complexes. Geostationary satellite imagery is one of the most essential tools for real-time monitoring of tropical

*Author for Correspondence
E-mail: spr.pukl2022@gmail.com; monamohanasp@pondiuni.ac.in

cyclones, although polar orbiting satellites have other complementary sensors.[8] Dvorak technique uses satellite imagery pattern matching and a classification scheme to assess the intensity of tropical cyclones. Tropical cyclone prediction agencies use this technique extensively to forecast tropical cyclones and issue accurate warnings to the public[9]. Super-resolution satellite images can enhance the visualization of various weather phenomena, thereby reducing the subjectivity in analysis. The three interpolation methods that are currently most frequently employed are bicyclic, bilinear, and nearest-neighbor. The fastest algorithm is the nearest-neighbour interpolation method.[10] Its rapid speed is a plus, however it can cause a lot of distortion. Compared to nearest-neighbour interpolation, bilinear interpolation is more complicated and requires more computing.[11] Unlike bilinear interpolation, which only takes into account 4 pixels (2 × 2), bicubic interpolation takes into account 16 pixels (4 × 4). Smoother and less distorted images resemble those produced by bicubic interpolation. However, traditional image enhancement algorithms cannot recover high-frequency details and produce poor quality results.[12]

Much progress has been made in images SR and reconstruction techniques as a result of deep learning. Deep Neural Networks (DNNs) have significantly improved performance over traditional methods.[13] Several works have been done on single image SR problems such as SRCNN[14], SRResNet[14], EDSR[15], and WDSR.[16] SRCNN uses a convolution neural network to discern relational mapping between images of low and high resolution. Its inception initiated a revolutionary transformation in the realm of image super-resolution by being the first to incorporate deep learning techniques.[17,18] A notable drawback of conventional CNNs is their limited capacity to revive intricate image textures, especially with significant upscaling levels.[19] Recognizing this, GANs emerged as a solution to image super-resolution conundrums. SRGAN uses a generative adversarial network to enhance low-resolution images to a higher quality. By avoiding the SRCNN framework, SRGAN unveiled the groundbreaking SRResNet generative architecture. Additionally, SRGAN adopts a perceptual loss function to counteract the excessive smoothening encountered when using the MSE loss function. The concomitant implementation of adversarial and content losses in SRGAN enables adaptable modifications to the weights. Based on the progress of SRResNet, the EDSR network architecture was established. To optimize memory usage and produce superior results, EDSR relies on the SRResNet generation mechanism in SRGAN, omitting its foundational BN layer. Meanwhile, WDSR, another super-resolution model, substitutes BN with weight normalization. Both methods extensively use numerous training samples to learn the mapping relationship between lower and higher resolution images.[20]

Related Works

Remote sensing data pose unique challenges for traditional machine learning methods due to the stringent requirements for image resolution, edge feature accuracy, and spectral information As artificial intelligence has advanced, numerous super resolution techniques based on deep learning have been developed to tackle these problems.[21,22] This section delves into an extensive review of GAN-based image super resolution methodologies and their relevant works.

The notion of Generative Adversarial Networks (GANs) was presented by Goodfellow.[23] This novel neural network framework involves simultaneously training two networks: a generating model and a discriminative model. The generative model, also known as the generator, takes in low-quality images with the goal of creating synthetic Super-Resolution (SR) images that are similar to the authentic SR images contained in the training dataset.[24] Meanwhile, the discriminator is being trained to discriminate between real images and those produced by the generator. In the same year, Mirza *et al.* developed the idea of conditional GANs.[25] This extension of the GAN architecture incorporates additional information, such as class labels, that is associated with the input images. These supplementary data are integrated into both the Generator and Discriminator networks during training. Unlike traditional GANs, cGANs offer users the ability to dictate the type of output.

Later, in 2015, Ronneberger *et al.*[26] conceptualized a U-shaped framework, dubbed U-Net, which encompasses encoder and decoder modules fortified with skip connections. [26–27] This design ensures the encoding and decoding of the input to yield an output that mirrors the resolution of the initial input.

Li and Wand *et al.*[28] proposed a PatchGAN technique, which cuts the image into small overlapping patches and discriminates them individually.[25–28] It is technically identical to blending data after cutting it into patches and is useful for data with a uniform distribution. Ledig *et al.*[29] proposed SRGAN, which produces images with rich texture and good quality.[26]

The SRGAN uses a Low Resolution (LR) input image to generate a High Resolution (HR) output image at a magnification of 4x. The SRGAN's architecture consists of a deep convolution network with residual blocks. The SRGAN approach uses a perceptual loss as its objective function, which includes both adversarial and content loss. The feature loss is calculated as the difference between the generated image's feature map (obtained from the pre-trained VGG19 network) and the real image. The GANs use skip connections in deep residual networks (ResNet). In SRGAN, the Mean Squared Error (MSE) is used to calculate the similarity between feature maps. Furthermore, this model uses the VGG network's high-level feature map to define a novel perceptual loss, which is paired with a discriminator to provide a solution that is perceptually indistinguishable from the HR reference image. Wang *et al.*[30] presented an ESRGAN to improve the visual quality outcome of SRGAN.[28] To achieve this, they have made three changes to the architecture of SRGAN. The author has removed the BN layer and inserted RDDB blocks, which helps to improve the training process. Relativistic GAN (RGAN) is introduced as a discriminator to improve classification.

A generative adversarial network (GAN)-based edge-enhancement network (EEGAN) was introduced by Jiang *et al.* in 2019 to tackle the issue of recovering high-frequency edge details in noise-contaminated imaging conditions while ensuring robust satellite image super resolution reconstruction.[31] The objective of this paper is to rectify the shortcomings of existing deep learning-based super resolution techniques when it comes to capturing boundary details in remote sensing satellite imagery. An Ultra-Dense Subnetwork (UDSN) is incorporated into the proposed EEGAN framework to facilitate feature extraction, while an Edge-Enhancement Subnetwork (EESN) is utilized to eliminate noise-contaminated components and improve image contours. Extensive experiments on a variety of datasets—including the Kaggle Open Source Data set, Jilin-1 video satellite images, and Digitalg lobe are conducted to illustrate the EEGAN framework's superior reconstruction performance in comparison to current-state super resolution methods.

