Enhancing Spatial Resolution of Sentinel-2 Imagery through Deep Learning and Generative Adversarial Networks: GS-SRGAN

Saad Farah [1], Hachem Saadaoui [2], Hatim Lechgar [1], Mehdi Maanan [2], Mohamed Maanan [3],
Anass Rghioui [1], Hassan Rhinane [2]

[1] LaGeS-SGEO Laboratory, Hassania School of Public Works, Casablanca, Morocco
[2] Geoscience Laboratory, Hassan II University, Casablanca, Morocco
[3] UMR 6554 CNRS LETG-Nantes Laboratory, Institute of Geography and Planning, Nantes University, 44312 Nantes, France

Keywords: Super-resolution, Generative Adversarial Network, Deep Learning, Sentinel-2, Google Earth Imagery.

Abstract

Sentinel-2 satellites provide multi-spectral images with 13 bands at resolutions of 10, 20, and 60 m/pixel, widely used for various applications due to their cost-free access and high revisit frequency. Their open data policy has made them a key resource in remote sensing. Nonetheless, the growing need for high-resolution images has highlighted the significance of super-resolution technology (SR), which improves spatial detail through enhanced sensor precision and density. Deep learning techniques are an effective solution for enhancing Sentinel-2 images through super-resolution, improving low-resolution images by retrieving fine-grained high-frequency details. This results in high-resolution outputs from freely available data. In this research, we propose an enhancement of single-image resolution model derived from a Generative Adversarial Network, commonly abbreviated as GAN. We implemented and trained a model, named **GS-SRGAN** (Google Sentinel - SRGAN), built on the foundation of the Super-Resolution GAN model (SRGAN), using pairs of Google Earth and Sentinel-2 images for generating super-resolved outputs of the RGB bands from the multispectral Sentinel-2 data using a 4x scaling factor. The results from our **GS-SRGAN** model surpass those of current best in class models when evaluated using standard metrics such as SSIM (Structural Similarity Index) and PSNR (Peak Signal-to-Noise Ratio), enabling the super-resolved Sentinel-2 imagery for use in studies that demand very high spatial resolution.

1. Introduction

Renewable energies, particularly photovoltaic (PV) solar energy, play a vital part in the worldwide shift away from fossil fuels, contributing significantly to mitigating climate change impacts. Solar photovoltaic systems are regarded as a potential of renewable energy sources. Due to their minimal footprint, ease of production, and simple maintenance (Selvaraj et al., 2024) unlike wind energy, which necessitates specific geographic conditions, PV solar panels can be deployed in diverse settings, offering a flexible solution for sustainable energy production (Popovych et al., 2023).

Solar PV power systems are seen as an important solution for the energy crisis and environmental problems (Goyal et al., 2024). There are various methods to capture solar energy, including solar heating, solar photovoltaics (PV), and solar hybrid photovoltaic thermal (PVT) systems. Because it can transform sunlight directly into electrical power. Owing to its ability to transform sunlight directly into electricity, flexibility, low cost, and extended lifespan, solar PV has developed significantly in recent decades (You, 2022).

However, a significant challenge is the extensive surface area required for effective energy generation. This issue can be addressed by harnessing the abundant unused roofs spaces available globally, making large-scale solar installations more feasible (Saripalli et al., 2024). As the expense of photovoltaic systems keeps declining and their efficiency enhances, solar energy's role in sustainable development is expected to grow (D. Li, 2023).

(Galagan & Frigerio, 2023) to efficiently identify and analyze suitable rooftops for solar panel deployment, satellite imagery, specifically Sentinel-2, can be employed. Sentinel-2 is an invaluable tool for this purpose due to its free and open-source nature, making it accessible for widespread use. Improving Sentinel-2 imagery from a spatial resolution of 10 meters per pixel to a finer resolution of 2.5 meters per pixel allows for more detailed and precise spatial analysis. This enhanced resolution is crucial for accurately mapping and assessing rooftop spaces. However, the upscaling process also introduces certain challenges, such as increased data processing requirements and potential inaccuracies in the analysis (Galagan & Frigerio, 2023).

The utilization of Very High Spatial Resolution (VHSR) images can be prohibitively expensive, especially for extensive area coverage or conducting multi-temporal analyses (L. Salgueiro Romero et al., 2020). Therefore, in such cases, it is advantageous to consider freely available data with adequate spatial quality, such as those provided by satellites like Landsat or Sentinel. Specifically, the Copernicus Sentinel-2 mission consists of two polar-orbiting satellites. Its multispectral instrument (MSI) captures data in 13 spectral bands, from visible to short-wave infrared (L. Salgueiro Romero et al., 2020). More specifically, this sensor captures images with a spatial resolution of 10 meters for the red, green, blue, and near-infrared bands (L. Salgueiro Romero et al., 2020).

(Xin Li & Orchard, 2001) Single-image spatial super-resolution (SR), which seeks to reconstruct a high spatial resolution image from a low spatial resolution one, can overcome the inherent spatial resolution limitations of imaging systems without relying on any additional prior knowledge or auxiliary information. The fundamental approach to single-image spatial super-resolution involves using nonlinear interpolators like bilinear and bicubic interpolation (Xin Li & Orchard, 2001), which directly utilizes the information from neighboring pixels. However, these methods often lead to edge blur and ringing effect (Baraskar et al., 2023a). To address this challenge, this study integrates Geographic Information System (GIS), remote sensing (RS), computer vision (CV) and Deep learning (DL) to establish a model for predicting a super resolution image. Hence, Super-resolution of satellite imagery is important for maximizing the utility of the available images (Baraskar et al., 2023b).

Over the past decade, various supervised super-resolution (SR) techniques have emerged, particularly those utilizing dictionary approaches. Notable contributions include Yang and al (L. Salgueiro Romero et al., 2020) and (Weisheng Dong et al., 2011), and (Gou et al., 2014), who effectively employed sparse coding methods. also, Pan et al. incorporated structural self-similarity combined with compressive sensing for SR tasks, while other researchers (Y. Zhang et al., 2014); (J. Li et al., 2016) explored diverse image feature spaces to enhance performance. Generally, these methods focus on the low-level features of images.

Recently, deep learning has significantly advanced the SR field (Z. Wang et al., 2019). High-level features from optical data

have notably improved performance through convolutional neural networks (CNNs) for image SR. Dong et al. introduced SRCNN (Dong et al., 2014), a pioneering three-layer CNN, which spurred the development of numerous other CNN approaches (Dong et al., 2014).

Generative Adversarial Networks (GANs) have recently attracted significant attention. Ledig et al. proposed SRGAN, which produces more photo-realistic outputs despite lower quantitative metrics, representing a seminal perceptual approach in super-resolution. SRGAN's generator comprises 16 residual blocks and 2 up-sampling layers to achieve a scaling factor of 4. The training process is divided into phases: initially training the generator and subsequently combining adversarial and perceptual losses (Johnson et al., n.d.); (X. Wang et al., n.d.) The enhancement of SRGAN through denser residual layers and the elimination of batch normalization layers to reduce artifacts, along with the use of a relativistic GAN, represents a major advancement in image super-resolution techniques. Significant work in this field includes "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks", which introduces the Residual-in-Residual Dense Block (RRDB) without batch normalization and uses a relativistic discriminator to boost visual quality. Moreover, the study "A Super-Resolution Generative Adversarial Network with Simplified Gradient Penalty and Relativistic Discriminator" proposes a streamlined residual network without batch normalization layers and improves the GAN using a relativistic discriminator for sharper image reconstruction (X. Wang et al., 2019). The "RBDN: ReRRDBsidual Bottleneck Dense Network for Image Super-Resolution" further optimizes SRGAN by employing dense cascading connections and a residual-in-residual bottleneck block, eliminating batch normalization to reduce negative effects and improve convergence (X. Wang et al., 2019). These advancements collectively enhance the super-resolution capabilities by improving feature utilization, reducing computational costs, and stabilizing training for higher quality image outputs.

