RGNZ/MGPB

Rudarsko-geološko-naftni zbornik
(The Mining-Geology-Petroleum Engineering Bulletin)
DOI: 10.17794/rgn.2025.3.8

Review scientific paper

Satellite Image Enhancement Using Deep Learning and GIS Integration: A Comprehensive Review

Dalia A. Hussein[1]*, Mohamed A. Yousef[1], Hassan A. Abdel-Hak[1], Yasser G. Mostafa[2]

[1] *Faculty of Engineering, University of Assiut, Egypt.*
[2] *Faculty of Engineering, University of Sohag, Egypt.*

Abstract

A comprehensive review of 32 studies (20 journals, 11 proceedings, and one book chapter) published from 2016 to 2023 in the fields of deep learning (DL), image enhancement, super-resolution image, and Geographic Information System (GIS) is presented, focusing on the integration of DL methodologies with GIS to improve the quality of satellite images. The review summarizes the background, principles, enhancement quality, speed, and advantages of these technologies, comparing their performance based on metrics such as Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Structural Similarity Index Measure (SSIM), and computation time. Satellite remote sensing technologies, which have provided an efficient means of gathering spatial information since the launch of Landsat 1 by NASA in 1972, have recently advanced to enable the collection of high-resolution satellite (HRS) images (≤30 cm). However, factors such as atmospheric interference, shadowing, and underutilization of sensor capacity often degrade image quality. To address this, satellite images require enhancement, and DL has emerged as a powerful tool due to its ability to model complex relationships and accurately recover super-resolution images. While DL and neural networks have demonstrated significant success in natural image enhancement, their application to satellite images presents unique challenges. These challenges include insufficient consideration of the distinct characteristics of satellite imagery, such as varying spatial resolutions, sensor noise, and spectral diversity, as well as the reliance on modelling assumptions that may not align with the complexities of satellite data. This highlights the need for further investigation into advanced DL approaches tailored specifically for this domain.

Keywords:
deep learning, GIS, neural networks, satellite images, image enhancement, super resolution

1. Introduction

During the past few decades, Geographical Information Systems (GIS) have become a fundamental tool for scientific and decision-making analysis in diverse fields such as agriculture, environment, telecommunication, water management, construction, sports, transportation, and others. A key reason for GIS's success lies in its ability to integrate, analyze, and visualize spatial data effectively. In recent years, advancements in artificial intelligence (AI) have further amplified GIS's capabilities, enabling more sophisticated analyses and predictions. Among these advancements, deep-learning-based methods, particularly Convolutional Neural Networks (CNNs), have shown exceptional potential in processing and interpreting satellite imagery – an integral data source for GIS applications (**Zhang et al., 2016**). However, the effective use of deep learning in GIS requires sufficient computational resources, robust datasets, and careful tuning of hyperparameters (**Goodfellow, 2016**). When such resources are available, these AI-driven models can enhance tasks such as evaluation, classification, clustering, function approximation, and data visualization, providing new opportunities for innovation in GIS-based applications.

CNNs use has greatly enhanced a number of GIS tasks, including model training and the identification of separated objects across vast areas. However, conventional approaches like image fusion, principal component analysis (PCA) (**Pohl et al., 1998**), and histogram equalization were frequently used for satellite image enhancement before deep learning techniques were adopted. In order to provide the foundation for further research, these techniques sought to preprocess and improve the visual quality or analytical utility of satellite imagery. Even though remarkably deep learning models have been developed, their integration with GIS has opened new avenues for spatial data analysis, such as enhancing satellite image interpretation, automating land use classification, improving environmental moni-

* Corresponding author: Dalia A. Hussein
e-mail address: dalia.alip31@eng.aun.edu.eg

Received: 12 September 2024. Accepted: 1 February 2025.
Available online: 3 July 2025

Table 1. Classification of satellite image enhancement techniques

Category	References	Key Contributions
1. Methods		
Generative Adversarial Networks (GANs)	- **(Jiang et al., 2019)** - **(Wang et al., 2018)** - **(Wenlong et al., 2021)**	- Introduced GAN-based models for super-resolution. - Enhanced edge quality. - ESRGAN showed perceptual quality improvements. - RankSRGAN introduced ranking for better results.
Convolutional Neural Networks (CNNs)	- **(Kattenborn et al., 2021)** - **(Yamashita et al., 2018)** - **(Lei et al., 2017)**	- Reviewed CNN applications in vegetation and radiology. - Developed local-global optimization networks for enhanced super-resolution.
Dense Skip Connections	- **(Tong et al., 2017a)**	- Introduced dense skip connections to improve image feature retention in super-resolution tasks.
Evolutionary Computation	- **(Wang et al., 2024)**	- Surveyed evolutionary algorithms in enhancing GANs for image applications.
Hybrid Networks	- **(Fırat et al., 2023)**	- Combined 3D/2D convolutional networks for hyper-spectral image classification.
2. Success Rates and Comparative Analysis		
High Success in Super-Resolution	- **(Shi et al., 2016)** - **(Kim et al., 2016a, 2016b)**	- Sub-pixel CNNs achieved real-time efficiency. - Very deep CNNs and recursive networks showed high accuracy in super-resolution.
Comparison Between Methods	- **(Singla et al., 2022)** - **(Zhang et al., 2022)**	- Reviewed single-image super-resolution GANs. - Developed terrain-specific super-resolution networks.
Limitations Addressed	- **(Sharma et al., 2021)**	- Proposed improvements in satellite image enhancement techniques.
3. Progress of Technology		
Emerging Techniques (2015–2017)	- **(Dong et al., 2015)** - **(Tong et al., 2017a)**	- Early CNNs for super-resolution. - Dense skip connections introduced.
Advanced Architectures (2018–2023)	- **(Wang et al., 2018)** - **(Lei et al., 2017)**	- Advanced GANs (e.g. ESRGAN) and global-local optimization for improved results.
Current Trends (2023–2024)	- **(Wenlong et al., 2021)** - **(Noshiri et al., 2023)**	- GANs with rankers for targeted applications (e.g. RankSRGAN). - Focus on explainability and computational efficiency.
4. Key Comparative Insights		
Best Performers	- **(Wang et al., 2018)** - **(Wenlong et al., 2021)**	- GANs like ESRGAN and RankSRGAN achieved state-of-the-art results in perceptual quality and accuracy.
Trade-Offs	- **(Shi et al., 2016)** - **(Kim et al., 2016a, 2016b)**	- Sub-pixel CNNs are efficient but may not achieve perceptual quality like GANs.
Specific Use Cases	- **(Kattenborn et al., 2021)** - **(Zhang et al., 2022)**	- Applications in vegetation monitoring and terrain super-resolution.

toring, and optimizing spatial decision-making processes (Zhang et al., 2016). However, the literature has not fully examined these crucial connections in recent years (Oveis et al., 2021). To give a thorough overview of the developments in deep learning methods for satellite image augmentation, Table 1 classifies recent studies based on their methodologies, success rates, and comparative insights. This classification highlights the range of techniques, from traditional convolutional neural networks (CNNs) to advanced generative adversarial networks (GANs) and hybrid approaches. The evolution of technology, the benefits and drawbacks of various strategies, and their suitability for particular use cases, such as vegetation monitoring and terrain, super-resolution are also highlighted. The table provides a succinct reference for comprehending cutting-edge methods and their influence on enhancing satellite imaging by arranging these investigations.

Deep learning (DL) with its deep hierarchical architectures for representation learning of patterns has shown strong capabilities as well as state-of-the-art performance in many visual recognition and machine learning tasks, especially in image processing, analysis, and reconstruction. The very well-known deep architectures integrated with DL technology include deep stacked autoencoders (DSAE), convolutional deep belief networks (CDBN), and convolutional neural networks (CNN) or simply deep convolutional networks (DCN) (**Chen et al., 2023**).

One of the time-consuming, large-scale, and complex data types widely used for visualization and exploration of the Earth's surface present in GIS is remote sensing data, especially satellite images. These images are taken by Earth observation systems aboard satellites orbiting the Earth. They are then used to detect and extract information in remotely located areas and time points on Earth's surface, especially in geography, geosciences, GIS, environmental research, and climate change (**Tsatsaris et al., 2021**).

In terms of recent research, the use of DL for satellite image reconstruction, enhancement, super-resolution, spectroscopy, pan-sharpening, shadow removal, cloud and fog removal, color correction, mosaicking, denoising, fusion, in-painting, time series generation, and geographic feature classification is a quickly growing research field. However, these DL satellite image enhancements must be combined with the real-world utility of GIS to create practical impacts on satellite image data for real-world application domains.

This article thus gave a comprehensive survey and comprehensive review of various integrated deep learning methods, state-of-the-art benchmarks, international database test cases, the potential of deep reinforcement learning, and transfer learning approaches. The main objective of this review is to cover and organize a substantial part of the research material, methodologies, deep learning techniques, and advancements focusing on the deep learning integrated with the GIS technique to enhance unsupervised and supervised satellite image classification, and also suggest future directions to this scientific community to work on. Researchers are encouraged to use this review to rapidly gain some know-how of this novel scientific area and also get clarity in thematic aspects. This article organizes research related to the application of deep learning solutions in satellite imagery and the integration of GIS techniques aimed at enhancing various types of satellite images.

The growing importance of applications that involve satellite images has naturally boosted research in this area. With the growing demand for fast and efficient solutions, the use of deep learning techniques stands out. In the search for applications that optimize the use of these techniques, researchers have increasingly sought to include geographic information in the processes, including characteristics of spatial data, and grouping various GIS concepts in the proposed applications. The aim is to contribute to this area, identifying what has been covered by the review, bringing together research clusters and possible developments. In addition to deep learning matters, in this review, we have also addressed the most important topics presented by the researchers.