In 2020, Wang *et al.* introduced the udGAN model, which integrates adversarial learning and Ultra-Dense Residual Blocks (UDRB).[32] This model demonstrates superior reconstruction performance compared to established CNN and GAN-based super-resolution methods, including direct and skip-connection-based approaches. The use of ultra-dense connections in the GAN generator, as introduced in the udGAN model, leads to improved feature expression and more photo-realistic reconstruction outcomes. The proposed udGAN algorithm demonstrates competitive reconstruction results even when the image degradation model is unknown, compared to other GAN algorithms. This paper introduces dense connections to GAN generators and highlights the novelty that leads to the development of a new GAN variant called udGAN.[32]

SWCGAN, a Generative Adversarial Network (GAN) introduced by Tu *et al.*[33], integrates the benefits of the convolution and Swin Transformer layers to enhance the resolution of remote sensing images. In relation to reconstruction performance, the suggested approach surpasses other contemporary methods, as evidenced by its assessment on the UCMerced benchmark dataset and authentic remote sensing images. When comparing the SWCGAN method to other GANs, it outperforms the latter on the LPIPS metric and the majority of other metrics. Additionally, an actual satellite remote sensing image is processed using SWCGAN, which yields an even higher quality of image in comparison to the other models.[33]

Multi-Attention Generative Adversarial Network (MA-GAN) for remote sensing image super-resolution was proposed by Jia *et al.*[34] in 2022. Using deep learning techniques, MA-GAN effectively learns the mapping from Low-Resolution (LR) to High-Resolution (HR) images and generates HR images that are perceptually pleasing. A GAN-based framework is presented for the image SR task, with an emphasis on the post-up sampling image generator. The MA-GAN framework incorporates a generator that utilizes Attention-based Up Sampling (AUP) and residual-dense block (PCRDB) pyramidal convolution to produce images of superior quality. The effectiveness of the proposed MA-GAN is demonstrated by the fact that it outperforms a number of state-of-the-art super-resolution methods on a dataset of remote sensing scene images.[34]

In conclusion, perceptual loss was enhanced by employing the pre-trained VGG-19 network. Conventional approaches to image SR, including SRGAN, CGAN, Patch GAN, and ESRGAN, depend on the unreliable Bicubic kernel interpolation.[29] These models fail to capture high-frequency details and produce images that are overly smooth. A recent study by Real-ESRGAN introduces a novel technique known as blind SR, which simulates the degeneration of real-world images with greater accuracy by employing a high-order degradation model (HDM).[30]

The objective of this research endeavor is to conduct an exhaustive assessment of five notable SR methodologies: Bicubic, SRCNN, SRGAN, ESRGAN, and Real-ESRGAN. The evaluation encompasses a diverse range of image datasets, which comprise Visible (VIS), Short Wave Infrared (SWIR), Water Vapour (WV), multiple infrared categories, and multispectral (RGB) images. By employing the Structural Similarity Index Measure (SSIM) and the Peak Signal-to-Noise Ratio (PSNR) as performance metrics, this analysis seeks to ascertain the approach that maintains the structure while ensuring optimal clarity.

The primary contributions that this work makes are as follows.

- Conducted a rigorous and detailed comparison of five leading super-resolution techniques across seven distinct image datasets, providing a holistic view of their relative performances.
- Presented empirical evidence supported by quantitative metrics, including PSNR and SSIM, in order to offer unbiased and transparent assessments of the effectiveness of each technique.
- Identified Real-ESRGAN's consistent superior performance in terms of PSNR across all image categories, marking a significant contribution to the domain of image upscaling.
- Revealed the varied landscape of SSIM values among the techniques, emphasizing the importance of structural preservation and the competitive nature of methods such as SRGAN.
- Highlighted that while Real-ESRGAN is generally top-notch, the choice of technique should be influenced by the specific requirements of the application, especially when structural preservation is paramount.
- Employed a methodical approach in designing the study and analyzing the data, guaranteeing that the findings may be replicated and used as a standard for future research in this field.
- The examined techniques were shown to be applicable and their performance varied across several types of images, such as visible, infrared, and multispectral. This makes them suitable for a wide range of real-world applications.

Analysis of Image Super Resolution Methods

Super Resolution Convolution Neural Network (SRCNN)

The architecture of SRCNN consists of three convolutional layers designed to generate high-resolution imagess.[35] The SRCNN framework consists of three components: feature extraction, non-linear mapping, and image reconstruction. These parts perform a series of steps in a certain order. First, they extract characteristics from inputs that have low resolution. Then, they convert these characteristics into features with high resolution. Finally, they reconstruct images with high resolution. The initial layer applies a 9 × 9 kernel to the input, taking into account padding, and converts the 3-channel image into 64 separate feature maps. The next layer utilizes a 1 × 1 kernel to compress the data into 32 feature maps. The last layer employs a 5 × 5 kernel to generate the resulting image. The ReLU activation function follows both the initial and middle convolution layers. The training objective of the model is to minimize the Mean Squared Error (MSE) between the reformed images and the actual images on a pixel-wise basis. We employed the Mean Squared Error (MSE) to train and validate the model's loss function. Furthermore, in order to assess the quality of image restoration, the Peak Signal-To-Noise Ratio (PSNR) was used as a benchmark metric.

Super Resolution Generative Adversarial Network (SRGAN)

Super-resolution refers to the application of diverse approaches to enhance or improve the quality of an image. Traditionally, SRCNNs have been employed for the purpose of producing high-resolution images that exhibit rapid training and reach a high level of accuracy. Nevertheless, in certain instances, it may not be possible to retrieve the intricate particulars, and it is frequently necessary to enhance the image's clarity. Hence, in this study, we utilized the Super Resolution Generative Adversarial Network (SRGAN) developed by Ledig *et al.*[29] The SRGAN architecture bears resemblance to the GAN architecture, as depicted in Fig. 1. Furthermore, alongside the SR model, it incorporates a perceptual loss function that effectively transforms low-resolution images into high-resolution images while preserving a significant amount of information. The system is composed of a generation network and a discrimination network. Generators have the task of producing data with a high level of detail, whereas discriminators are employed to identify counterfeit data. The generator architecture consists of a fully convolutional SRResNet model designed to produce high-quality, high-resolution images. By using a discriminator model as an image classifier, the entire architecture is able to adjust to image quality, resulting in a significant enhancement of the final image.