CNN-based SR models encounter challenges in remote sensing due to insufficient training datasets. (Ma et al., 2018) introduced TGAN (Transferred Generative Adversarial Network) to address this issue. (Haut, Paoletti, et al., 2019); (Haut, Fernandez-Beltran, et al., 2019) and (Lei et al., 2017) synthesized LR-HR pairs from public remote sensing images and tested various network architectures. However, models trained solely on synthetic LR-HR pairs often fail to generalize due to the complex nature of LR images, which result from more than just Gaussian noise, blurring, or compression artifacts (Bulat et al., 2018). More recently, Ma et al. proposed DRGAN (Dense Residual Generative Adversarial Network) for remote sensing SR. Pouliot et al. utilized Landsat (30 m GSD) and Sentinel-2 (10m GSD) imagery to train CNN architectures for SR of Landsat imagery. Beaulieu et al. experimented with various CNNs and GANs using Sentinel-2 (10m GSD) and WorldView (2.5 m GSD) images, achieving a scaling factor of 4. These models, trained with red, green, and NIR channels, showed promising results, particularly with GAN-based networks. They speculated that geometric and radiometric disparities between satellite images contributed to lower metric values. Several researchers have also applied ESRGAN architectures to various remote sensing SR applications (Chen et al., 2019); (L. F. Salgueiro Romero et al., 2020).

For the purpose of expanding the work of super-resolution to be employed on geographical satellite imagery, the proposed work has adopted the approach of generative models to create visually appealing images, generative models, also known as GANs, work to enhance perceptual quality. SRGAN, for instance, uses a GAN-based architecture to produce visually

appealing images. It employs the SRResnet network architecture as the backend and leverages a multi-task loss function to improve the results. Three terms make up the loss: MSE loss capturing pixel similarity, Perceptual similarity loss used to capture high-level information by using a deep network. Adversarial loss from the discriminator. The current study employs the Real ESRGAN technique, leveraging the advantages of a standard SRGAN, and presents the algorithm's results through a user-friendly GUI developed with Flask on the backend (Baraskar et al., 2023b).

2. Satellite Images

Satellite images are digital images of the Earth's surface captured by sensors aboard satellites circling the Earth. These images provide valuable data across various spectral bands, such as visible, infrared, and thermal wavelengths (Roy et al., 2014a). Satellite imagery is essential for monitoring Earth's environmental changes, (Drusch et al., 2012a); (Wulder et al., 2019). The Landsat 8 and 9 satellites offer images with a spatial resolution of 30 meters, making them ideal for monitoring environmental changes, land use, and natural disasters (Roy et al., 2014b). Sentinel-2, part of the European Space Agency's Copernicus program, captures high-resolution optical imagery with a spatial resolution of up to 10 meters (Fernández-Manso et al., 2012). Super-resolution applications in Sentinel-2 imagery further enhance the detection of small-scale features, supporting precision agriculture and urban planning. Google Earth, known for its very high-resolution imagery (VHR) from various commercial sources, offers spatial resolutions as fine as 30 centimeters (Vuolo et al., 2015). This makes it particularly useful for detailed urban planning and infrastructure monitoring. In our case, we will apply super-resolution techniques to Sentinel-2 imagery. Together, if these satellite platforms are enhanced by super-resolution, they will represent powerful tools for roofs observation and analysis.

2.1. Sentinel-2

Sentinel-2 is part of the Copernicus program, an initiative by the European Space Agency (ESA), the European Environment Agency (EEA), and the European Commission (EC) aimed at delivering operational data about our planet for environmental and safety applications. The mission includes two twin satellites, named Sentinel-2A and Sentinel-2B, launched in June 2015 and March 2017, respectively. These satellites share the same orbit but are spaced 180° apart, allowing ensuring a frequent revisit cycle of 5 days at the Equator. Each satellite is equipped with a multispectral instrument (MSI) capable of capturing 13 spectral bands spanning from visible and near infrared (VISNIR) to shortwave infrared (SWIR). Among these bands, four (B2, B3, B4, and B8) have a spatial resolution of 10 meters, six bands have a 20-meter resolution, and three bands have a 60-meter resolution. The lower resolution bands are primarily used for assessing vegetation health, differentiating snow, ice, and clouds, and gathering data on aerosols, water vapor, and cirrus clouds (Drusch et al., 2012b; Gascon et al., 2017a).

To enhance the spatial resolution of the Sentinel-2 10-meter bands through super-resolution techniques, data from the multispectral Worldview satellites was utilized to create the LR-HR training dataset. WorldView-2, launched in October 2009, was the first commercial satellite to offer high-resolution 8-band multispectral imaging, operating at an altitude of 770 km and providing a 1.84-meter resolution for VISNIR bands and a 46 cm panchromatic band. WorldView-3, launched in August 2014, operates at approximately 617 km altitude and delivers 1.24-meter resolution for VISNIR bands, along with additional SWIR bands at 3.7 meters and a 31 cm panchromatic band

(Drusch et al., 2012b); (Gascon et al., 2017b). Table 1 below presents the technical specifications of Sentinel-2's spectralbands in the visible and the near-infrared (VIS-NIR) range.

	Wavebands	Spectral Band Edges	Spatial Resolution
B1	Coastal Aerosol	443 nm	60m
B2	Blue band	490 nm	10m
B3	Green band	560 nm	10m
B4	Red band	665 nm	10m
B5	Red-edge 1	705 nm	20m
B6	Red-edge 2	740 nm	20m
B7	Red-edge 3	783 nm	20m
B8	Near-infrared	842 nm	10m
B8A	Near-infrared narrow	865 nm	20m
B9	Water Vapor	945 nm	60m
B10	SWIR-Cirrus	1375 nm	60m
B11	SWIR-1	1610 nm	20m
B12	SWIR-2	2190 nm	20m

Table 1. Technical Specifications of Sentinel-2's Spectral Bands in the Visible and the Near-Infrared Spectrum.

2.2. High-Resolution Imagery Sources

In this study, the high-resolution imagery utilized is derived from the Google basemap and Bing basemap in QGIS software, also known collectively as Google Earth and Bing Maps Imagery. This type of imagery is classified as Very High Resolution (VHR) satellite imagery, which typically offers a ground sample distance (GSD) ranging from less than 1 meter to 5 meters. VHR imagery is essential for detailed analysis and is commonly used in applications such as mapping, urban planning, and environmental monitoring (Gómez et al., 2016); (J. Zhang et al., 2020). Key providers of VHR satellite imagery include Maxar Technologies (formerly DigitalGlobe), which operates satellites like WorldView-3, WorldView-4, GeoEye-1, and QuickBird; Airbus Defence and Space with Pleiades-1A, Pleiades-1B, SPOT 6, and SPOT 7; and Planet Labs' SkySat constellation.

The high-resolution images from the Google basemap and Bing basemap are composites sourced from multiple satellite and aerial imagery providers. These composites often include data from DigitalGlobe/Maxar's satellite constellation (WorldView, GeoEye, QuickBird), Landsat (with enhanced products), Sentinel-2 (offering 10m GSD optical imagery), and Airbus (through Pleiades and SPOT satellites), alongside various high-resolution aerial photographs (Hu et al., 2019); (Wulder et al., 2016). Such integrated imagery ensures extensive and detailed coverage, which is vital for the accuracy and reliability of our super-resolution analysis. The use of these high-quality composites is key for enhancing the spatial resolution of

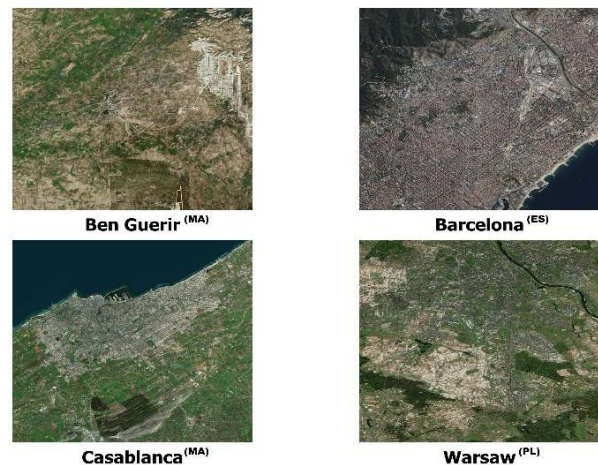satellite data and supports the objectives of this project (Gómez et al., 2016); (J. Zhang et al., 2020).