2. Enhanced remote sensing image processing techniques

Enhancing satellite images traditionally involves a sequence of procedures such as image distortion, noise removal, enhancement, filtering, contrast stretching, and morphological processing and their associated inversion operation to improve the spatial resolution of images. **Lee (1980)** refined the low-pass-filtered image using a nonlinear function, followed by a median filtering operation to suppress grain noise. **Sharma et al. (2021)** proposed a method for enhancement of remote sensing images based on the combination of the picking rule that preserves the radiometric mean value and entropy based on spatial frequency structures. However, while these traditional methods improve certain aspects of image clarity, they often struggle with enhancing image contrast effectively without introducing noise or distortion.

To achieve clearer image details, modern techniques focus on amplifying contrast while maintaining spatial and radiometric fidelity. One such approach combines multiple techniques, such as Compound Contrast Limited Adaptive Histogram Equalization (CLAHE), which reduces noise by limiting amplification in homogeneous regions, and the change-intensity-histogram method integrated with a sigmoid function for non-linear contrast enhancement (**Zhang et al., 2024**). These techniques aim to provide a more refined enhancement by addressing the limitations of traditional methods, especially in terms of contrast enhancement and noise suppression. Fusion methods, such as those based on the ratio of limiting variance and differential mean techniques, are also used to merge complementary features for better visual quality. Additionally, Poisson noise reduction techniques are employed to handle noise inherent in low-light or low-quality images (**Chatterjee et al., 2011**). These methods can be useful to enhance specific geographic features, but the proposed general outcome is based on a combination of multi-image wavelet and other deep learning. Hyper-spectral Pixel Matching (HPM) was used to ensure accurate mapping and classification by matching pixel information from hyper-spectral images to a reference, and learners should be informed about this technique with applications. It ensures that the class of the pixel generated is the same as the reference land cover class from which it was sampled (**Fu et al., 2015**).

The following subsections explore advanced techniques for remote sensing image processing, including super-resolution (Section 2.1), which leverages deep learning to enhance low-quality satellite images; pan-sharpening (Section 2.2), which integrates panchromatic and multi-spectral data for high-resolution outputs; and image fusion (Section 2.3), which combines spatial, spectral, and temporal information to enhance image quality.

2.1. Super-Resolution

The spatial resolution of satellite images is an essential factor in remote sensing. The high spatial resolution of satellite data enhances the ability for land-cover classification and change detection. However, most countries receive satellite data with lower resolution due to the higher cost of high-resolution satellite data. The cost of the satellite images should be affordable to allow frequent acquisition, which is essential for monitoring and management tasks (**Chen, 2007; Fisher et al., 2018**).

Deep learning has been gaining in popularity recently in a variety of domains, including natural language processing and computer vision. This study investigates the use of deep learning techniques to improve satellite imagery, namely in the areas of categorization and super-resolution. Super-resolution is the process of producing higher-quality images by computationally increasing the resolution of low-resolution images, usually with the use of sophisticated algorithms like deep learning. High-resolution satellite data, on the other hand, is directly recorded by sophisticated satellite sensors with a finer native spatial resolution. When high-resolution data is limited, costly, or unavailable, super-resolution techniques are extremely helpful. Applications like land cover classification, which aids in monitoring natural disasters, identifying unlawful farms or structures, and other vital duties, require high-resolution data. Since most satellite-captured images have a lower spatial resolution than those from ground-based tools, super-resolution techniques are crucial for improving them including super-resolution (**Ma et al., 2019**).

2.2. Pan-Sharpening

Pan-sharpening, as a process, results in creating high spatial resolution multi-spectral images with panchromatic images. This operation has become important, especially with the increase in the use of satellite images for any changes or disasters. Reducing the spatial resolution of an image provides a better image, and merging this process with multi-spectral information provides better results. In the previous studies carried out for both pan-sharpening and deep learning, the aim was to improve the results. The image created by deep neural network models would not be pan-sharpened; it only enhances the low-resolution image (**Santurri et al., 2012**).

According to various studies, Landsat ETM+ and OLI-TIRS imagery have been utilized for different applications in remote sensing, including landslide prediction and land-cover classification (**Roy et al., 2014**). A review of the literature reveals that a total of 24 spectral models have been identified, each tailored to specific regions by combining different spectral bands. These combinations of bands are selected to highlight particular features relevant to the region, such as soil moisture or vegetation, which are crucial for accurate land-cover mapping and hazard assessment (**Jensen, 2009**).

Newer algorithms like wavelet-based fusion techniques and Gram-Schmidt spectral sharpening were created to overcome these constraints. Wavelet-based methods are useful for preserving spectral and spatial features since they enable multiscale picture decomposition. Recent developments in wavelet-based methods have made strides in multi-resolution image fusion, especially in the context of satellite and remote sensing imagery, where maintaining high spectral fidelity while enhancing spatial resolution is crucial. For example, recent work has explored how the spectral response of sensors can be better integrated into these fusion algorithms, significantly improving their utility in practical applications like satellite image enhancement and geospatial analysis (**Dibs et al., 2023**). Further innovation is required because, in spite of these developments, these conventional approaches remain limited by their incapacity to adjust to different sensor kinds and spatial resolutions.

To get over the drawbacks of conventional techniques, deep learning techniques have been used more and more in pan-sharpening in recent years. Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs) have become effective tools for this. Deep learning models, in contrast to conventional techniques, are data-driven and are capable of learning complex associations between multi-spectral and panchromatic images without the need for preset mathematical adjustments (**Masi et al., 2016**).

The application of CNNs to improve the spatial resolution of multi-spectral photographs is one such instance. To teach these models the mapping needed for image enhancement, paired high-resolution and low-resolution images are used for training. Conversely, GANs use adversarial training, in which a discriminator assesses the quality of high-resolution images produced by a generator. The integrity and quality of the fused images are enhanced by this adversarial approach (**Deng et al., 2020**).

Deep learning techniques are not without difficulties, though. The need for vast quantities of high-quality training data, which might not always be accessible for particular sensors or geographical areas, is one of the main problems. Widespread adoption may also be hampered by the computing requirements of training and implementing these models.

2.3. Image Fusion

Image fusion is the combination of information from different images/sources of a scene. Remote sensing image fusion can be performed in three domains, namely, spatial, spectral, and temporal domain. Spatial domain fusion is the process of using different resolutions to represent the area of interest (**Hong et al., 2016**). For example, higher spatial resolution images of one sensor can be fused with coarser spatial resolution images of other sensors. Another image fusion technique in this domain is to use a single imaging sensor to acquire images which are taken at slightly different viewing geometries such that the area of interest can have different illumination areas with different spatial resolutions. Spectral domain fusion is the fusion of spectral information from different sensors, i.e. merging spectral bands with high spatial resolution and low spatial resolution. Temporal domain fusion is the process of combining images taken at different times to create an image with improved qualities, either by combining images captured closely or by combining images captured at different times (**Belgiu et al., 2019**).

3. Deep Learning techniques in super resolution satellite images

Earth exploration has long made use of remote sensing imagery, and many techniques have been developed to recover lost data because of satellite technology limits and data transmission channels. The resolution and quality of gradient maps have greatly increased recently, improving item classification accuracy. Researchers now have access to high-quality surface images thanks to advancements in satellite and aerial photography, which can be used to test and train machine learning algorithms to improve image quality. This has also made it possible to recover small details lost during compression by using super-resolution methods. In order to produce high-resolution (HR) outputs with a higher pixel density, super-resolution (SR) techniques use deep neural networks to train artificial intelligence models that predict objects and elements in low-resolution (LR) images. **Dong et al. (2015)** first proposed a three-layer autoencoder as the foundation for these techniques. Based on the generative adversarial networks (GAN) framework, **Radford (2015)** presented a novel image processing architecture in 2015 termed DCGAN (Deep Convolutional Generative Adversarial Network). Despite not being created especially for image upsampling, this work had a big influence on the advancement of super-resolution methods. The size of the necessary training dataset was decreased by streamlining the training procedure with the help of the GAN framework. **Ledig et al. (2017)** proposed SRGAN, one of the first successful SR architectures utilizing the GAN technique. When it came to recovering details in general-purpose photos, this neural network performed admirably. Nevertheless, it had several shortcomings, including the tendency for the network to produce artefacts at abrupt gradient changes during training, and the frequency and quantity of these artifacts were correlated with the number of Residuals in Residual Dense Blocks (RRDB). Since deeper networks are necessary to successfully recreate surface features, this constraint made the architecture unsuitable for satellite image restoration. In order to enhance SRGAN's performance, recent studies have concentrated on reducing the creation of artefacts in high-resolution photos. The architecture was improved by the **Wang et al. (2018)** and **Rakotonirina et al. (2020)** teams, which increased its stability during training. The new ESRGAN+ model has Gaussian noise and Residual Scaling (RS) layers instead of Batch Normalization (BN) layers, although it still has a comparable high-level structure. These changes have made it possible to employ ESRGAN+ as a general-purpose super-resolution technique that may enhance missing portions while maintaining the high reconstruction accuracy of the original image. Furthermore, a new architecture called Progressive Enhanced Generative Adversarial Network (PEGAN) was presented by **Jing et al. (2022)** for super-resolution with high-level image amplification. PEGAN uses an ensemble of tiny networks to process the image, building on earlier ESRGAN work. Prior to being processed by two networks – one that uses an autoencoder to recover structural information and another that uses a multi-pass straight network to restore high-frequency features – it first extracts low-frequency components. The resulting resolution remained less than one meter per pixel after amplification, making this method inadequate for satellite image processing with a precision of 1 pixel every 8 meters, even if it produced images of nearly original quality when scaling by a factor of 4. By jointly learning a low-rank dictionary pair from overlapping hyper-spectral and multi-spectral regions, **Gao et al. (2020)** presented a novel technique to improve the spectral resolution of multi-spectral images. But, as the authors pointed out, their method isn't appropriate for complicated or large-scale instances. **Sharshov et al. (2024)** present a new method for super-resolution in remote sensing by combining CNN-based SRGAN with GIS data. In order to improve picture reconstruction, this technique uses ancillary data including Normalized Difference Vegetation Index (NDVI) composites, water extent, and terrain elevation in conjunction with a GIS Data Features block to increase residual scaling. With gains in the Peak Signal-to-Noise Ratio (PSNR) from 27.3630 to 27.3986 and the Structural Similarity Index (SSIM) from 0.9632 to 0.9673, as well as better visual quality in reconstructed images, the suggested GESRGAN model outperformed ESRGAN. The approach improved training efficiency and decreased the requirement for huge datasets by processing GIS data and implementing an effective gating mechanism. These developments open up the opportunity for more improvements in GIS data processing and larger-scale applications by demonstrating the promise of GIS-integrated super-resolution techniques for satellite image reconstruction, especially in situations where high-resolution data is scarce.