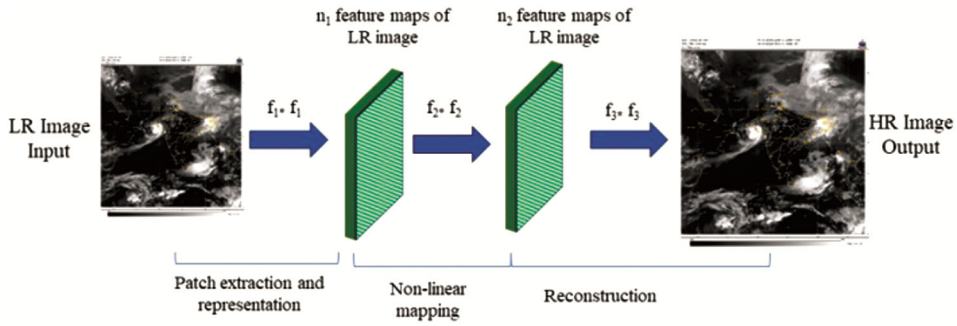

Fig. 1 — Architecture of SRCNN proposed by Dong *et al*. (2015)[35]

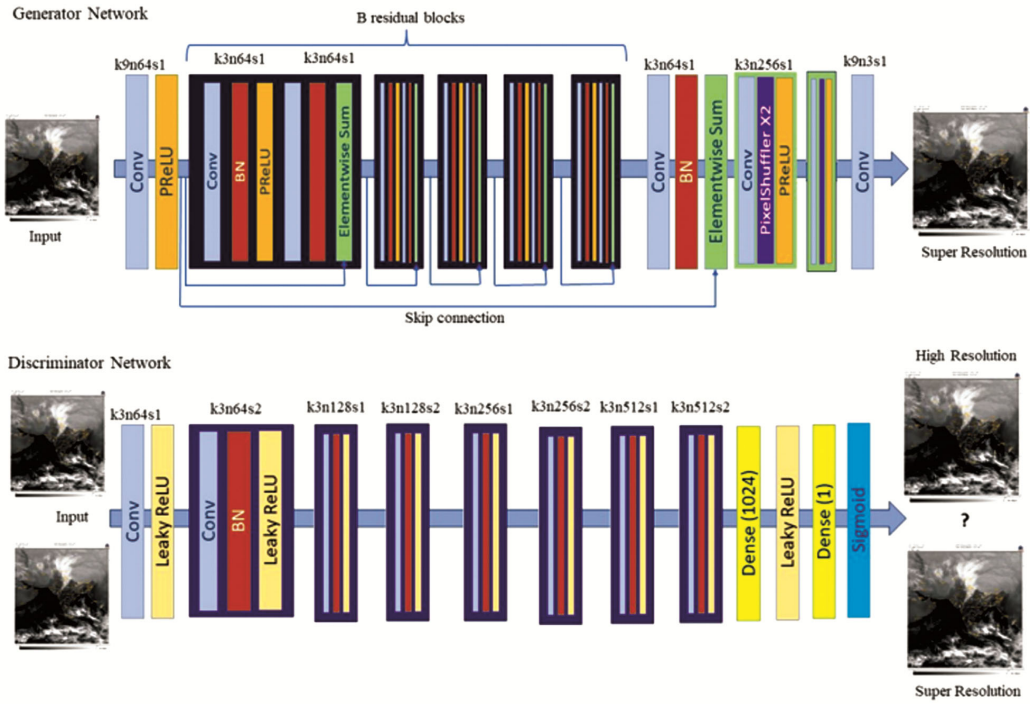

Fig. 2 — Architecture of SRGAN proposed by Ledig *et al*. (2017)[29]

The SRResNet generating network utilizes a low-resolution image input which undergoes a preliminary convolution layer with 9 kernels and 64 channels in a single phase. The fully convolutional feed forward SRResNet model consists of several layers, including a residual block with a convolution layer, 64 feature mappings, batch normalization, a ReLU parametric activation function, another convolution layer with batch normalizing, and the element-by-element sum approach is utilized. The element-by-element summing approach utilizes the feed forward output and the skip connection output as the ultimate output. A pixel shuffler is employed following the 4x upsampling of the convolution layers to provide images with enhanced resolution.

Discriminator: An image classification model distinguishes between genuine and counterfeit images. The discriminator receives two inputs: the generated images from the generator and the real high-resolution image. The developed discriminator model is designed to address the adversarial Min-Max problem.

The loss function of SR incorporates an adversarial loss and a content loss, which enhances the VGG network's high-level feature mappings to be less affected by variations in the pixel space, as depicted in Fig. 2.

Enhanced Super-Resolution Generative Adversarial Network (ESRGAN)

The Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) technology is developed to convert low-resolution images into high-resolution images while addressing the issue of

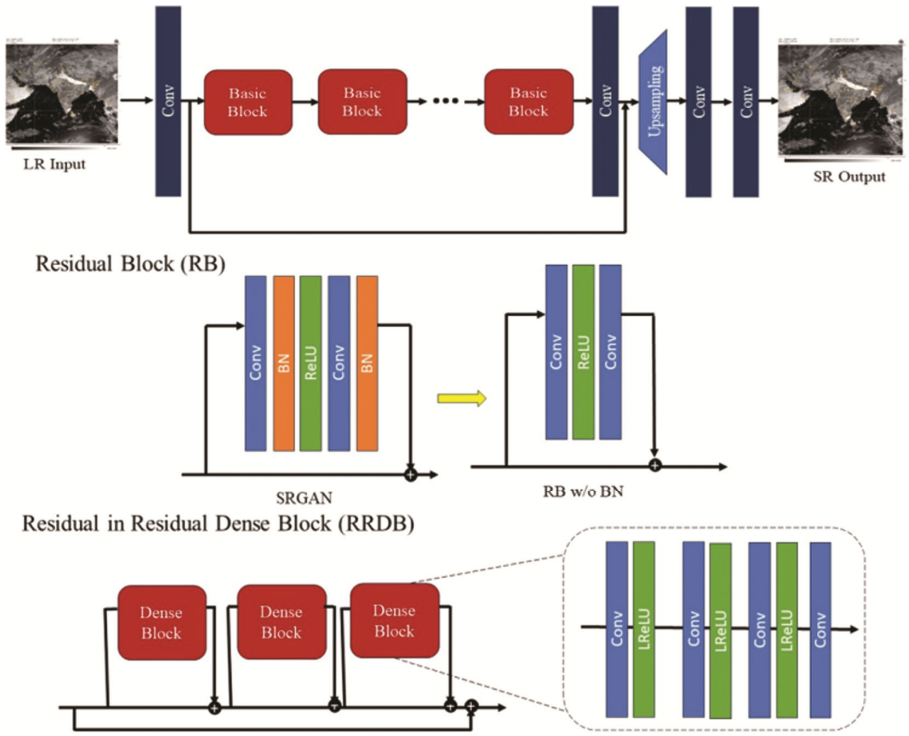

Fig. 3 — Architecture of ESRGAN proposed by Wang *et al*. (2018)[30]

unwanted artifacts.[30] The research primarily investigates three fundamental aspects of SRGAN: the structure of the network, the adversarial loss function, and the perceptual loss function, which enhances the visual attractiveness of the output. ESRGAN is a method that is guided by perception, although the use of perceptual loss can occasionally lead to uneven brightness between the reconstructed image and the original. In order to address this issue, the authors developed a novel foundation network unit called Residual Dense Block in Residual (RRDB), which eliminates the use of batch normalization. The discriminator has been improved by implementing a relativistic average GAN (RaGAN), which is taught to evaluate the realism of images, rather than only its authenticity. This enhancement allows the generator to accurately reproduce finer details in the resulting images. The enhanced perceptual loss leverages the VGG feature prior to activation, providing more effective guidance for preserving brightness consistency and restoring texture.[30]