Figure 1. High-Resolution Satellite Imagery from Google and Bing Basemaps for Various Cities.

3. Data preparation

The preparation of traditional RGB imagery differs significantly from that of satellite images, especially when using deep learning algorithms like SRGAN. In models such as CNNs with fixed input sizes, for instance, 512×512×3, RGB images can be simply resized to this dimension. Due to the small range of pixel values (0 to 255), the average values do not change drastically during resizing.

On the other hand, satellite images require a more intricate approach. Firstly, they need to be georeferenced, ensuring accurate spatial representation. Secondly, resizing satellite images can result in significant changes to pixel values because of their larger bit depth. For example, pixel values in Sentinel-2A images are 16-bit compared to the 8-bit depth in traditional RGB images. This greater range means that averaging pixel values during resizing can lead to substantial changes, potentially affecting the output inference.

In our experiment, we set up two different test cases, as illustrated in the accompanying graphs, to address these challenges in preparing satellite imagery for the SRGAN model. These test cases involved creating high-resolution and low-resolution image pairs from Sentinel-2 and Google Earth data, accounting for the unique properties of satellite images to ensure accurate model training and evaluation.
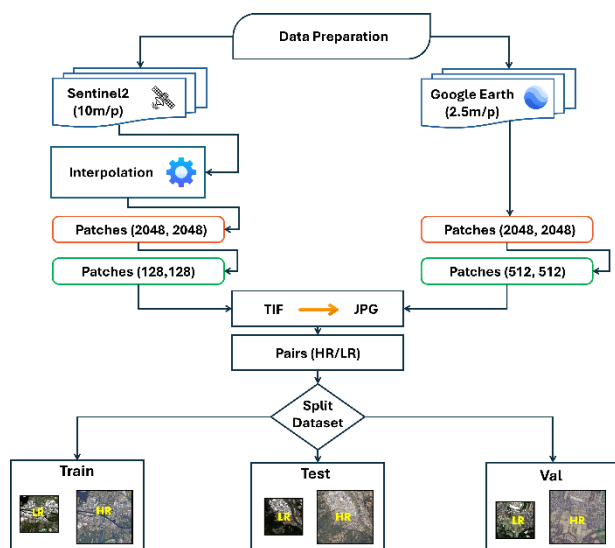
Figure 4. High-Resolution and Low-Resolution Satellite Image Pairs.

These images illustrate the final output of the data preparation process, showcasing the visual differences between low-resolution and high-resolution images used in the study. On the left portion of each pair, the larger image provides a broader geographic context, while the red box highlights the specific area of interest. The right portion of each pair presents a zoomed-in comparison between the low-resolution and high-resolution versions of the same area. The low-resolution images display noticeable blurring and loss of detail, which underscores the necessity for super-resolution techniques. In contrast, the high-resolution images reveal finer details and sharper features, demonstrating the value of high-resolution imagery in detailed analysis and applications. These prepared image pairs serve as the foundation for training and evaluating the GS-SRGAN model.

4. Methodology

To enhance the resolution of satellite imagery, we will utilize the Super-Resolution Generative Adversarial Network (SRGAN) model, a state-of-the-art approach in the field of image super-resolution. GS-SRGAN employs a deep learning framework that generates high-resolution images from low-resolution inputs by training a generator network to produce realistic textures and details. In general, the SRGAN model is composed of two primary components: a generator that creates the super-resolved images and a discriminator that distinguishes between generated and real images, thus driving the generator to improve its outputs continually (Huang et al., 2017a). This adversarial process results in images that not only have higher pixel density but also capture fine details and complex structures, making them visually convincing. The application of SRGAN to satellite images, such as those from Sentinel-2 and Landsat, allows for significant improvements in spatial resolution, which enhances the ability to analyze and interpret subtle features in environmental monitoring, urban planning, and resource management (Bauer-Marschallinger et al., 2019). Moreover, the SRGAN model's ability to reconstruct fine details is particularly beneficial for high-resolution platforms such as Google Earth, where precision is paramount for applications such as infrastructure assessment and disaster management (Mushkin et al., 2020). Overall, GS-SRGAN's innovative architecture and capability to produce realistic super-resolved images make it an invaluable tool in the ongoing efforts to maximize the utility of satellite imagery across various domains.

Figure 2. Super-Resolution Image Processing and Dataset Creation. The graph illustrates in Figure 2 a data preparation workflow for generating high-resolution (HR) and low-resolution (LR) image pairs used in super-resolution modeling. The process begins with acquiring satellite imagery from Sentinel-2 and Google Earth. Sentinel-2 images, with a 10-meter ground sample distance (GSD), undergo interpolation to enhance their resolution, creating 2048×2048-pixel patches that offer greater detail. Interpolation methods such as bicubic or spline interpolation are used to estimate unknown pixel values, maintaining spatial properties while increasing resolution. High-resolution patches from Google Earth (2.5m GSD) are extracted directly. These patches are resized into smaller LR versions, such as 128×128 for Sentinel-2 and 512×512 for Google Imagery, allowing the model to learn the upsampling process. The image patches are converted from TIFF to JPG and paired to form HR/LR datasets, which are then split into 70% of training, 20% of testing, and 10% of validation sets. This comprehensive preparation ensures the model learns effectively from detailed and accurately represented data, enhancing its capability to perform super-resolution tasks for applications in urban planning and environmental monitoring.
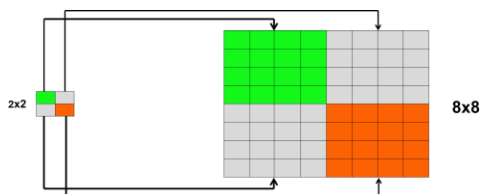

Figure 3. Visualizing 2x2 to 8x8 Pixel Upsampling.

Figure 3 illustrates the pixel relationship between a 2×2 tile and its corresponding 4x up-sampled 8×8 tile. Directly resizing images to dimensions like 32×32 or 128×128 can disrupt their spatial properties, so it's not an ideal approach. Instead, the optimal method is to crop sections from the original tile using dimensions that are multiples of 32 and 128. This approach preserves the pixel relationships, allowing the model to effectively learn the 4x up-sampling of each pixel. By employing data augmentation methods such as rotating and flipping the images, we were able to generate thousands of samples for our model.

Figure 5. Google-Sentinel Super-Resolution Generative
Adversarial Network (GS-SRGAN) architecture.

As depicted in Figure 5, the use of our Google-Sentinel Super-Resolution Generative Adversarial Network (GS-SRGAN) for enhancing and analysing satellite images. The process begins with low-resolution satellite images that are fed into the GS-SRGAN, where the generator component enhances the spatial resolution, producing high-resolution outputs with finer details and improved clarity. These enhanced images are then processed through a classification and feature extraction pipeline. The classifier categorizes the high-resolution images into predefined classes, useful for applications such as land use classification and urban planning, while the feature extractor identifies meaningful patterns and details for further analysis, such as object detection and change detection. This workflow demonstrates the capability of GS-SRGAN to significantly improve the quality of satellite imagery, enabling more accurate and detailed analysis for various practical applications.