3.1. Convolutional Neural Networks (CNNs)

CNNs are suitable for working with grid data, including raster map data. Image data can have different pattern types because images are multidimensional data. A regular CNN model extracts the features based on the filter; however, feature extraction ignores the spatial relationship between the features. In other words, CNN does not change the spatial data, but the feature extraction is changed. For spatial data, searching for local geospatial structures (such as patches and objects) is an interesting concept, and deep learning introduces this by analyzing the spatial data with an extra layer. Thus, a deep CNN model uses hierarchical features, which represent the entire structure of the spatial relationship. Since remote sensing images require high computation

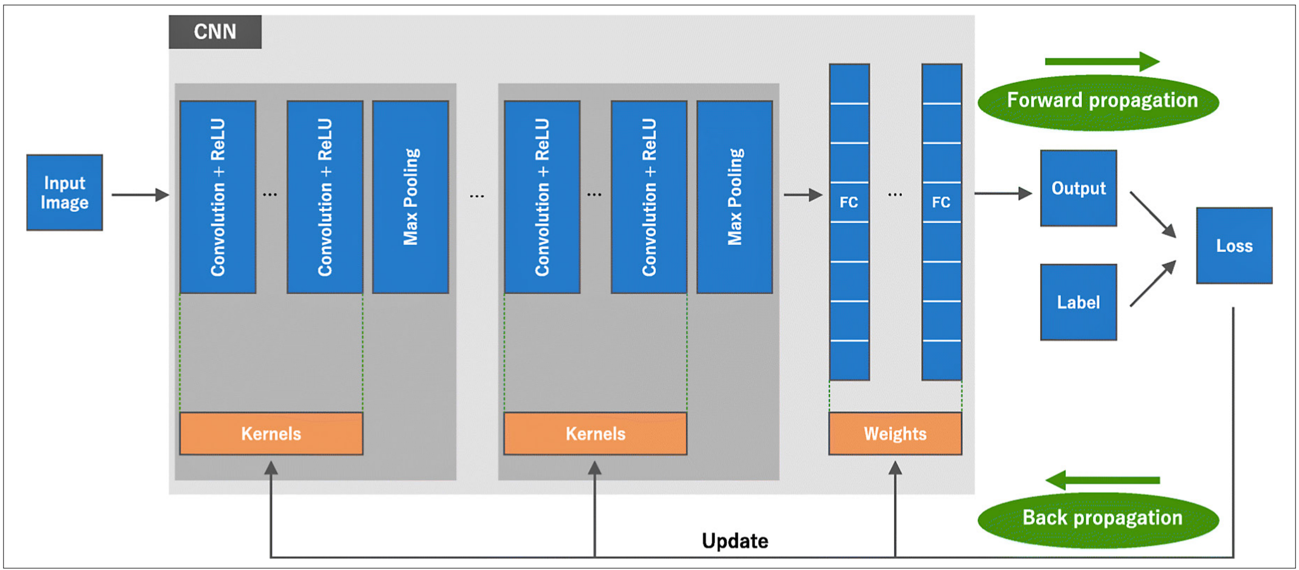

Figure 1. Architecture of a convolutional neural network (CNN) (from **Yamashita et al., 2018**)

time and complex classification algorithms, trying to find spatial patterns in the high data is time-consuming (**Yamashita et al. 2018**). **Figure 1** illustrates the architecture of a convolutional neural network (CNN).

The primary building parts of CNN models are the Convolutional Layer, Pooling Layer (subsample Layer), Fully Connected Layer, Output Layer, and Flatten Layer, aside from feature extraction. Convolutional kernels, also known as learned features, are the particular filters that are applied to the input image by the convolutional layer. Every time, the input and recommended filter are smaller than the learned filter kernels. Subsequently, the dot products between each object inside the filter's influence range and the learned features are calculated by the convolutional layer. The item will be split and compared to the features in the convolutional layer if it is within the filter's range. The outcome is a feature map that displays the positions and intensities of the input image's learned features. Its weight is its most crucial characteristic. This weight is input earlier rather than later. As an illustration, we can express the weight values as a matrix, where each entry in the matrix is an element that needs to be multiplied by the input components.

3.2. Autoencoders and Generative Adversarial Networks (GANs)

Autoencoders are a type of artificial neural network commonly used for dimensionality reduction and data compression tasks. They excel in unsupervised learning for convolutional or recurrent networks and effectively reduce training costs without requiring labeled data. For instance, **Hinton et al., (2006)** demonstrated the use of autoencoders for dimensionality reduction in high-dimensional data, significantly improving visualization tasks. Later advancements, such as denoising autoencoders, **Vincent et al. (2008)** showed improved robustness by reconstructing corrupted input data, while the variational autoencoders **Kingma (2013)** introduced probabilistic modeling to learn latent variable distributions effectively.

Generative Adversarial Networks (GANs) extend deep learning applications by enabling the generation of high-quality synthetic data. **Goodfellow et al. (2020)**, who introduced GANs, highlighted their ability to generate photorealistic images by adversarial training a generator and a discriminator. Comparatively, **Radford (2015)** proposed the Deep Convolutional GAN (DCGAN), which improved GAN stability and scalability using convolutional layers. Additionally, **Karras (2019)** introduced StyleGAN, which significantly advanced image synthesis by offering control over image style at various levels. Despite these successes, GANs face challenges like mode collapse and training instability, as highlighted by **Salimans et al. (2016)**, who proposed techniques such as feature matching to address these issues.

While autoencoders focus on dimensionality reduction and data reconstruction, GANs aim at high-quality data generation. Studies comparing the two such as **Dumoulin et al. (2016)** suggest that autoencoders perform better in encoding compact representations, whereas GANs excel in creating realistic data, making them complementary in tasks like semi-supervised learning and image enhancement.

3.3. Experimental case studies and analysis

We reviewed four different architectures of Deep Learning specifically focusing on four types of CNN and three types of GANs. The selected CNN models include the widely used Super Resolution Convolutional Neural Network (SRCNN), Local Global Combined Network (LGCNet), Progressively Enhanced Convolutional Neural Network (PECNN), and Deep Distillation Recursive

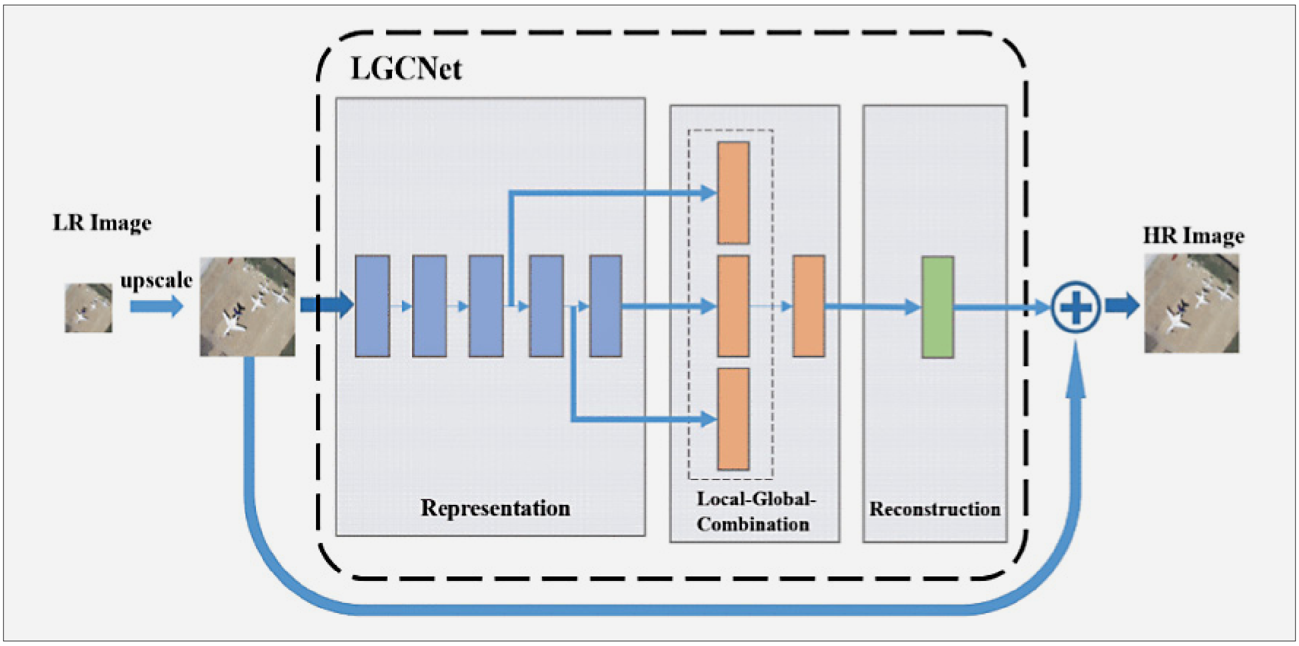

Figure 2. Diagram illustrating the suggested super-resolution technique for remote sensing images (from **Lei et al., 2017**)

Network (DDRN). For GANs, we examined Super-Resolution Generative Adversarial Networks (SRGAN), Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN), and Edge-enhanced GAN (EEGAN) due to their relevance in remote sensing image enhancement tasks.

This section explores how each of these architectures contributes to improving satellite image quality, providing insight into their strengths and limitations. To ensure clarity, we have structured the discussion under separate headings for each model, focusing on its specific applications and performance in super-resolution, pan-sharpening, and image fusion tasks.