ESRGAN_Generator: Initially, the Residual-in-Residual Dense Block (RRDB), which serves as the main construction block of the network, was launched without the utilization of batch normalization. Removing batch normalization layers from the SRGAN and incorporating the residual-in-residual dense block layers, as depicted in Fig. 3.

The ESRGAN discriminator utilizes relativistic average GAN to determine the relative realism of two images, rather than simply discerning if an image is real or fake. The Relativistic Average GAN discriminator evaluates the likelihood that a real image is more plausible than a counterfeit one, rather than determining the probability that an input is a genuine image.

Objective function: In order to have more precise control over the brightness and texture of an images, it is recommended to assess the perceptual loss using the VGG features before the activation utilized in the SRGAN.

Real Enhanced Super-Resolution Generative Adversarial Network (Real-ESRGAN)

Wang *et al.*[36] introduced Real-ESRGAN in their investigation. The primary objective of this super-resolution model is to address a wider spectrum of image degradation compared to the traditional ESRGAN. The initial design of the discriminator in ESRGAN is inadequate for Real-ESRGAN due to its insufficient ability to distinguish complex training outputs. The discriminator in Real-ESRGAN is responsible for providing accurate gradient feedback for complex local textures and distinguishing global styles. To address this difficulty, the VGG-style discriminator

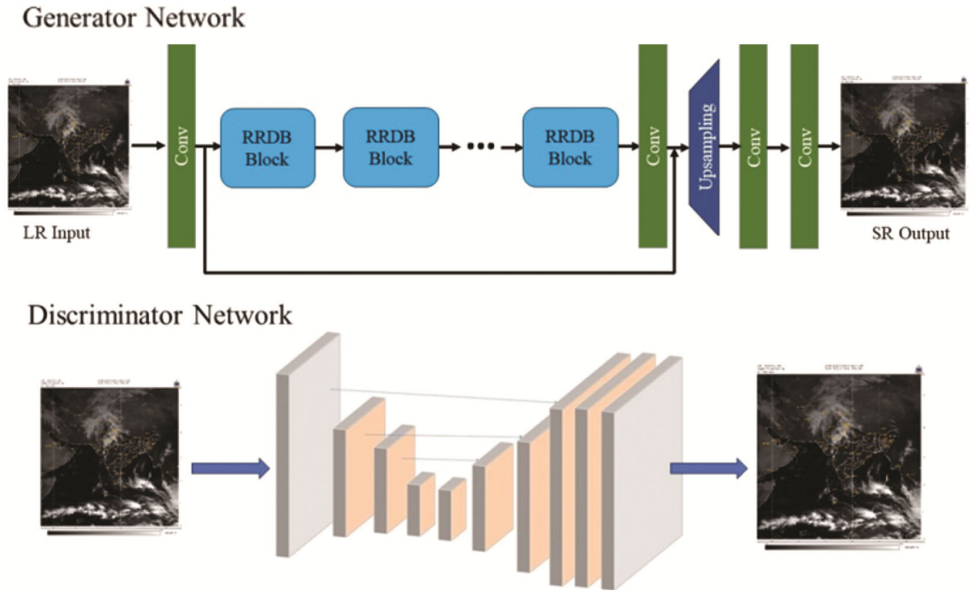

Fig. 4 — Architecture of Real-ESRGAN proposed by Wang *et al*. (2022)[36]

in ESRGAN has been improved by adopting a U-Net architecture that includes skip connections. The U-Net configuration generates authenticity values for each individual pixel, providing the generator with specific feedback for each pixel. The inclusion of skip connections in the U-Net design enables the discriminator to effectively collect both the overall and specific characteristics of the image, which are crucial for successfully distinguishing between genuine and counterfeit images. The enhanced discriminator in Real-ESRGAN offers more precise feedback to the generator, hence enabling it to generate high-quality super-resolved images with reduced artifacts.

Real-ESRGAN_Generator: As seen in Fig. 4, they employed the same generator as ESRGAN. Before feeding the input into the Real-ESRGAN design, we used the pixel-unshuffled approach to decrease the model, computational complexity, and memory use because ESRGAN is a heavy network. It facilitates increasing the channel size and decreasing the spatial resolution size.

Real-ESRGAN_Discriminator: To reduce the over-sharpness and undesirable artifacts produced by ESRGAN, a U-Net discriminator with spectral normalization regularization was used in place of the original VGG-style discriminator.

Loss Function: L1 loss, perceptual loss, and GAN loss {1,1,0.1} are combined in Real ESRGAN. Before activating in the pre-trained VGG19 network, we use the {conv1,...conv3} feature maps (with weights {0.1, 0.1, 1, 1, 1} as the perceptual loss.

Image Quality Assessment

The aim of IQA is to precisely and automatically forecast the quality of an image. It emphasizes the visually important characteristics of images.[35–37] Image Quality Assessment (IQA) can be categorized into two types: Objective image quality assessment (O-IQA) and Subjective image quality assessment (S-IQA). S-IQA refers to the evaluation of human viewers based on their perception. O-IQA relies on computational models for its foundation.

A human observer's subjective assessment results in a Mean Opinion Score (MOS). This score approach was formerly employed as a subjective quality metric for image quality assessment (IQA).[38] Nevertheless, the process of subjective judgment is both laborious and costly, making it impractical for real-time situations. Hence, it is imperative to have a systematic approach for evaluating the standard of quality. Objective Image Quality Assessment (O-IQA) approaches yield measurable and comparable outcomes, but they may not precisely reflect human visual perception.[37] The reference images quality assessment methods can be categorized into full reference, reduced reference, and non-referenced categories based on the level of information supplied, as explained earlier.