4.1. Generator

The generator in the Google Sentinel Super-Resolution Generative Adversarial Network (GS-SRGAN) plays a crucial role in transforming low-resolution images into high-resolution outputs. It uses a deep neural network architecture that consists of multiple convolutional layers, which work together to learn and reconstruct high-frequency details lost in lower-resolution images. The generator begins with a convolutional layer followed by a Parametric ReLU (PReLU) activation function, which introduces non-linearity and allows for more complex mappings from input to output (Huang et al., 2017b). Subsequent layers are organized into residual blocks, which are instrumental in improving image quality by preserving essential information and refining image features progresses through a sequence of skip connections. These connections mitigate issues like vanishing gradients, which are common in deep networks, by allowing the gradient to bypass some layers and reach earlier layers directly (He et al., 2016). Batch normalization layers are also used to stabilize learning and enhance convergence speed, ensuring that the output maintains high visual fidelity. Finally, the generator incorporates deconvolution layers, also referred to as transposed convolutional layers, which upscale the image to the desired high resolution by effectively learning how to fill in missing details and textures (Lu et al., 2018). Overall, the GS-SRGAN generator is designed to produce visually appealing images that are close to natural photographs, making it highly effective for applications that require high-quality image reconstruction from low-resolution inputs.
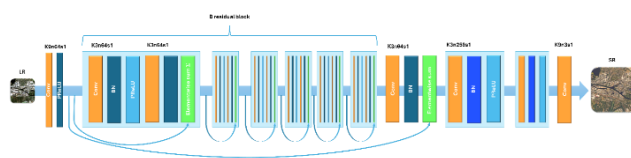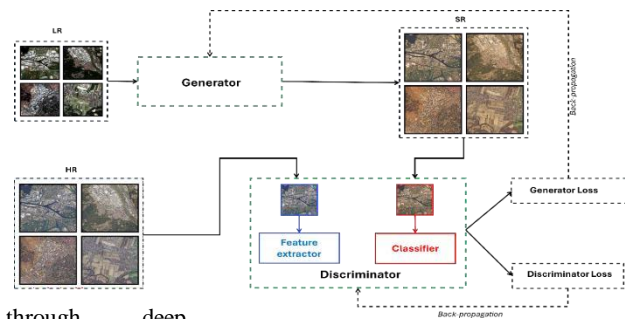
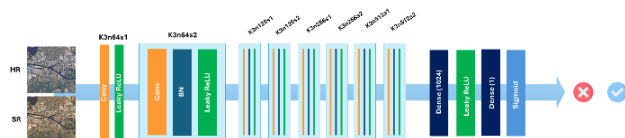Figure 6. GS-SRGAN Generator Architecture for Super-Resolution.

This schematic illustrates the architecture of the generator within the Super-Resolution Generative Adversarial Network (GS-SRGAN) as presented in Figure 6. It visually depicts the data flow through the generator, starting from a low-resolution input image and culminating in a high-resolution output. Key components of the architecture are highlighted, including the initial convolutional layer, which processes the input image, and the series of residual blocks, which enhance the image quality

through deep learning techniques. Each residual block is depicted with its convolutional layers and activation functions, illustrating how they contribute to refining the image details. The architecture also shows the use of batch normalization to maintain stable learning conditions and the final transposed convolutional layers responsible for upscaling the image. This comprehensive structure allows the GS-SRGAN generator to excel in producing detailed, high-resolution images from lower-quality inputs.

4.2. Discriminator

The discriminator in the GS Super-Resolution Generative Adversarial Network (GS-SRGAN) serves a critical function in distinguishing between real high-resolution images and those generated by the network, thereby guiding the generator to improve its output. The discriminator uses a deep convolutional neural network (CNN) architecture, consisting of multiple convolutional layers followed by leaky ReLU (Rectified Linear Unit) activations, which introduce non-linearity and allow the network to capture complex patterns and textures in the image data. This architecture effectively extracts hierarchical features at varying scales, making it adept at detecting discrepancies between generated images and real ones (Huang et al., 2017c). As the generator produces images, the discriminator evaluates them alongside actual high-resolution images, learning to classify them as either real or fake. This adversarial process creates a feedback loop in which the generator iteratively improves its image quality to fool the discriminator

(Goodfellow et al., 2014). To stabilize training and enhance performance, batch normalization layers are incorporated within the discriminator to normalize the feature maps and promote faster convergence (Radford et al., 2015). The discriminator's ability to accurately differentiate between high-quality and low-quality images is indispensable for refining the super-resolution output, ultimately leading to visually convincing results that closely resemble real photographs. This mechanism enhances the realism and fidelity of images produced by GS-SRGAN, making it a powerful tool in applications requiring high-resolution image reconstruction from low-resolution inputs.

Figure 7. GS-SRGAN Discriminator Architecture with Real and Generated Image Pair Evaluation.

The Figure 7 showcased the architecture of the discriminator component within the Google-Sentinel Super-Resolution Generative Adversarial Network (GS-SRGAN). The discriminator is designed to distinguish between real high-resolution images and the ones generated by the GS-SRGAN model. At the input, the discriminator receives pairs of images, one real and one generated, which are fed through a series of

convolutional layers. These layers, highlighted as the feature extractor, play an essential role in a role in analyzing and evaluating the authenticity of the images by capturing intricate details and patterns.

The feature extractor is integral to the discriminator's ability to differentiate image quality, and it functions similarly to well-known CNN architectures like VGG16 and VGG19. Both VGG16 and VGG19 are deep convolutional neural networks used for image classification, and they serve as powerful feature extractors by processing images through multiple layers of convolutions, pooling, and activation functions to extract hierarchical features. Within the scope of GS-SRGAN, such feature extractors are often pre-trained on extensive datasets such as ImageNet to ensure they can effectively recognize complex textures and structures in the image (Simonyan & Zisserman, 2014a).

In this architecture, each convolutional layer is succeeded by a Leaky ReLU activation function, which introduces non-linearity and helps model complex patterns. Batch normalization (BN) layers are used to stabilize the learning process by normalizing the feature maps, which improves convergence speed. After the feature extraction layers, the network includes dense (fully connected) layers, culminating in a sigmoid activation that outputs a binary classification. This final step indicates whether an image is classified as real (check mark) or fake (cross).

4.2.1. VGG16 Feature Extractor:

VGG16 is a well-known deep convolutional neural network architecture created by the Visual Geometry Group at the University of Oxford. It is recognized for its straightforward yet powerful design, comprising 16 layers, including 13 convolutional layers with 3×3 filters and 3 fully connected layers. This uniform architecture allows VGG16 to effectively capture hierarchical features, ranging from edges to complex textures, making it highly effective for image classification tasks. The network utilizes ReLU (Rectified Linear Unit) activation functions after each convolutional layer to introduce non-linearity and mitigate the vanishing gradient problem (Simonyan & Zisserman, 2014b).
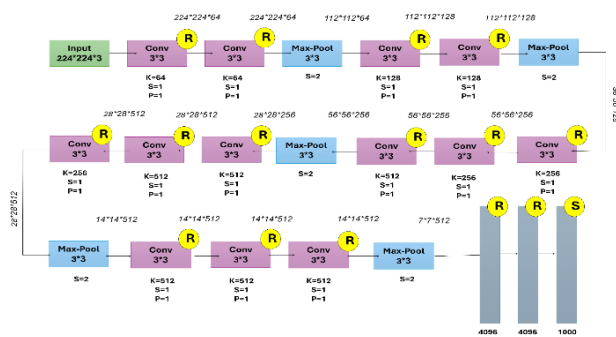

Figure 8. VGG16 Convolutional Neural Network Architecture Diagram.