3.3.1. Types of CNN (convolutional neural network-based methods)

3.3.1.1. Super Resolution Convolutional Neural Network (SRCNN)

CNNs consist of three primary components: pooling, nonlinear mapping, and convolution. By using these processes, CNNs can be trained and supervised to adaptively convert input picture space to an efficient feature space for a given job. The author only uses convolution and nonlinear mapping operations in the model since it is known that the image super resolution job, the image with low resolution will lose more detailed information after pooling, leading to a worse reconstructed outcome (**Lei et al., 2017**).

The size of inputs in this study is indicated by the symbol X = HHW×C, where C stands for the channel number of the remote sensing images. The outputs following convolution and nonlinear mapping for a network of L convolutional layers can be calculated as follows:

$$f_1(X;\ W_1,\ b_1) = \sigma\ (W_1 * X + b_1) \tag{1}$$

$$f_l\ (X;\ W_l, b_l) = \sigma\ (W_l * f_l - 1(X) + b_l) \tag{2}$$

where:

W_l, b_l,

$l \in$ (1,…, L) are the network weights and bias, that will be learned.

W_1 is the tensor with the size of $k_l \times k_l \times n_l - 1 \times n_l$, in which k_l denotes kernel size at layer one,

and n_l denotes the number of the feature maps at the same layer (n_0=C).

b_l is the vector whose size equals n_l.

Recently, most nonlinear functions are rectified linear functions (max (0, x)), and the nonlinear function σ is an element-wise operation. This allows CNNs to converge considerably faster than previous saturating nonlinearities (**Krizhevsky et al., 2012**).

3.3.1.2. Local Global Combined Network (LGCNet)

Lei et al. (2017) suggest local global combined networks (LGCNet), a novel single image super resolution technique for remote sensing pictures based on deep CNNs. With its complex "multifork" architecture, LGCNet is able to train multilevel representations of remote sensing images that incorporate both global environmental priors and local features. Experimental results over multiple state-of-the-art algorithms show an overall improvement in both the accuracy and the visual performance on a public remote sensing data set (UC Merced) (**He et al., 2016; Kim et al., 2016a; Xu et al., 2016**).

Multilevel data has a lot of potential for jobs involving image super-resolution, particularly when it comes

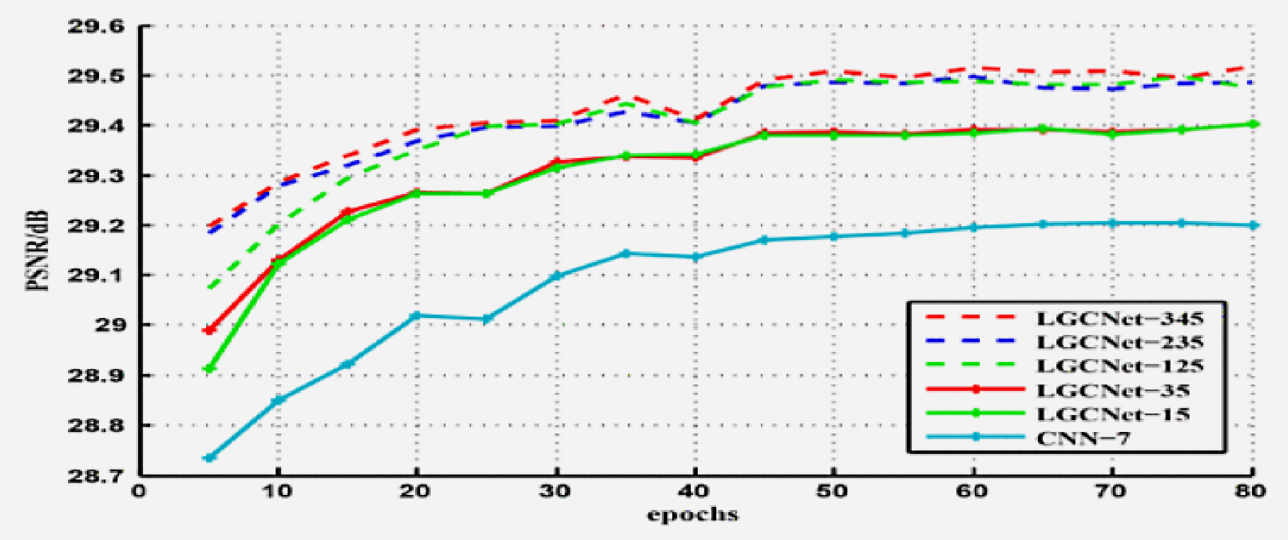

Figure 3. Results of the experiment of (mean PSNR) for validation set using various training epochs. Every model is trained using the same training configuration and an upscale factor of 3 (from **Lei et al., 2017**)

to photos from remote sensing. Deep CNNs, which have many convolution layers, are hierarchical models that by nature, produce multilevel representations of the input images. At lower layers, the representations concentrate on local details, such as an object's edge and contours. In comparison, at higher layers, they incorporate a greater degree of global prior information, such as the type of environment, see **Figure 2**. LGCNet consists of three primary components that are thoroughly explained as follows: it utilizes both the local and global representations to the fullest extent possible (**Lei et al., 2017**).

1- Representation: In the first section, inputs are adaptively transformed into effective feature space using L convolutional layers, with each layer being followed by a nonlinear mapping, to provide various level representations. They selected the filter size k_l and the number of the feature mappings n_l in each layer relatively small: k_l=3 and n_l=32, respectively, because of a big convolutional filter size that would make the network be redundant and slow.

2- Local Global Combination: The foundation of multiscale learning in this section. The primary method of doing local-global combination is by concatenating the convolutional results of various layers using a "multi-fork" structure. For the final reconstruction, these aggregated representations are merged using a single convolutional layer. By making the filter size and the number of feature maps relatively large, where k=5 and n=64, a richer representation of the combined layer can be obtained. The concatenated representation fc can be defined as follows:

$$fc=[fi,fj,fk,\ldots] \tag{3}$$

where:

fi , fj , fk are representations of different levels.

Then flgc the overall local global combined representation can be determined as follows:

$$flgc=\sigma(Wlgc*fc+blgc) \tag{4}$$

3- Reconstruction: The last section of LGCNet directly uses one convolutional layer to recover residuals, or high-frequency components, from the previously described local-global combined representation.

$$R=Wf*flgc+bf \tag{5}$$

By adding its low-resolution component, the ultimate high-resolution image Y^ can be obtained further.

$$Y\hat{}=X+R \tag{6}$$

Setting L=5 to 5 in LGCNet facilitates a quick inquiry and verification of the suggested concept. Padding is used for each convolutional layer with a size of 1 for k=3 and 2 for k=3 to ensure that the output feature maps match the inputs in size.

Figure 3 shows experiment results which are measured by the mean PSNR of the validation set with the training epochs proceeding. The models with related names are those that were created using various methodologies. Taking LGCNet-345 as an example, this model incorporates the third, fourth-, and fifth-layer representations. The combination of Layer, which incorporates more local and global representations, as predicted produces superior super-resolution results for the remote sensing images with more of layers combined.

The experimental findings demonstrate that combining several layers yields more precise reconstruction outcomes. Compared to many cutting-edge algorithms, our technique achieves overall increases in accuracy and visual performance (for all 21 classes). Additionally, re-

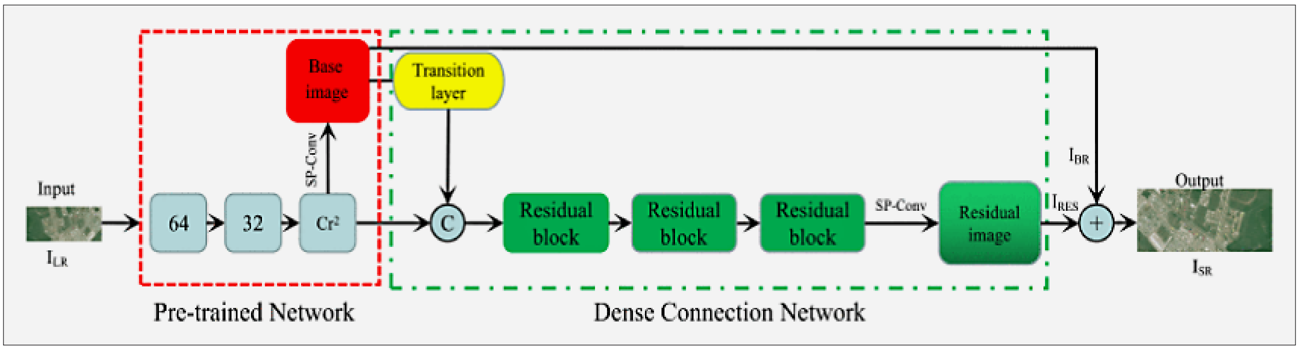

Figure 4. A summary of the suggested gradually upgraded network (PECNN). The pretrained portion is indicated by the components in the red box, and the dense connection subnetwork is indicated by the components in the green box (from **Jiang et al., 2018**).

al-world data studies validate the stability of our LGC-Net, and the adoption of additional layers in the representation portion results in slower quality gains.

3.3.1.3. Progressively Enhanced Convolutional Neural Network (PECNN) (**Jiang et al., 2018**)

The Progressively Enhanced Convolutional Neural Network (PECNN), a unique deep learning technique for super-resolving video satellite images, is presented in another previous study. The complicated imaging conditions of satellite photos, which frequently contain small ground targets, poor texturing, and severe compression distortion, provide issues that the network is built to handle (**Dong et al., 2015; Haeusser et al., 2017; Kim et al., 2016a; Tong et al., 2017b**). The low-resolution image (ILR) and super-resolution image (ISR) are viewed as the suggested PECNN's input and output, respectively, in this study. The term "up-scaling ratio" (r) is used. The C color channels of low- and high-resolution image are depicted as real-valued tensors with sizes of WHH×C and Wr×Hr×C, respectively.

A pre-trained CNN-based network and a dense connection network are the two subnetworks that make up the PECNN. From the low-resolution (LR) input image, the pre-trained network extracts low-level features to create a basic high-resolution (HR) image that is called the base image. In order to improve fine details, the dense connection network extracts high-frequency residuals from the base image. This network employs a transition unit to gather structural data, which is essential for enhancing the output of super-resolution (SR).