Objective IQA Methods

This subsection presents a concise summary of three separate methodologies employed in the evaluation of image quality. The Full Reference (FR) technique provides a highly thorough method where test images

are directly compared to fully accessible reference images, sometimes referred to as the ground truth. By directly comparing, we can thoroughly evaluate the quality of the test image. In contrast, the Reduced Reference method lacks the ability to obtain a complete reference image. Instead, it depends on a limited perspective, utilizing specific characteristics of the reference image to assess the quality of the test image. An advantage of this strategy could be its usefulness in situations when it is not practical to store or access the complete reference image. The final strategy, known as No Reference (NR), is particularly fascinating as it deviates entirely from the utilization of reference photographs. The NR technique is assessed exclusively on the inherent characteristics of the test image, without relying on reference images that may not be available in real-world situations. The selection of the appropriate strategy is contingent upon the particular circumstances and the available resources, as each approach possesses its own advantages.

Peak Signal-to-Noise Ratio (PSNR)

This metric is frequently employed in image processing according to Wang *et al.*[36] The PSNR values are computed by utilizing the mean squared error (MSE) as defined in Eq. 2. The total error is computed by comparing the original low-resolution image with the reconstructed high-resolution image, where higher values indicate superior image quality.

$$MSE(r,g) = \frac{1}{MN}\sum_{i=1}^{M} \quad \sum_{j=1}^{N} \quad \left(r_{ij}, g_{ij}\right)^2 \qquad \ldots (1)$$

The variables M and N denote the number of rows and columns in the input image, while r represents the low-resolution input image and g represents the image generated during the super-resolution process. The Mean Squared Error (MSE) value obtained from Eq. 1 is utilized in Eq. 2.

$$PSNR = 10\ log\ log\ 10\ \left(\frac{Max_f}{\sqrt{MSE}}\right) \qquad \ldots (2)$$

Here, Max_f is a term used to represent the maximum possible pixel value in the input image.

Structural Similarity Index (SSIM)

SSIM is a quantitative measure employed to assess the structural similarity of two images in terms of their structure, brightness, and contrast. The calculation involves utilizing the mean, variance, and covariance of the images in combination. The SSIM value is a numerical measure that varies from 0 to 1. A value of 0 indicates that the images are not identical, while a value of 1 indicates that the images are same. A greater value of Structural Similarity Index (SSIM) signifies superior image quality.

$$l_x(r,f) = \frac{2\mu_r\mu_f + con_1}{\mu^2{}_r + \mu^2{}_f + con_1} \qquad \ldots (3)$$

$$C_x(r,f) = \frac{2\sigma_r\sigma_f + con_2}{\sigma^2{}_r + \sigma^2{}_f + con_2} \qquad \ldots (4)$$

$$S_x(r,f) = \frac{2\sigma_{rf} + con_3}{\sigma_r + \sigma_f + con_3} \qquad \ldots (5)$$

In the Eq. 4 measures the luminance closeness of the two images. They are expressed as the average values r and f of two images. If r = f, then both images have the same mean and mean luminance equal to 1.

In the Eq. 5 measures the contrast closeness of two images. It is represented by the standard deviation values r and f of two images. If r = f, both images have equal standard deviation values and a contrast standard deviation value 1.

It is represented by the co variance value rf between the two images. It measures the correlation coefficient between two images determined using Eq. 6.

If r = f, both images have equal co variance values and structural co variance values equal to 1. Finally, the SSIM is a weighted average of brightness, contrast, and texture values determined using Eq. 7

$$7SSIM(r,f) = l_x(r,f)C_x(r,f)S_x(r,f) \qquad \ldots (6)$$

Experimental Results and Discussions

The INSAT-3DR imager is an advanced multispectral optical radiometer with the ability to capture images in six distinct bands: Visible (VIS), Shortwave infrared (SWIR), Mid wave infrared (MWIR), Water Vapour (WV), and Thermal infrared-1 and -2 (TIR-1/2). The images obtained from the INSAT-3D / 3DR bands are of great importance in analytical research.[35–37] There are six unique bands in INSAT 3D

1) Visible (0.55–0.75 μm): The band is of a reflecting nature; hence its utility is restricted to daytime. The images are derived from the target's albedo. Clouds with a higher albedo than the land surface seem brighter than the land. This band is utilized for monitoring meso-scale meteorological characteristics.

2) Shortwave-infrared (1.55–1.70 μm): The band is of a reflective nature and has limited functionality during daylight hours. Water, ice, and snow have a high absorption rate for radiation, while clouds reflect

Table 1 — Comparison of quantitative results of PSNR(dB) and SSIM on Water vapour, Visible, Infrared and Multispectral image datasets

Images	Bicubic		SRCNN		SRGAN		ESRGAN		Real-ESRGAN	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Visible	30.5	0.9	29.9	0.9	28.0	1.0	29.7	1.0	32.5	1.0
SWIR	30.2	0.9	29.9	1.0	28.0	1.0	28.0	0.9	32.1	1.0
Water Vapour	33.1	0.9	32.2	1.0	27.8	1.0	27.9	1.0	35.1	1.0
MIR	30.5	0.9	32.1	1.0	28.1	1.0	28.1	1.0	34.4	1.0
IR1	31.3	0.9	31.8	1.0	27.9	1.0	27.9	1.0	35.0	1.0
IR2	31.8	0.9	32.2	1.0	27.9	1.0	27.9	1.0	35.0	1.0
Multispectral	29.9	0.8	31.7	1.0	27.9	1.0	27.9	1.0	31.9	1.0

it. Therefore, it is employed to distinguish between the cloud, rain-producing cloud, and snow. This band is utilized for the purpose of monitoring the snow cover, daytime fog, and the radiative properties of clouds in the nearby area.

3) M id infra-red (3.9) μm: It has greater sensitivity compared to the thermal infrared. The water droplets in fog can be distinguished from the land or sea surface at the same temperature due to a difference in emissivity. Due to the modulation of the band by reflected sunlight, the daytime image seems warmer compared to the night time image.

4) Water vapour Image (6.5–7.1 μm):

This band belongs to the infrared spectrum and is primarily absorbed by water vapour, making it the dominant gas for absorption. The satellite receives radiation from the mid and upper troposphere, resulting in a limited depiction of the lower half of the troposphere in WV Imagery.