The diagram portrays the flow of data through VGG16, starting with an input image of size 224×224×3, followed by convolutional and max-pooling layers, which reduce the size of the feature maps while retaining essential information. The network concludes with fully connected layers and a softmax layer, which outputs class probabilities for image classification. VGG16's straightforward architecture, along with its strong performance in benchmark competitions like ImageNet, makes it a popular choice for transfer learning, serving as a robust feature extractor for various tasks beyond image classification,

including object detection and segmentation. The softmax layer, indicated by S in the diagram, transforms network outputs into a probability distribution across classes, enabling classification by choosing the class with the highest probability. Overall, VGG16's depth and uniform design allow it to efficiently learn and represent complex visual features, solidifying its role as a foundational model in deep learning research and applications.

4.2.2. VGG19 Feature Extractor:

VGG19 is an enhanced version of the VGG16 architecture, also created by the Visual Geometry Group from the University of Oxford. Known for its simple and consistent design, VGG19 consists of 19 layers with learnable parameters, including 16 convolutional layers and 3 fully connected layers, while VGG16 includes just 13 convolutional layers and 3 fully connected layers. The additional convolutional layers in VGG19, particularly in the last three blocks, enable it to capture finer details and more intricate patterns, offering enhanced feature learning capabilities. Both architectures utilize small 3×3 filters in convolutional layers and 2×2 max-pooling layers for down-sampling, preserving spatial hierarchies and allowing deep feature extraction across scales (Simonyan & Zisserman, 2014c).
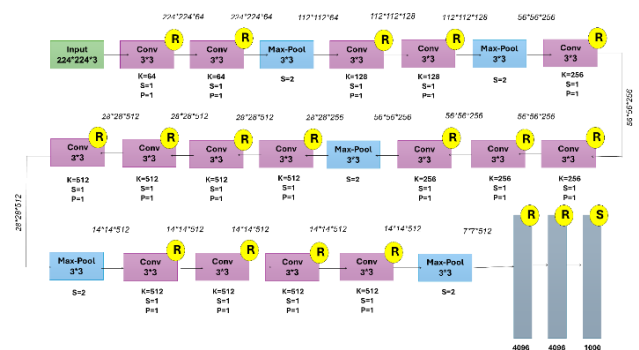

Figure 9. VGG19 Convolutional Neural Network Architecture Diagram.

The VGG19 diagram as revealed in Figure 9 highlights these convolutional layers, with ReLU activation functions, denoted by **R**, adding non-linearity to facilitate complex feature learning. The architecture concludes with fully connected layers and a softmax activation, indicated by **S**, to output class probabilities. While VGG19's increased depth allows for improved performance in tasks requiring detailed feature representation, it demands more computational resources and longer training times compared to VGG16. Therefore, the choice between VGG16 and VGG19 typically depends on the specific task requirements, balancing computational efficiency against the need for deeper feature extraction. Both models remain effective for image classification and transfer learning, with pre-trained versions available on large datasets such as ImageNet, making them valuable tools for deep learning applications. In our model, we will use VGG19.

We chose VGG19 over VGG16 due to its deeper architecture, which includes three additional convolutional layers. These extra layers allow VGG19 to learn more complex features and representations, crucial for tasks like perceptual loss in SRGAN. The richer feature maps provided by VGG19 improve the

quality of generated images, capturing subtle details and textures, resulting in more realistic and visually appealing super-resolved images.

5. Quantitative Metrics

To evaluate the accuracy of the generator's results, denoted as G(I), using the input low-resolution (LR) image I from the test set, we compare them to the corresponding high-resolution (HR) target image J. The quantitative assessment of the super-resolved (SR) bands involves calculating several parameters, including the Root Mean Square Error (RMSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM).

- The Root Mean Square Error (RMSE) quantifies the average deviation between the predicted pixel values and the actual pixel values in an image, offering a measure of the average error in the pixel-value space.

$$RMSE(I,J) = \sqrt{E[(I-J)^2]} \qquad (1)$$

Where I = *the predicted pixel values,*
   J = *the actual pixel values,*

   E = *the expected value*

- Peak Signal to Noise Ratio (**PSNR**): is a standard metric used to assess the quality of a reconstructed image. In this context, MaxVal represents the maximum value of the high-resolution (HR) image (J).

$$PSNR(G(I),J) = 20 \cdot \log_{10}\left(\frac{MaxVal}{RMSE[G(I),J]}\right) \qquad (2)$$

- The Structural Similarity Index Measure (**SSIM**) is a metric that evaluates the similarity between two images by considering three essential components: *luminance*, *contrast*, and *structure*. The SSIM value ranges from -1 to 1, where a value of 1 indicates that the images are *identical*. The metric includes two constants, $C_1 = k_1 L$ and $C_2 = k_2 L$, which are depend on the dynamic range L of the pixel values, with default values $k_1 = 0.01$ and $k_2 = 0.03$ (Z. Wang et al., 2004).

$$SSIM(G(I),J) = \frac{(2\mu_I \times \mu_J + C_1)(2\sigma_{IJ} + C_2)}{(\mu_I^2 + \mu_J^2 + C_1)(\sigma_I^2 \times \sigma_J^2 + C_2)} \qquad (3)$$

Where $_I$, $_J$= *the mean intensities of images* I *and* J,

   σ_I^2 , σ_J^2= *the variances,*

   σ_{IJ} = *the covariance between the two images.*

6. Loss functions in GS-SRGAN

In this section, we delve into the various loss functions used in our GS-SRGAN (GS - Super-Resolution Generative Adversarial Network) training process. Each loss function plays a crucial role in ensuring the generator and discriminator networks learn effectively to produce high-quality super-resolution images. The following are detailed descriptions and purposes of each loss function:

- **MSE** stands for Mean Squared Error. It is a commonly used loss function that measures the average of the squares of the differences between predicted and actual values. In the context of neural networks and machine learning, MSE is used to quantify the error between the predicted output and the true target.

$$MSE = \frac{1}{n} \sum_{i=1}^{n}(I_i - J_i)^2 \qquad (4)$$

Where n = *number of data points,*
   I_i = *the predicted value for the i-th data point,*

   J_i = *the actual value for the i-th data point.*

- The adversarial loss ($Loss_{GAN}$) is crucial in training the generator in GS-SRGAN framework. It measures how well the generator produces high-resolution images that the discriminator mistakenly classifies as real. This loss is calculated using the Mean Squared Error (MSE) between the discriminator's output on generated images and a label indicating "real". The generator aims to minimize this loss to improve its ability to create images that convincingly resemble real high-resolution images, effectively fooling the discriminator.

$$Loss_{GAN} = MSE(discriminator(gen\_hr), valid) \qquad (5)$$

- The content loss ($Loss_{content}$) measures the perceptual difference between the generated image and the real high-resolution image. This loss is calculated using the $L1$ loss between the features of the generated image ($gen_{features}$) and the real image ($real_{features}$). These features are typically extracted from a pretrained VGG network, which captures high-level details important for human perception. The primary purpose of content loss is to ensure that the generated images retain the same perceptual content as the real images, making them visually similar in terms of structure and details that are significant for human observers.

$$Loss_{content} = L1(gen_{features}, real_{features}) \qquad (6)$$

- The $Loss_G$ is the total loss for the generator, calculated by combining the adversarial loss ($Loss_{GAN}$) and the content loss ($Loss_{content}$), where λ is a weight factor, typically set to 1e-3, that balances the two components. The purpose of $Loss_G$ is to guide the generator in producing high-resolution images that are indistinguishable from real ones. The adversarial loss component ensures that the generated images can successfully deceive the discriminator, while the content loss component guarantees that these images retain important visual features, making them perceptually similar to the high-resolution originals. By minimizing $Loss_G$, the generator learns

to create images that are both realistic and true to the content of the original images.

$$\text{Loss}_G = \text{Loss}_{content} + \lambda \times \text{Loss}_{GAN} \quad (7)$$