A subpixel convolution layer is added to the pre-trained network after a basic three-layer CNN to enable upscaling of the picture without the need for conventional interpolation techniques. An initial SR image is produced by this network and is used as the dense connection network's input. A more thorough representation can be obtained due to the intricate patterns of the several feature maps in the final filters. The following is a description of this operation:

$$I_{BR}=PS(f_n(I_{LR}))=PS(W_n\times f_n-1(I_{LR})+b_n) \quad (7)$$

Where:

f_n-1(ILR) denotes output of the $(n-1)^{th}$ layer, and

W_n and b_n represent the weight and bias of the nth layer, respectively.

f(·) refers to the convolution operation, followed by the rectified linear unit for activation.

PS(·) represents a periodic shuffling operator (SP-conv) (**Shi et al., 2016**) that can rearrange the elements of a $H\times W\times C\cdot r^2$ tensor into a renewed tensor of rH×rW×C.

To obtain the base image by the factor of r, it has been mathematically optimized the loss function as follows:

$$\theta\hat{}(I_{BR},\theta_1)=\arg\min_{\theta 1}\sum\|I_{HR}-f(I_{LR},\theta_1)\|^2_2 \quad (8)$$

The ground truth picture is denoted by IHR, the pre-trained network's base output is indicated by f(ILR,θ1), and the model parameters are represented by θ1. Specifically, after the pretrained network was optimized, the model parameters were fixed and maintained.

A transition unit in the dense connection network performs pixel offset, projection, and concatenation on the basis picture to produce structure-related information. In order to improve information propagation and provide improved feature extraction and expression, the network makes use of dense connections between residual blocks.

By combining the residual image generated by this network with the base image, the final output is created. The issue that lessens the impact on the front layers arises as depth grows. In order to extract and express features, a transition unit and a densely connected network have been used, as **Figure 5** illustrates.

The transition unit is made up of operations such as projection, pixel offset, and concatenation, as seen in **Figure 4** (top panel). The pixel offset is used to extract a set of images with comparable structures from the underlying image. After that, we project them from the HR space to the LR space in order to produce structure-related information. This involves down sampling the base images using a bicubic kernel with a down sampling value of r and applying a Gaussian filter to each basic

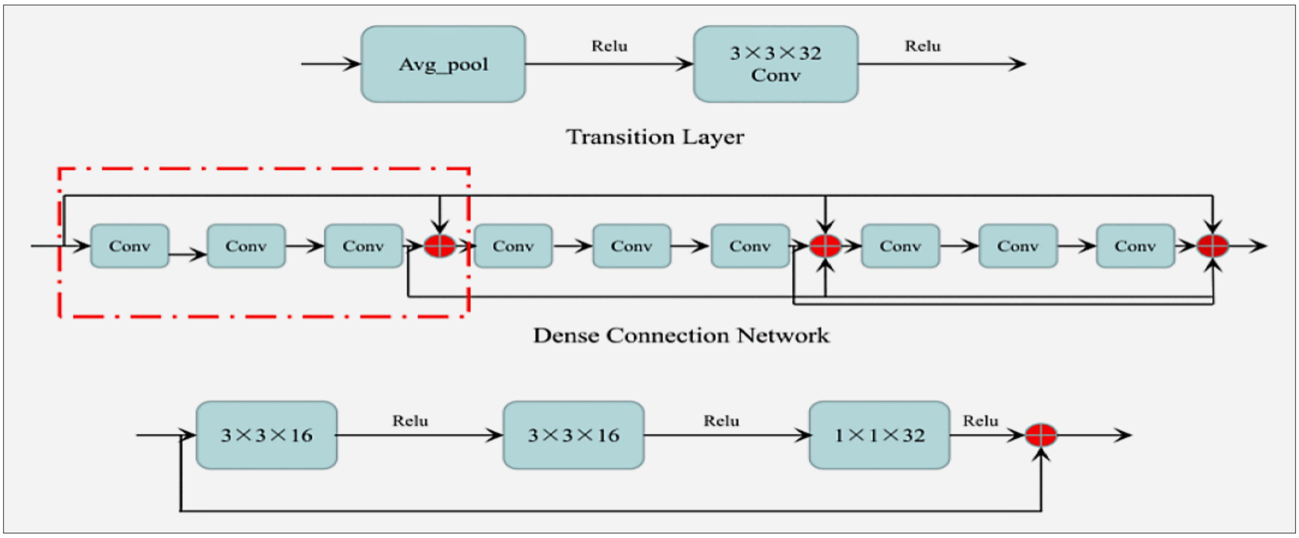

Figure 5. Overviews of ResNet and the suggested transition unit. Concatenation and addition are indicated by the symbols "C" and "+," respectively (from (**Jiang et al., 2018**)).

image in order to extract structure information. Ultimately, the transition unit's structure-related data and the pretrained network's low-level characteristics are combined for a final extraction in residual blocks (**Jiang, Wang, Yi, & Jiang, 2018**).

Second, the primary elements of the dense connection network are depicted in **Figure 4** (middle panel). In contrast to traditional skip-connection-based networks, a more beneficial dense connection pattern has been used to improve information propagation between every two blocks, hence improving the fine feature expression.

Lastly, a residual block is depicted in **Figure 5** (bottom panel), which extracts local features using three convolution layers. In order to minimize the parameters and combine the data, a 1x1 convolution layer is incorporated into the conclusion.

Image SR is a naturally occurring ill-posed problem with a non-unique solution. The network model's convergence rate and approximation accuracy during training are significantly impacted by the loss function. The commonly used loss functions include pixel-based l1-norm (**Lai et al., 2017**) and l2-norm (**Dong et al., 2015; Kim et al., 2016b, 2016a**), and feature-based cosine distance. Typically, they match the target results exactly with the learned visuals.

This study proposes a new progressively upgraded network for super resolving video satellite imagery, called PECNN. It is made up of a dense link network and a pretrained network. Specifically, an efficient transition unit is incorporated into the network's core to capture information pertaining to the profile structure and also encourages the expression of features through the use of progressive feature learning and more efficient dense connections. Large-scale satellite imagery can benefit from the built network's strong SR performance within tolerable computing complexity because it requires fewer depths and filters but denser connections between layers. The Jilin-1 video satellite imagery and the Kaggle Open Source Dataset experimental results demonstrate their superiority over the SRCNN, VDSR, and Bicubic (**Jiang et al., 2018**).

To compare the suggested ESRGAN with the original SRGAN and other cutting-edge techniques, a number of in-depth studies have been carried out. The DIV2K, Flickr2K, and OST datasets, which offered a wide range of textures, were used to train their models.

- The results of the investigation demonstrated that BN layer removal enhanced generalization and decreased computational complexity.
- Textures with greater detail and sharper edges are a result of the Relativistic Discriminator.
- Accurate brightness and crisper edges were caused by features present prior to activation in the perceptual loss.
- Texture recovery was further improved, and noise was decreased by combining the RRDB architecture with a deeper network.
- ESRGAN continuously outperforms earlier techniques in terms of quantitative measures and visual quality:
- Perceptual Quality: Based on comparisons on benchmark datasets such as Set5, Set14, BSD100, and Urban100, ESRGAN produced more natural textures with fewer artefacts than SRGAN and other approaches.
- PIRM-SR Challenge: The ESRGAN variant demonstrated its superior ability to produce high-quality super-resolution images by winning the first place in PIRM-SR Challenge (region 3) with the best perceptual index.

The following are the ramifications of the ESRGAN enhancements discussed in the paper:

Visual Clarity Against PSNR: A key topic is the trade-off between perceptual quality and PSNR, which may be

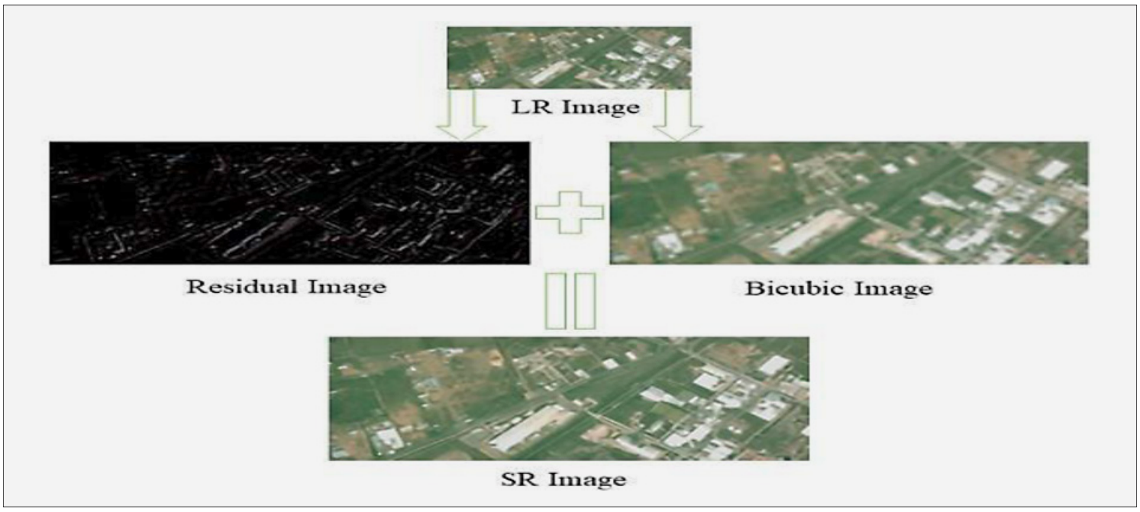

Figure 6. Graphical summary (from **Jiang et al., 2018**).