5) Thermal infra-red 1(10.3–11.3 μm), At around 10.7 μm, most of the radiation from the surface is seen in the image as this region is called the atmospheric window where no absorption takes place in the atmosphere. This band is utilized for the purpose of monitoring the temperature of both the top layer of clouds and the Earth's surface. Thermal infrared images exhibit distinct contrast among clouds at varying altitudes. It is applicable for use in both daytime and night time.

6) Thermal infra-red 2 (11.5–12.5 μm). This band is sensitive to low levels of water vapour and is therefore utilized for the purpose of identifying moisture in the lower troposphere.

The imaging radiometer detects both radiant and sunlight-reflected energy from the target area. It employs a scanning mirror positioned on a dual-axis gimbal for this task. The resolution for the visible band and SWIR stands at 1 km, while the infrared channel boasts a resolution of 4 km, and the water vapour channel has a resolution of 8 km. This device can capture an entire image of the Earth in just 26 minutes. The multispectral images are generated from the individual bands as RGB composites. The final images of different channels are generated in the normal-resolution and high-resolution modes. Each mode has different pixel resolutions for different bands. The machine learning-driven methods SRCNN, SRGAN, ESRGAN, and Real ESRGAN outperform the Bicubic interpolation approach in varied weather scenarios, including water vapour, visible, and infrared conditions.[27,39] Experimentally, Real SRGAN exhibits superior performance relative to other SRGAN techniques. The robustness of these experimental results was assessed using metrics such as MSE, PSNR, and SSIM. Quantitative data can be viewed in Table 1, which establishes Real ESRGAN as the leading standard in Super-Resolution (SR) for photo realistic imagery.[36]

The high-resolution images provide clear delineation of the boundaries of the clouds, enabling the identification of their pattern and shape. Additionally, the normalizing of the pixels is observed, resulting in the reduction of slight variances among the cloud mass. This can potentially cause inaccurate assessment of the intensities. This phenomenon is more pronounced in black-and-white bands compared to multispectral (RGB) images. The utilization of super-resolution images enhances the distinction between clouds at varying altitudes, hence facilitating their categorization. This research primarily focuses on using an adversarial learning-based super-resolution approach to satellite images in order to improve image resolution. The effectiveness of this method is measured using PSNR (power signal-to-noise ratio) and SSIM (structural similarity index) metrics. The results demonstrate an increase in image sharpness without sacrificing quality. This section presents a combination of visual and numerical information.

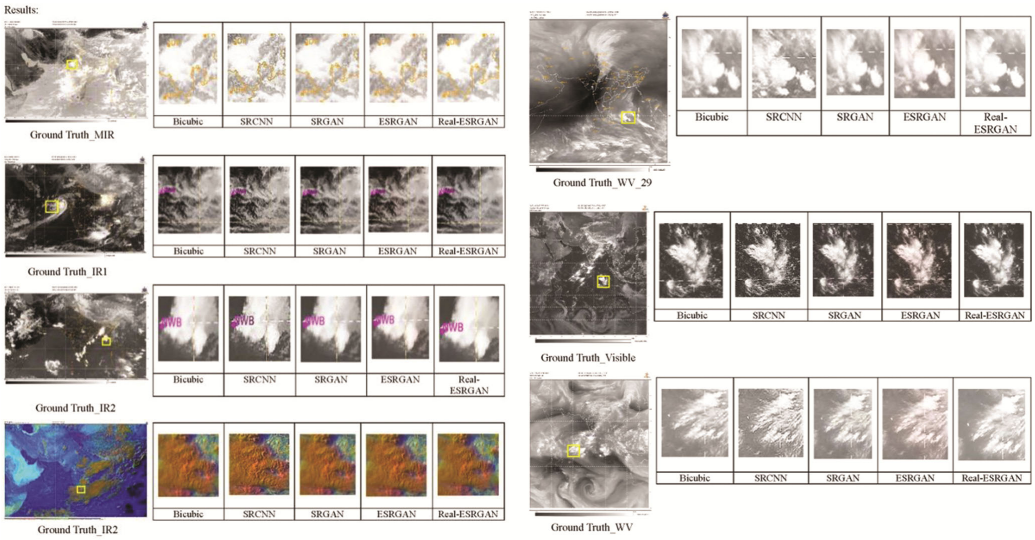

Fig. 5 — Visual Comparison of the results of 4x Up scaling Super Resolution for Bicubic, SRCNN, SRGAN, ESRGAN, Real-ESRGAN and the satellite images of VIS, SWIR, WV

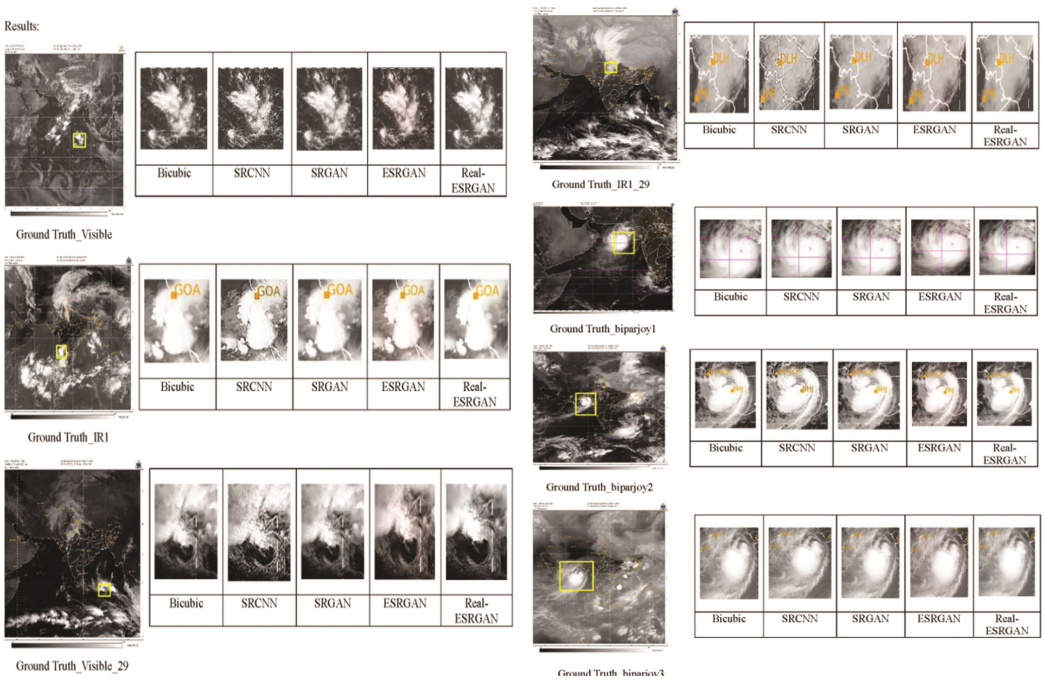

Fig. 6 — Visual Comparison of the results of 8x upscaling Super Resolution for Bicubic, SRCNN, ESRGAN, Real-ESRGAN and the satellite images of MIR, TIR1, TIR2, Multispectral