- The $Loss_{real}$ measures the performance of the discriminator when evaluating real high-resolution images. It is calculated using the Mean Squared Error (MSE) between the discriminator's output for real images (discriminator(imgs_hr)) and the label indicating "real" (typically a tensor of ones, represented as valid). This loss quantifies how accurately the discriminator identifies real images as authentic. A lower $Loss_{real}$ indicates that the discriminator is effectively recognizing real images as real, contributing to the overall training dynamics between the generator and discriminator in the GS-SRGAN framework.

$$\text{Loss}_{real} = \text{MSE}(\text{discriminator}(\text{imgs\_hr}), \text{valid}) \quad (8)$$

- The $Loss_{fake}$ is the loss associated with the discriminator's performance on fake (generated) images. It is calculated using the Mean Squared Error (MSE) between the discriminator's output for generated images (discriminator(gen_hr.detach())) and the label indicating "fake" (typically a tensor of zeros, represented as fake). This loss measures how well the discriminator can identify generated images as fake. By reducing $Loss_{fake}$, the discriminator becomes more effective at distinguishing generated images from real ones, which in turn encourages the generator to improve the realism of its output to better fool the discriminator.

$$\text{Loss}_{fake} = \text{MSE}(\text{discriminator}(\text{gen\_hr}. \text{detach}(\,)), \text{fake}) \quad (9)$$

- The $Loss_D$ is the total loss for the discriminator, representing its performance in distinguishing between real and fake images. This loss is calculated by averaging the losses from classifying both real ($Loss_{real}$) and fake ($Loss_{fake}$) images. The purpose of $Loss_D$ is to ensure that the discriminator can accurately differentiate between authentic high-resolution images and those generated by the generator. By minimizing this loss, the discriminator improves its ability to correctly classify real and fake images, maintaining the balance in the adversarial training process between the generator and discriminator.

$$\text{Loss}_D = \frac{\text{Loss}_{real} + \text{Loss}_{Fake}}{2} \quad (10)$$

7. Experiments and Results

In this section, we present the experimental setup and results of training the GS-SRGAN model with and without learning rate decay. The effectiveness of the model is evaluated using various loss metrics and image quality metrics such as PSNR and SSIM. Additionally, we provide visual comparisons of the generated super-resolution images against the low-resolution and high-resolution ground truth images.

7.1. GS-SRGAN with Learning Rate Decay

In this experiment, we trained the GS-SRGAN model with a learning rate decay strategy. This approach aims to stabilize the training process and enhance the quality of generated high-resolution images over time.
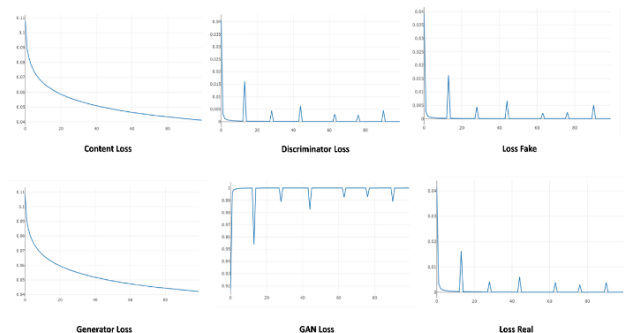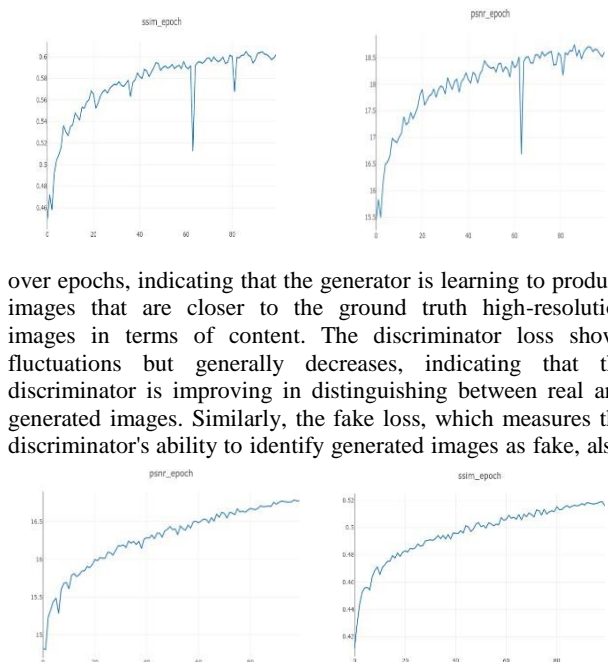

Figure 10. Training Progress with Learning Rate Decay.

As presented in Figure 10, the content loss decreases steadily

over epochs, indicating that the generator is learning to produce images that are closer to the ground truth high-resolution images in terms of content. The discriminator loss shows fluctuations but generally decreases, indicating that the discriminator is improving in distinguishing between real and generated images. Similarly, the fake loss, which measures the discriminator's ability to identify generated images as fake, also

fluctuates but generally decreases, showing improvement in the discriminator's performance. The generator loss decreases steadily, indicating that the generator is improving in producing realistic high-resolution images. Although the GAN loss shows some fluctuations, it remains relatively high, which might indicate that the adversarial training is challenging, but the generator is still improving. Lastly, the real loss, which measures the discriminator's ability to identify real images as real, shows fluctuations but generally decreases, indicating improvement in the discriminator's performance.

Figure 11. Image Quality Metrics with Learning Rate Decay.

In Figure 11, the PSNR (Peak Signal-to-Noise Ratio) demonstrates a steady increase, indicating that the quality of the generated images is improving over epochs. Likewise, the SSIM (Structural Similarity Index) exhibits a consistent upward trend, suggesting that the structural similarity between the generated and real high-resolution images is progressively enhancing.

7.2. GS-SRGAN without Learning Rate Decay

In this experiment, we trained the GS-SRGAN model without a learning rate decay strategy. This approach helps in understanding the importance of learning rate decay in stabilizing the training process.
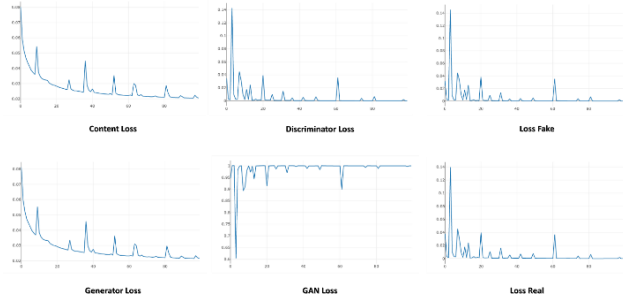

Figure 12. Training Progress without Learning Rate Decay.

Without Learning Rate Decay, as shown in Figure 12, the content loss decreases over epochs, but there are noticeable spikes, indicating instability during training. The discriminator loss is more erratic, suggesting that the discriminator struggles to learn consistently without learning rate decay. Similarly, the fake loss, which measures the discriminator's ability to identify generated images as fake, shows significant spikes, reflecting instability in training the discriminator. The generator loss decreases, but like the content loss, it has multiple spikes indicating instability. The GAN loss is fairly stable with minor fluctuations, suggesting that the adversarial training remains challenging. The real loss shows erratic behaviour, similar to other discriminator-related losses, further indicating instability in the training process.

Figure 13. Image Quality Metrics without learning rate decay.

Without Learning Rate Decay, as illustrated in Figure 13, PSNR increases over epochs but with noticeable drops. This indicates that while image quality improves, it does so inconsistently. Similarly, SSIM also increases with fluctuations, suggesting that structural similarity improves but remains unstable.