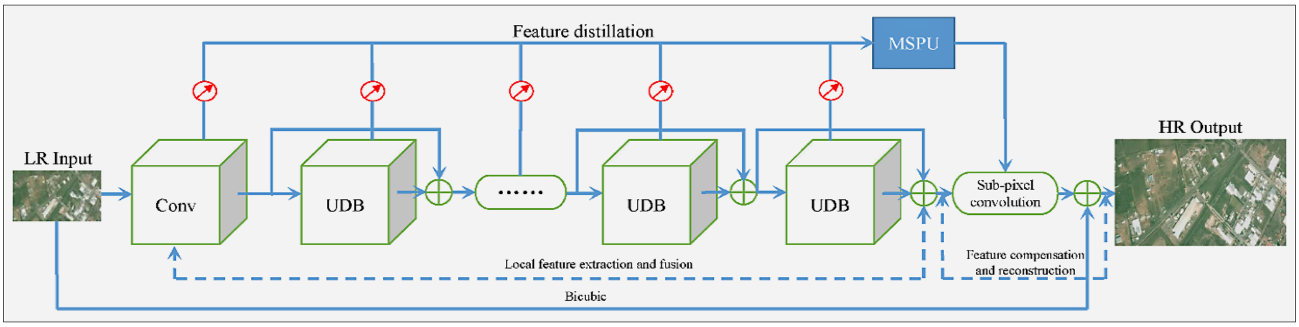

Figure 7. A summary of the deep distillation recursive network (DDRN) that is being suggested. The distillation process with a unique distilled ratio of α is represented by the red distillation symbol that comes after the UDB (from (**Jiang et al., 2018**)

fine-tuned based on the application using the network interpolation technique.

3.3.1.4. Deep Distillation Recursive Network (DDRN) (**Jiang et al., 2018**)

This paper presents the deep distillation recursive network (DDRN), a straightforward yet powerful CNN framework for video satellite image SR. A collection of ultra-dense residual blocks (UDB), a reconstruction module, and a multiscale purifying unit (MSPU) comprise the DDRN. In particular, it is possible to efficiently communicate features produced from several parallel convolution layers by adding rich interactive linkages within and between multiple-path units in each UDB. When compared to traditional models based on dense connections, DDRN has the following primary characteristics.

(1) More linking nodes with the same convolution layers can be found in DDRN.

(2) Additionally, a method for feature distillation and compensation is created that operates at various network levels. Specifically, MSPU allows for the compensation of high-frequency components lost during the information propagation.

(3) The feature maps taken from UDB and the corrected components acquired from MSPU can help the final SR image.

DDRN performs better than various state-of-the-art feature extraction techniques as well as the traditional CNN-based baselines in experiments conducted using the Kaggle Open-Source Dataset and Jilin-1 video satellite images.

The suggested model, depicted in **Figure 6**, is a deep recursive neural network that may be broadly classified into three substructures: feature distillation, feature extraction and fusion, and feature compensation and SR reconstruction. Except for the upsampling process, which is inspired by earlier SISR research (**Dong et al., 2015; Kim et al., 2016b; Shi et al., 2016; Tong et al., 2017a**), the entire process of the local feature extraction and fusion is in the LR space. ILR and ISR are considered the LR input and HR output of the proposed DDRN. Fi and Bj refer to output in the i^{th} layer and the j^{th} block, respectively. In this study, the LR RGB images are directly fed into the network and processed with initial convolutional layers (two layers with 3 × 3 kernel) to extract features as follows:

$$F_1=H(I_{LR}), \tag{9}$$

$$F_2 = H(F_1), \quad (10)$$

Where:

$H(\cdot)$ denotes convolution operation.

F_1 and F_2 represent the shallow feature maps extracted through the initial convolutional layers, served as the input of the UDB.

Additionally, the suggested residual block UDB serves as the foundational module in DDRN for the local feature extraction. Not only may the data for each UDB be exchanged between layers and the multiple-path units, but it can also be used as an input for the ultra-densely connected residual blocks that follow. Through the sequential combination of multiscale coarse-and-fine features, these strategies enforce the information transmission and result in fine feature expression.

As illustrated in **Figure 7** and **Figure 8**, our suggested approach may adaptively distil and preserve feature maps by partially selecting information from current output while maintaining its integrity. This is in contrast to the traditional network, whose output in each of the blocks is directly sent to the following segment. Prior to the reconstruction process, these feature maps that were gathered from various phases are then combined and refined in MSPU in order to deduce and account for the high-frequency information.

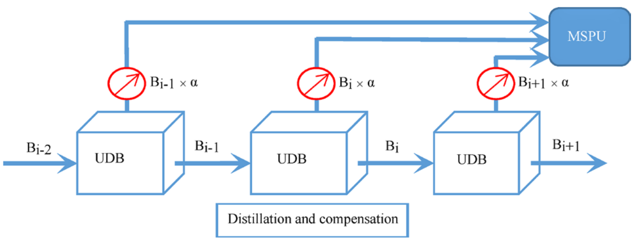

Figure 8. The process of distillation and compensation. The distilled feature maps $B_i \times \alpha$ in the current UDB are adaptively conserved, as indicated by the red components. The distillation ratio for the current UDB output, $B_i B_i$, is shown by αα. The further purification operation is referred to as MSPU (from **Jiang et al., 2018**).

3.3.2. GAN-based methods:

3.3.2.1. Super-Resolution Generative Adversarial Networks (SRGAN) (Ledig et al., 2017)

In order to produce photo-realistic high-resolution images and enhance the quality of single-image super-resolution (SISR), the research presents a novel technique called the Super-Resolution Generative Adversarial Network (SRGAN). The process includes a few essential elements:

The generator network is composed of 16 residual blocks and is based on a deep residual network (ResNet) with skip connections. Two convolutional layers with batch normalization and parametric ReLU as the activation function are present in each block. The generator uses sub-pixel convolution layers to boost resolution. The purpose of the discriminator network is to distinguish between actual high-resolution images and the super-resolved images that the network produces.

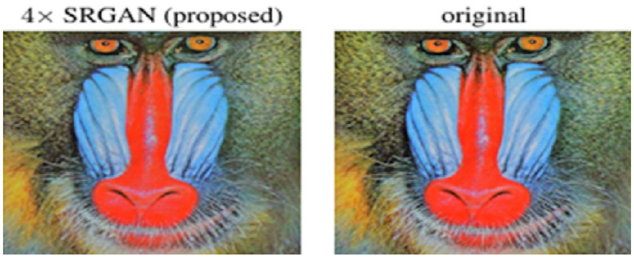

Figure 9. Comparing super-resolved images with SRResNet, bicubic interpolation, and SRGAN demonstrates how SRGAN improves perception (from **Ledig et al., 2017**)

Using the feature maps of the VGG19 network, a perceptual loss is used to calculate the content loss (see **Figure 10**). Instead of pixel-by-pixel similarity, this loss is intended to be more akin to perceived similarity. By using the adversarial loss, the generator is pushed towards the natural image manifold and encouraged to produce images that are indistinguishable from real images (see **Figure 9**). The adversarial loss and content loss are added together to form the combined loss function, also known as perceptual loss.

A sizable ImageNet image dataset was used to train the networks. Using a bicubic kernel, the high-resolution images were down scaled to produce the low-resolution versions. In-depth tests are conducted in the research using scale factors of 4×, which means that the photos were upscaled by four times their original size using benchmark datasets (Set5, Set14, and BSD100).

Quantitative Metrics: The Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR) are two metrics used in the paper to evaluate the performance of SRGAN with alternative approaches. While not attaining the maximum PSNR, SRGAN considerably enhances the perceived quality of images.

MOS, or mean opinion score: human raters evaluated the visual quality of the photographs in a subjective assessment. In comparison to alternative approaches, SRGAN achieved better MOS ratings, indicating superior perceptual quality as shown in **Figure 11**.

SRGAN produced high-quality, photo-realistic images better than state-of-the-art techniques, especially for large upscaling factors (4×). According to the MOS tests, images generated using SRGAN were judged to be more similar to the original high-resolution photographs than images created using other techniques. Additionally, the study discovered that improving texture detail and perceptual quality required careful consideration of the content loss option (VGG loss from deeper layers).

The authors talk about the shortcomings of conventional measurements like PSNR and SSIM, which frequently fall short of capturing an image's perceived quality. This gap is filled by SRGAN, which emphases

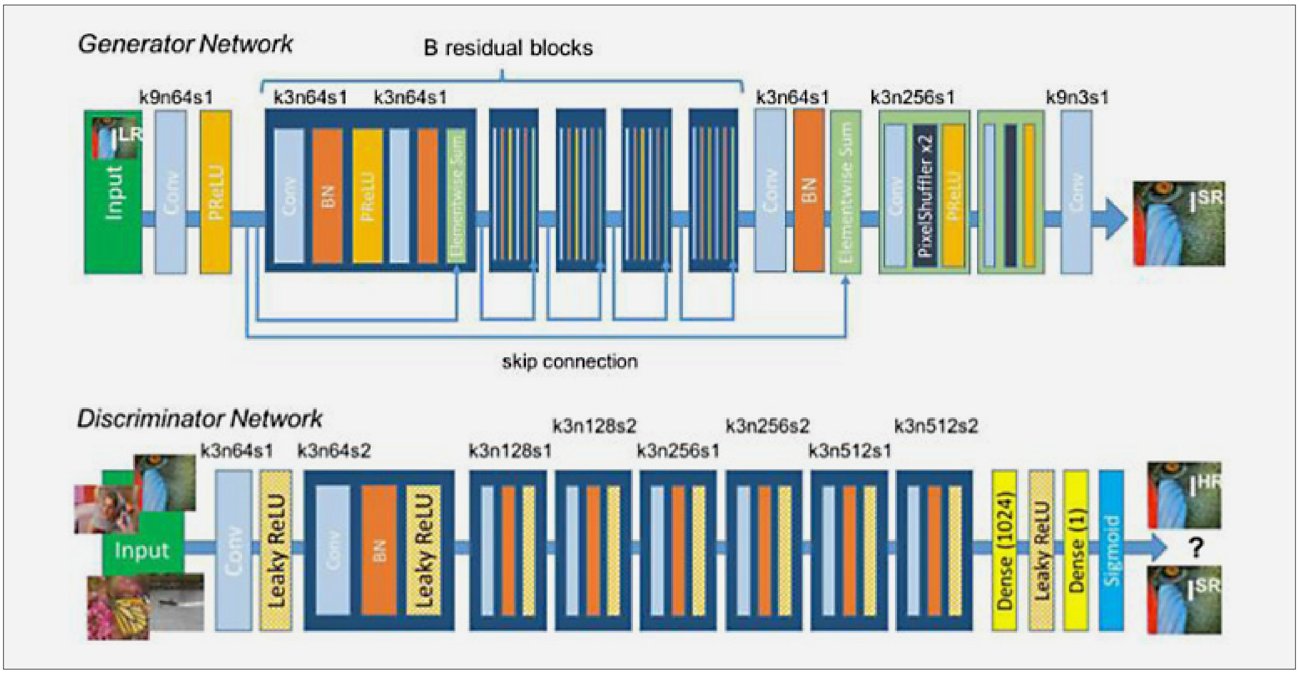

Figure 10. Structure of the SRGAN Generator and Discriminator Networks (from **Ledig et al., 2017**)