A detailed comparison of alternative image super resolution approaches, with a focus on their performance on different types of image datasets is given in Table 1. This performance is measured using the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index Measure (SSIM). PSNR and SSIM measures have been employed to assess the efficacy of five distinct super-resolution (SR) algorithms, namely Bicubic, SRCNN, SRGAN, ESRGAN, and Real-ESRGAN, across a range of image types. Greater PSNR and SSIM values imply superior image quality after super-resolution. Real-ESRGAN has proven to be the most effective method for enhancing visual images, as it achieved the highest PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index) values of 32.464 dB and 0.9837 respectively. However, SRGAN had the lowest PSNR value of 27.990 dB, despite its relatively high SSIM of 0.9741. Real-ESRGAN also performed well on

Short-Wave Infrared (SWIR) images, with PSNR and SSIM values of 32.053 dB and 0.9887 respectively, similar to its performance on visible images. SRGAN had the lowest PSNR, while ESRGAN had the lowest SSIM of 0.9491. Real-ESRGAN also outperformed other techniques in the category of water vapour, registering the highest PSNR (35.072 dB) and a competitive SSIM of 0.9834. SRGAN, while presenting a lower PSNR, showed impressive SSIM values of 0.9771, almost on par with ESRGAN. For Medium Infrared (MIR), IR1, and IR2 images, Real-ESRGAN consistently displayed the highest PSNR values, emphasizing its robustness across infrared datasets. The SSIM values for SRGAN were impressively high, especially for IR1 and IR2, challenging the dominance of Real-ESRGAN in SSIM. In the multispectral domain, SRCNN exhibited a competitive performance with an SSIM value of 0.9802, although Real-ESRGAN still led the PSNR metric at 31.862 dB. The Bicubic method showed the lowest SSIM value of 0.7940, indicating potential structural differences in the upscaled images compared to the original is shown in Figs 5 & 6.

Real-ESRGAN consistently outperformed other methods in terms of PSNR in all image categories, indicating its superiority in preserving image clarity during super resolution. However, when considering the SSIM metric, which emphasizes structural preservation, the competition becomes tighter, especially with SRGAN showing competitive values in several categories. This highlights the importance of considering multiple metrics and real-world applications when evaluating the effectiveness of super-resolution methods.

Conclusions

This paper compares the effectiveness of high-resolution techniques as tools to improve low-resolution satellite imagery. Comparative analysis of five image interpolation algorithms such as Bicubic, SRCNN, SRGAN, ESRGAN, and Real-ESRGAN are implemented. The Real-ESRGAN technique provides the best trade-off between visual quality and computing complexity for real-time image interpolation applications out of the five image interpolation algorithms tested. Though there is an improvement in terms of metrics in Real-ESRGAN based upscaling, challenges still remain in operational classification of synoptic weather systems with up scaled images. Visual interpretation of shape, size, pattern and texture of the clouds are an important aspect in determination of the classification. Upscaling of images may likely lead to smoothening which can affect the pattern and texture of clouds leading to different interpretations. More research needs to be undertaken to identify the improvement in the accuracy of the derived meteorological products from up scaled images with the ground truth. As satellite based remote sensing has a wider spatial coverage and availability, increase in the reliability of the meteorological products derived from high resolution data can have a greater impact on weather forecasting.

Acknowledgement

The authors would like to thank their institutions for their support in undertaking the research.

References

1 Agustsson, E & Timofte, R, NTIRE 2017 Challenge on single image super-resolution: Dataset and study, *IEEE Comp Soc Conf Comp Vis Pattern Recog Workshops*, **2017** (2017) 1122–1131, doi: 10.1109/cvprw.2017.150.

2 Yue L, Shen H, Li J, Yuan Q, Zhang H & Zhang L, Image super-resolution: The techniques, applications, and future, *Signal Proc*, **128** (2016) 389–408, doi:10.1016/j.sigpro.2016.05.002.

3 Han L, Zhao Y, Lv H, Zhang Y, Liu H, Bi G & Han Q, Enhancing remote sensing image super-resolution with efficient hybrid conditional diffusion model, *Remote Sens*, **15(13)** (2023), doi: 10.3390/rs15133452.

4 Kumawat N & Babu K N, INSAT-3D and -3DR relative radiometric performance over bright target, *Remote Sens Appl Soc Environ*, **25** (2022) 100672–100672, doi:10.1016/j.rsase.2021.100672.

5 Patel P N, Babu K N, Prajapati R P, Sitapara V & Mathur A K, Day-1 INSAT-3DR vicarious calibration using reflectance-based approach over great rann of kutch, *J Indian Soc Remote Sens*, **46(6)** (2018) 885–894, doi: 10.1007/s12524-017-0729-z.

6 Lim A H N N, Jung S E, Daniels J A, Bailey J M, Bresky W, Bi L & Mehra A, Optimizing the assimilation of the goes-16/-17 atmospheric motion vectors in the hurricane weather forecasting (HWRF) Model, *Remote Sens*, **14(13)** (2022), doi: 10.3390/rs14133068.

7 Umakanth N, Satyanarayana G C, Naveena N, Srinivas D & Rao D V B, Statistical and dynamical based thunderstorm prediction over southeast India, *J Earth Syst Sci*, **130(2)** (2021), doi: 10.1007/s12040-021-01561-x.

8 Meng F, Song T & Xu D, TCR-GAN: Predicting tropical cyclone passive microwave rainfall using infrared imagery via generative adversarial networks, *http://arxivorg/abs/220107000*, (2022), https://doi.org/10.48550/.

9 Meng F Song T & Xu D, Simulating tropical cyclone passive microwave rainfall imagery using infrared imagery via generative adversarial networks, *IEEE's-GRSS*, **19** (2022), doi: 10.1109/lgrs.2022.3152847.

10 Vayadande K, Adsare T, Dharmik T, Agrawal N, Patil A & Zod S, in *Cyclone Intensity Estimation on INSAT 3D IR Imagery Using Deep Learning* (IEEE) 2023, 592–599.