The comparison of results between the two training approaches highlights the importance of learning rate decay in stabilizing the training process and achieving consistent improvements in image quality. The model trained with learning rate decay exhibits smoother loss curves and more stable improvements in PSNR and SSIM, indicating better overall performance.
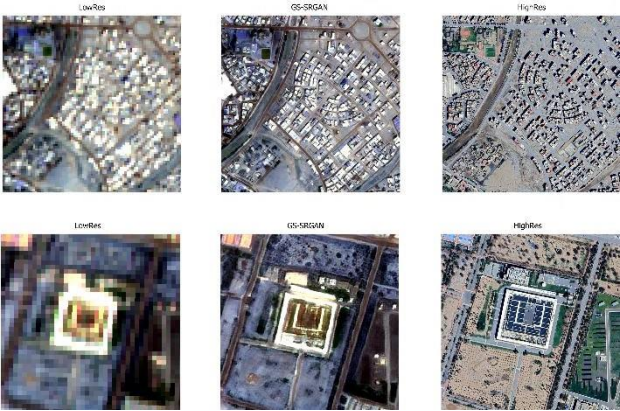
7.3. Visual Comparison

Figure 14. Results and Visual Comparison of Images Quality

Figure 14 shows a comparison between low-resolution (LowRes), super-resolved (GS-SRGAN) generated by the GS-SRGAN model, and high-resolution (HighRes) images. The super-resolved images demonstrate a significant improvement in visual quality over the low-resolution images, with finer details and reduced artifacts, closely resembling the high-resolution ground truth images.

8. Discussion

The analysis of training GS-SRGAN with and without learning rate decay, as summarized in Table 2, reveals central insights into model performance and training stability. With learning rate decay, the training process is significantly more stable, as evidenced by steady decreases in content and generator losses and smoother declines in discriminator-related losses. Conversely, without learning rate decay, the training process shows noticeable spikes and erratic behavior across various loss metrics, indicating instability and the discriminator's struggle to adapt consistently. Image quality metrics, such as PSNR and SSIM, also show more consistent improvements with learning rate decay, reflecting the model's effective learning of high-quality, structurally similar images.

Although the GAN loss remains relatively high with learning rate decay, indicating the ongoing challenge of adversarial training, the overall improvements in other metrics suggest the generator's progressive enhancement. In contrast, the stable GAN loss without decay, coupled with erratic behaviour in other losses, highlights the necessity of learning rate decay for consistent refinement. Thus, learning rate decay proves vital for stable, consistent training and superior image quality in GS-SRGAN. Additional research could explore adaptive learning rates, gradient clipping, and advanced regularization methods to further enhance the performance and stability of GS-SRGAN.

Metric	With LR Decay (Epoch 100)	With LR Decay	Without LR Decay (Epoch 100)	Without LR Decay
Loss Content	0.0411766	Steady decrease	0.0205375	Decreases with spikes
Discriminator Loss	3.96541655e-05	Fluctuations, generally decreases	8.3210883e-05	More erratic
Fake Loss	3.30387001e-05	Fluctuates, generally	6.6042612e-05	Significant spikes

			decreases	
Generator Loss	0.0421767	Steady decrease	0.0215375	Decreases with multiple spikes
GAN Loss	1.0000285	Fluctuates, remains high	1.0000235	Fairly stable, minor fluctuations
Real Loss	4.6269630e-05	Fluctuations, generally decreases	0.0001003	Erratic behavior
PSNR	16.7754355	Steady increase	18.6049153	Increases with noticeable drops
SSIM	0.5160742	Steady increase	0.6022491	Increases with fluctuations

Table 2. Comparison of Losses and Metrics Between Training Results **with** and **without** Learning Rate Decay.

9. Conclusion

In conclusion, this study demonstrates the effectiveness of the GS-SRGAN model in the super-resolution of satellite images, establishing a significant relationship between Geographic Information Systems (GIS) and deep learning (DL). The integration of these fields facilitates a more detailed and accurate spatial analysis, crucial for various environmental and urban planning applications.

One of the primary obstacles in spatial analysis lies in the inherent limitations of satellite image resolution, which can obscure fine details and complicate the extraction of precise information. By employing the GS-SRGAN model, we successfully upscaled Sentinel-2 images, significantly enhancing their resolution and enabling more detailed analysis. The ultimate aim of this research is to accurately identify and delineate rooftops in urban green cities. By estimating their surface areas, we can analyze the potential for solar installations, thereby promoting the utilization of renewable energy sources. This approach not only supports sustainable energy initiatives but also contributes to more efficient urban planning and resource management.

The findings of this study underscore the importance of advanced image processing techniques in overcoming the limitations of traditional satellite imagery. The enhanced resolution achieved through GS-SRGAN not only improves the accuracy of spatial analyses but also opens new avenues for research and application in the fields of GIS (Geographic Information System), DL (Deep learning), and PV (photovoltaic) technology. Future work will focus on refining these methods and exploring their broader implications for environmental sustainability and urban development.

References

Baraskar, T., Raman, V. C., & Panchal, P. (2023a). Implementation of Super Resolution Techniques in Geospatial Satellite Imagery. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(9s), 37–42. https://doi.org/10.17762/ijritcc.v11i9s.7394

Baraskar, T., Raman, V. C., & Panchal, P. (2023b). Implementation of Super Resolution Techniques in Geospatial Satellite Imagery. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(9s), 37–42. https://doi.org/10.17762/ijritcc.v11i9s.7394

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., & Wagner, W. (2019). Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(1), 520–539. https://doi.org/10.1109/TGRS.2018.2858004

Bulat, A., Yang, J., & Tzimiropoulos, G. (2018). *To Learn Image Super-Resolution, Use a GAN to Learn How to Do Image Degradation First* (pp. 187–202). https://doi.org/10.1007/978-3-030-01231-1_12

Chen, H., Zhang, X., Liu, Y., & Zeng, Q. (2019). Generative Adversarial Networks Capabilities for Super-Resolution Reconstruction of Weather Radar Echo Images. *Atmosphere*, *10*(9), 555. https://doi.org/10.3390/atmos10090555

Dong, C., Loy, C. C., He, K., & Tang, X. (2014). *Image Super-Resolution Using Deep Convolutional Networks*. http://arxiv.org/abs/1501.00092

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012a). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012b). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012c). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Fernández-Manso, A., Quintano, C., & Roberts, D. (2012). Evaluation of potential of multiple endmember spectral mixture analysis (MESMA) for surface coal mining affected area mapping in different world forest ecosystems. *Remote Sensing of Environment*, *127*, 181–193. https://doi.org/10.1016/j.rse.2012.08.028

Galagan, M., & Frigerio, A. (2023). The Contribution of Renewable Energy to International Security: A Preliminary Study. *International Relations and International Law Journal*, *101*(1). https://doi.org/10.26577/IRILJ.2023.v101.i1.012

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., … Fernandez, V. (2017a). Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sensing*, *9*(6), 584. https://doi.org/10.3390/rs9060584

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., … Fernandez, V. (2017b). Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sensing*, *9*(6), 584. https://doi.org/10.3390/rs9060584

Gómez, C., White, J. C., & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *116*, 55–72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Networks. *Science Robotics*, *3*(January), 2672–2680. https://arxiv.org/abs/1406.2661v1

Gou, S., Liu, S., Yang, S., & Jiao, L. (2014). Remote Sensing Image Super-Resolution Reconstruction Based on Nonlocal Pairwise Dictionaries and Double Regularization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(12), 4784–4792. https://doi.org/10.1109/JSTARS.2014.2328596

Goyal, R., Patel, S., & Sharma, A. (Eds.). (2024). *Renewable Energy: Accelerating the Energy Transition*. Springer Nature Singapore. https://doi.org/10.1007/978-981-99-6116-0

Haut, J. M., Fernandez-Beltran, R., Paoletti, M. E., Plaza, J., & Plaza, A. (2019). Remote Sensing Image Superresolution Using Deep Residual Channel Attention. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(11), 9277–9289. https://doi.org/10.1109/TGRS.2019.2924818

Haut, J. M., Paoletti, M. E., Fernandez-Beltran, R., Plaza, J., Plaza, A., & Li, J. (2019). Remote Sensing Single-Image Superresolution Based on a Deep Compendium Model. *IEEE Geoscience and Remote Sensing Letters*, *16*(9), 1432–1436. https://doi.org/10.1109/LGRS.2019.2899576