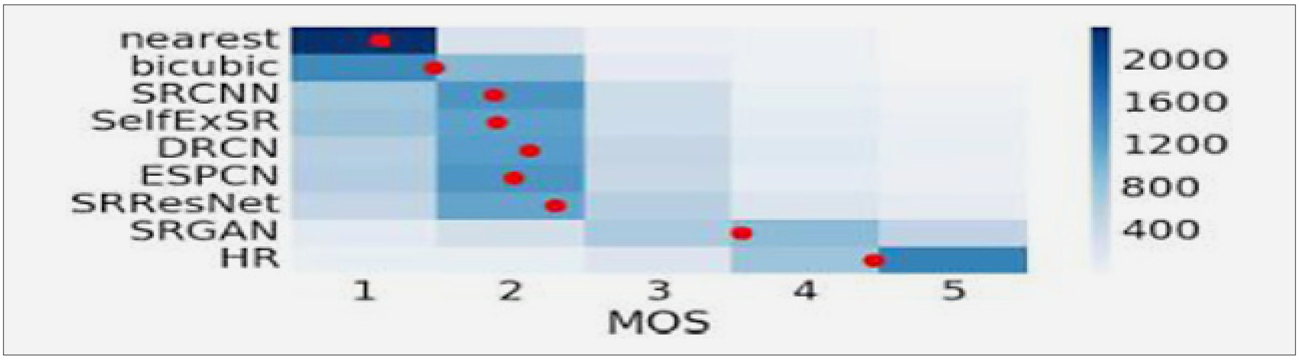

Figure 11. Distribution of MOS scores on the BSD100 dataset across approaches, demonstrating the superiority of SRGAN (from **Ledig et al., 2017**).

perceptual loss and yields more aesthetically acceptable outcomes. The study recommends that in order to further enhance the harmony between fidelity and perceptual quality, future research should investigate various network topologies and loss functions.

3.3.2.2. Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) (**Wang et al., 2018**)

The authors want to improve the Super-Resolution Generative Adversarial Network (SRGAN) in order to improve the visual quality of super-resolution for a single image. They emphasize three essential elements:

The primary building unit they introduce is the Residual-in-Residual Dense Block (RRDB), which increases the model's capacity without requiring the use of Batch Normalization (BN). By combining dense connections and residual networks, the RRDB improves performance in recovering high-resolution information and enables a deeper design.

Relativistic Average GAN (RaGAN), which the authors use, is a discriminator that forecasts relative realness between pictures as opposed to absolute realness. With this change, the generator is able to recover textures that are more realistic.

By utilizing attributes prior to activation, they enhance perceptual loss by offering more robust supervision for texture recovery and brightness consistency. They present a network interpolation technique that seamlessly transitions between a PSNR-oriented network and a GAN-based network in order to strike a compromise between perceived quality and PSNR (Peak Signal-to-Noise Ratio). To compare the suggested ESRGAN with the original SRGAN and other cutting-edge techniques, the authors carried out a number of in-depth studies. The DIV2K, Flickr2K, and OST datasets, which offered a wide range of textures, were used to train their models.

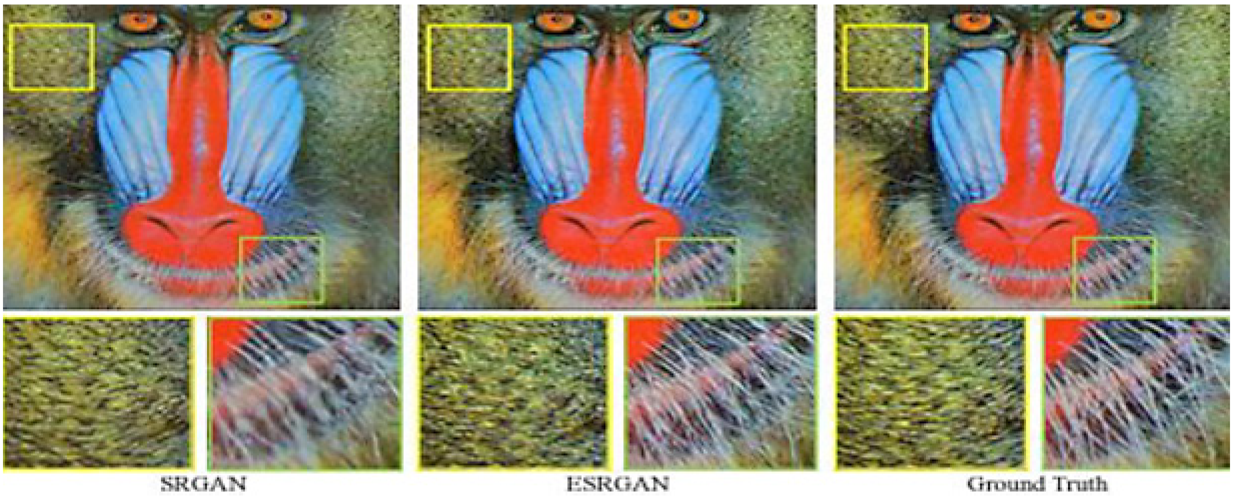

Figure 12. Super-resolution findings are compared between SRGAN, ESRGAN, and ground truth; ESRGAN displays textures that are sharper and more detailed (from: **Wang et al., 2018**)

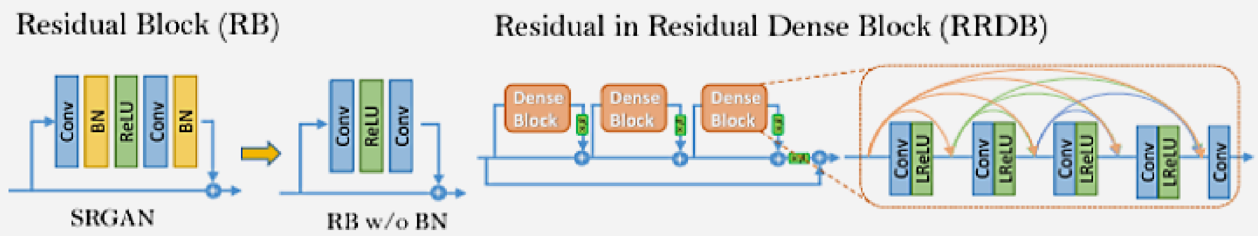

Figure 13. Diagram showing how the RRDB architecture differs from the original SRGAN (**Wang et al., 2018**)

The results of the investigation demonstrated that BN layer removal enhanced generalization and decreased computational complexity. Relativistic discriminator: a tool for producing textures with higher detail and sharper edges. More precise brightness and sharper edges were produced by features present before to activation in the perceptual loss. A deeper network and the RRDB architecture improved texture recovery and decreased noise even further, as shown in **Figure 12** and **Figure 13**.

In terms of both quantitative metrics and visual quality, ESRGAN continuously performs better than previous methods. In comparison to SRGAN and other techniques, it produces more realistic textures with fewer artefacts on benchmark datasets including Set5, Set14, BSD100, and Urban100. When ESRGAN won first place in the PIRM-SR Challenge (region 3) with the best perceptual index, demonstrating its remarkable capacity to generate high-quality super-resolution images, this excellent performance was further confirmed.

The following are the ramifications of the ESRGAN enhancements discussed in the paper: Visual Clarity Against PSNR: A key topic is the trade-off between PSNR and perceived quality, which may be fine-tuned based on the application using the network interpolation technique. The authors propose that texture recovery-focused perceptual loss functions be investigated further, and that ESRGAN could be employed for various picture restoration applications, see **Figure 14**.

3.3.2.3. Edge-enhanced GAN (EEGAN) (**Jiang et al., 2019**)

In this work, an adversarial learning method that is immune to noise was combined with a generative adversarial network (GAN)-based edge-enhancement network (EEGAN) for robust satellite image SR reconstruction. Specifically, EEGAN is composed of two primary subnetworks: the edge-enhancement subnetwork (EESN) and the ultra-dense subnetwork (UDSN). A collection of 2-D dense blocks is constructed in UDSN in order to extract features and provide an intermediate high-resolution result that appears sharp but is weakened by noise and artifacts, much as earlier GAN-based techniques. Then, using mask processing to remove the noise-contaminated components, EESN is built to extract and improve the image contours. A result with strong credibility and well-defined contents can be produced by combining the improved edges with the restored intermediate image. Extensive studies using Jilin-1 video satellite pictures, Digital globe, and Kaggle Open Source Data set demonstrate better reconstruction performance than the state-of-the-art SR techniques.