11 Gao C, Zhou R G, Li X & Li Y C, Asymmetric scaling of a quantum image based on bilinear interpolation with arbitrary

scaling ratio, *Quantum Inf Process*, **21(7)** (2022), doi: 10.1007/s11128-022-03612-8.
12 Sharma R P & Malik A, A Comparative review on image interpolation-based reversible data hiding, *Lecture Notes in Networks and Systems*, **664 LNNS** (2023) 687–699, doi: 10.1007/978-981-99-1479-1_51.
13 Krishna D & Jena B, A trained CNN based resolution enhancement of digital images, *Int J Modern Trends Sci Technol*, **05(7)** (2019) 9–16.
14 Mohan A, Dwivedi R & Kumar B, Image restoration of landslide photographs using SRCNN, *Lecture Notes in Elec Engg*, **777** (2022) 1249–1259, doi: 10.1007/978-981-16-2761-3_108.
15 Lin K, In *The Performance of Single-Image Super-Resolution Algorithm:* ICISCAE 2022/09//, IEEE: 2022, 964–968.
16 Lee K M, Lee P J & Bui T A, Edge detection cascaded with simplified WDSR of IR super resolution, *ICSSE 2021*, (2021) 378–381, doi: 10.1109/icsse52999.2021.9538439.
17 Zhao H, Kong X, He J, Qiao Y & Dong C, Efficient image super-resolution using pixel attention, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, **12537** (2020) 56–72, doi: 10.1007/978-3-030-67070-2_3.
18 Li K, Yang S, Dong R, Wang X & Huang J, Survey of single image super-resolution reconstruction, *IET Image Proc*, **14(11)** (2020) 2273–2290, doi: 10.1049/iet-ipr.2019.1438.
19 Zhang J, Liu M, Wang X & Cao C, Residual net use on fsrcnn for image super-resolution, *Chinese Control Conf*, **2021** (2021) 8077–8082, doi: 10.23919/ccc 52363. 2021.9550205.
20 Zhang M & Ling Q, Supervised pixel-wise GAN for face super-resolution, *IEEE Trans on Multimedia*, **23** (2021) 1938–1950, doi: 10.1109/tmm.2020.3006414.
21 Bashir S M A, Wang Y, Khan M & Niu Y, A comprehensive review of deep learningbased single image super-resolution, *PeerJ Comp Sci*, **7**(2021) 1–56, doi: 10.7717/peerj-cs.621.
22 Chang K, Li M, Ding P L K & Li B, Accurate single image super-resolution using multi-path wide-activated residual network, *Signal Proc*, **172** (2020), doi: 10.1016/ j.sigpro.2020.107567.
23 Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A & Bengio Y, Generative adversarial networks, *Adv In Neu Info Proc Sys*, **27** (2014) 139–144, doi: 10.1145/3422622.
24 Yuan X, Shi J & Gu L, A review of deep learning methods for semantic segmentation of remote sensing imagery, *Expert Syst Appl*, **169** (2021), doi: 10.1016/j.eswa.2020.114417.
25 Mirza M & Osindero S, *Conditional generative adversarial nets*, 2014, http://arxivorg/abs/14111784.
26 Ronneberger O Fischer P & Brox T, U-net: Convolutional networks for biomedical image segmentation, *Lec Notes in Comp Sci*, **9351** (2015) 234–241, doi: 10.1007/978-3-319-24574-4_28.
27 Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A & Bengio Y, Generative adversarial networks, *Comm ACM*, **63(11)** (2020) 139–144, doi: 10.1145/3422622.
28 Li C & Wand M, Precomputed real-time texture synthesis with markovian generative adversarial networks, *Lec Notes in Comp Sci*, **9907 LNCS** (2016) 702–716, doi: 10.1007/978-3-319-46487-9_43.
29 Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z & Shi W, Photo-realistic single image super-resolution using a generative adversarial network, *CVPR 2017*, **2017** (2017) 105–114, doi: 10.1109/cvpr.2017.19.
30 Wang X, Yu K, Wu S, Gu J, Liu Y, Dong C, Loy C C, Qiao Y & Tang X, *ESRGAN: Enhanced super-resolution generative adversarial networks*, 2018, http://arxivorg/ abs/180900219.
31 Jiang K, Wang Z, Yi P, Wang G, Lu T & Jiang J, Edge-enhanced gan for remote sensing image superresolution, *IEEE Transactions on Geoscience and Remote Sensing*, **57(8)** (2019) 5799–5812, doi: 10.1109/tgrs.2019.2902431.
32 Wang Z, Jiang K, Yi P, Han Z & He Z, Ultra-dense GAN for satellite imagery super-resolution, *Neurocomputing*, **398** (2020) 328–337, doi: 10.1016/j.neucom.2019.03.106.
33 Tu J, Mei G, Ma Z & Piccialli F, SWCGAN: Generative adversarial network combining swin transformer and CNN for remote sensing image super-resolution, *IEEE J Sel Top Appl Earth Obs Remote Sens*, **15** (2022) 5662–5673, doi: 10.1109/jstars.2022.3190322.
34 Jia S, Wang Z, Li Q, Jia X & Xu M, Multiattention generative adversarial network for remote sensing image super-resolution, *IEEE Transactions on Geoscience and Remote Sensing*, **60** (2022), doi: 10.1109/tgrs.2022.3180068.
35 Dong C, Loy C C, He K & Tang X, Image super-resolution using deep convolutional networks, *IEEE transactions on pattern analysis and machine intelligence*, **38(2)** (2015) 295–307, doi: 10.1109/tpami.2015.2439281.
36 Wang X, Xie L, Dong C & Shan Y, Real-esrgan: Training real-world blind super-resolution with pure synthetic data, in *Proc IEEE/CVF Int Conf Comput Vis* (IEEE) 2021, 1905–1914.
37 Cheon M, Yoon S J, Kang B & Lee J, Perceptual image quality assessment with transformers, *IEEE Comp Soc Conf Comp Vis Pattern Recog Workshops* (IEEE) 2021, 433–442, doi: 10.1109/cvprw53098.2021.00054.
38 De Tommaso S F N, Rodrigues I & Pinsky V C, Interactive qualitative analysis: A methodological contribution to researches in sustainability, *Administração: Ensino e Pesquisa*, **22(2)** (2021) 271–293.
39 Wang Y, Bashir S M A, Khan M, Ullah Q, Wang R, Song Y, Guo Z & Niu Y, Remote sensing image super-resolution and object detection: Benchmark and state of the art, *Expert Syst Appl*, **197** (2022), doi: 10.1016/j.eswa.2022.116793.