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hu, J., Hong, D., Wang, Y., & Zhu, X. X. (2019). A Comparative Review of Manifold Learning Techniques for Hyperspectral and Polarimetric SAR Image Fusion. *Remote Sensing*, *11*(6), 681. https://doi.org/10.3390/rs11060681

Huang, Y., Shao, L., & Frangi, A. F. (2017a). Simultaneous Super-Resolution and Cross-Modality Synthesis of 3D Medical Images Using Weakly-Supervised Joint Convolutional Sparse Coding. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5787–5796. https://doi.org/10.1109/CVPR.2017.613

Huang, Y., Shao, L., & Frangi, A. F. (2017b). Simultaneous Super-Resolution and Cross-Modality Synthesis of 3D Medical Images Using Weakly-Supervised Joint Convolutional Sparse Coding. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5787–5796. https://doi.org/10.1109/CVPR.2017.613

Huang, Y., Shao, L., & Frangi, A. F. (2017c). Simultaneous Super-Resolution and Cross-Modality Synthesis of 3D Medical Images Using Weakly-Supervised Joint Convolutional Sparse Coding. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5787–5796. https://doi.org/10.1109/CVPR.2017.613

Johnson, J., Alahi, A., & Fei-Fei, L. (n.d.). *Perceptual Losses for Real-Time Style Transfer and Super-Resolution*.

Lei, S., Shi, Z., & Zou, Z. (2017). Super-Resolution for Remote Sensing Images via Local–Global Combined Network. *IEEE Geoscience and Remote Sensing Letters*, *14*(8), 1243–1247. https://doi.org/10.1109/LGRS.2017.2704122

Li, D. (2023). Opportunities and Challenges for Solar Cells. *Highlights in Science, Engineering and Technology*, *59*, 129–136. https://doi.org/10.54097/hset.v59i.10072

Li, J., Yuan, Q., Shen, H., Meng, X., & Zhang, L. (2016). Hyperspectral Image Super-Resolution by Spectral Mixture Analysis and Spatial–Spectral Group Sparsity. *IEEE Geoscience and Remote Sensing Letters*, *13*(9), 1250–1254. https://doi.org/10.1109/LGRS.2016.2579661

Lu, J., Li, J., Yan, Z., Mei, F., & Zhang, C. (2018). Attribute-Based Synthetic Network (ABS-Net): Learning more from pseudo feature representations. *Pattern Recognition*, *80*, 129–142. https://doi.org/10.1016/j.patcog.2018.03.006

Ma, W., Pan, Z., Guo, J., & Lei, B. (2018). Super-Resolution of Remote Sensing Images Based on Transferred Generative Adversarial Network. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 1148–1151. https://doi.org/10.1109/IGARSS.2018.8517442

Mushkin, A., Gillespie, A. R., Abbott, E. A., Batbaatar, J., Hulley, G., Tan, H., Tratt, D. M., & N. Buckland, K. (2020). Validation of ASTER Emissivity Retrieval Using the Mako Airborne TIR Imaging Spectrometer at the Algodones Dune Field in Southern California, USA. *Remote Sensing*, *12*(5), 815. https://doi.org/10.3390/rs12050815

Popovych, V., Rudniev, Y., Brozhko, R., Tarasov, V., Tymchenko, O., & Aliekperov, I. (2023). Photovoltaics: Problems, Challenges, Environmental Feasibility and Efficiency. *2023 IEEE 5th International Conference on Modern Electrical and Energy System (MEES)*, 1–6. https://doi.org/10.1109/MEES61502.2023.10402385

Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*. https://arxiv.org/abs/1511.06434v2

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf, C. B., Schott, J. R., Sheng, Y., Vermote, E. F., Belward, A. S., Bindschadler, R., Cohen, W. B., Gao, F., … Zhu, Z. (2014a). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172. https://doi.org/10.1016/j.rse.2014.02.001

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf, C. B., Schott, J. R., Sheng, Y., Vermote, E. F., Belward, A. S., Bindschadler, R., Cohen, W. B., Gao, F., … Zhu, Z. (2014b). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172. https://doi.org/10.1016/j.rse.2014.02.001

Salgueiro Romero, L. F., Marcello, J., & Vilaplana, V. (2020). Comparative study of upsampling methods for super-resolution in remote sensing. In W. Osten & D. P. Nikolaev (Eds.), *Twelfth International Conference on Machine Vision (ICMV 2019)* (p. 44). SPIE. https://doi.org/10.1117/12.2557357

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

Saripalli, B. P., Guduru, A. K., & Yeruva, J. R. (2024). *Opportunities and Challenges in Solar Photovoltaic Waste Management* (pp. 217–238). https://doi.org/10.4018/979-8-3693-1018-2.ch015

Selvaraj, V., Thamizhchelvan, A., Franklin, S., & Isaac, J. R. (2024). *Experimental analysis on enhanced photovoltaic energy generation through concentrated optics and passive liquid cooling*. 040020. https://doi.org/10.1063/5.0190144

Simonyan, K., & Zisserman, A. (2014a). Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://arxiv.org/abs/1409.1556v6

Simonyan, K., & Zisserman, A. (2014b). Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://arxiv.org/abs/1409.1556v6

Simonyan, K., & Zisserman, A. (2014c). Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://arxiv.org/abs/1409.1556v6

Vuolo, F., Mattiuzzi, M., & Atzberger, C. (2015). Comparison of the Landsat Surface Reflectance Climate Data Record (CDR) and manually atmospherically corrected data in a semi-arid European study area. *International Journal of Applied Earth Observation and Geoinformation*, *42*, 1–10. https://doi.org/10.1016/j.jag.2015.05.003

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Loy, C. C. (n.d.). *ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks*. https://github.com/xinntao/ESRGAN.

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Loy, C. C. (2019). *ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks* (pp. 63–79). https://doi.org/10.1007/978-3-030-11021-5_5

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. https://doi.org/10.1109/TIP.2003.819861

Wang, Z., Chen, J., & Hoi, S. C. H. (2019). *Deep Learning for Image Super-resolution: A Survey*.

Weisheng Dong, Lei Zhang, Guangming Shi, & Xiaolin Wu. (2011). Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization. *IEEE Transactions on Image Processing*, *20*(7), 1838–1857. https://doi.org/10.1109/TIP.2011.2108306

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., Allen, R. G., Anderson, M. C., Belward, A. S., Cohen, W. B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J. D., Hostert, P., Hughes, M. J., … Zhu, Z. (2019). Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, *225*, 127–147. https://doi.org/10.1016/j.rse.2019.02.015

Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., Fosnight, E. A., Shaw, J., Masek, J. G., & Roy, D. P. (2016). The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment*, *185*, 271–283. https://doi.org/10.1016/j.rse.2015.11.032

Xin Li, & Orchard, M. T. (2001). New edge-directed interpolation. *IEEE Transactions on Image Processing*, *10*(10), 1521–1527. https://doi.org/10.1109/83.951537

You, D. (2022). *Overview of Photovoltaic Product Benefits and Photovoltaic Development Suggestions*. https://doi.org/10.2991/aebmr.k.220307.168

Zhang, J., Tian, H., Wang, D., Li, H., & Mouazen, A. M. (2020). A Novel Approach for Estimation of Above-Ground Biomass of Sugar Beet Based on Wavelength Selection and Optimized Support Vector Machine. *Remote Sensing*, *12*(4), 620. https://doi.org/10.3390/rs12040620

Zhang, Y., Du, Y., Ling, F., Fang, S., & Li, X. (2014). Example-Based Super-Resolution Land Cover Mapping Using Support Vector Regression. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(4), 1271–1283. https://doi.org/10.1109/JSTARS.2014.2305652