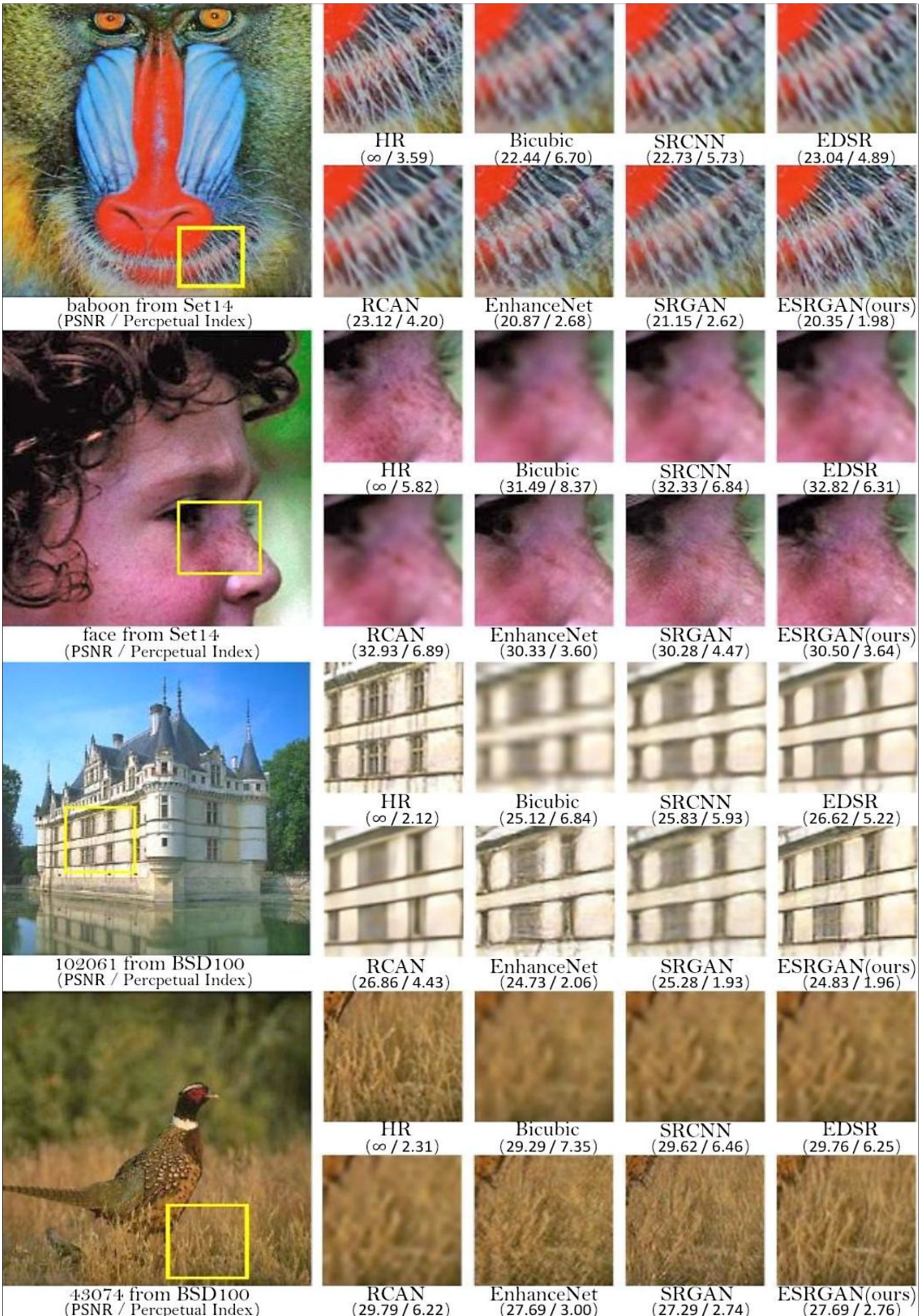

Figure 14. Comparative analyses using benchmark datasets that highlight ESRGAN's enhanced visual quality (from: **Wang et al., 2018**).

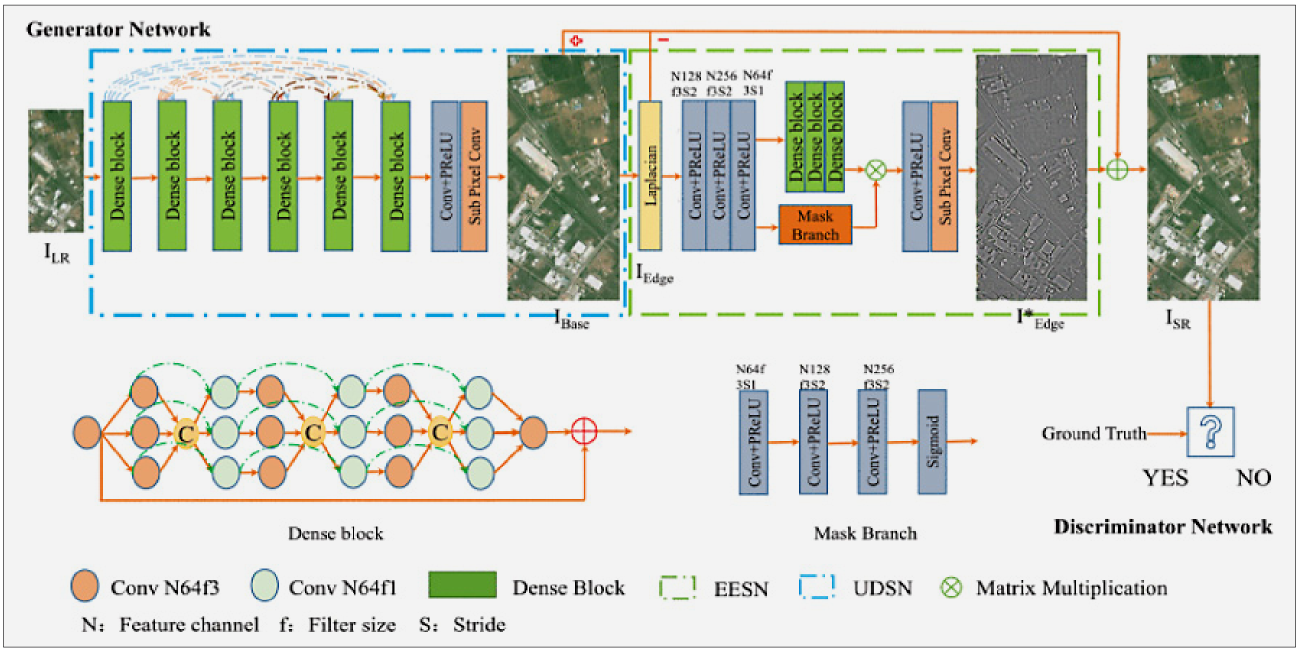

Figure 15. The suggested edge-enhancement network (EEGAN) in its schematic form (from: **Jiang et al., 2019**)

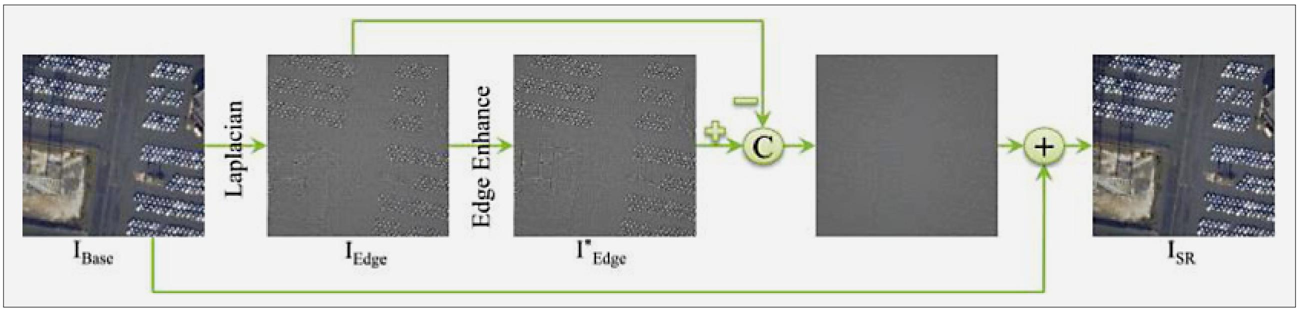

Figure 16. Description of the suggested edge improvement (from: **Jiang et al., 2019**)

The Ultra-Dense Subnetwork (UDSN) and the Edge-Enhancement Subnetwork (EESN) are the two subnetworks that make up the EEGAN structure as a whole, as seen in **Figure 15**. It demonstrates how the intermediary steps convert the low-resolution (LR) input into the final super-resolved (SR) output.

The EESN's edge extraction and enhancement procedure is shown in **Figure 16**. It describes how to extract the intermediate edge maps, clean them up with the mask branch, and create enhanced edge maps.

An Edge-Enhanced Generative Adversarial Network (EEGAN) for super resolution (SR) remote sensing image processing is proposed in this research. The technique is meant to deal with issues like recovering high-frequency edge details in imaging settings where noise is present. The two primary parts of the EEGAN are as follows:

UDSN, or ultra-dense subnetwork: The intermediate high-resolution (HR) image produced by this subnetwork, which is in charge of feature extraction, appears crisp but could include noise and artifacts.

Edge-Enhancement Subnetwork (EESN): Using a mask-processing technique, this subnetwork refines the edges derived from the intermediate HR picture. Next, the intermediate image and the improved edges are combined to create the final SR image.

The EEGAN framework uses a generative adversarial network (GAN) approach in which the discriminator distinguishes between the generated image and the actual HR image, and the generator attempts to create an HR image that is close to the ground truth. The Edge-Enhancement Subnetwork (EESN) architecture is the subject of **Figure 17**, which illustrates the progression of activities from edge extraction to noise reduction and enhancement. Several loss functions are involved in the process:

1. Content Loss: promotes the generation of an intermediate HR image that is close to the actual data.
2. Adversarial Loss: by training the discriminator to recognize phony images, this technique aids in the creation of realistic textures.
3. Consistency Loss: verifies that the final SR picture and the ground truth are consistent.

The comparison of EEGAN with cutting-edge SR techniques like SRCNN, VDSR, and SRGAN is the main emphasis of the analysis. Metrics such as the Feature Similarity Index (FSIM), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR) are used to assess the performance. In order to verify the contributions of the various EEGAN components, ablation studies are also included in the publication. The outcomes show that EEGAN performs noticeably better than competing SR techniques on a number of datasets, including Digital Globe imagery, Jilin-1 video satellite pictures, and Kaggle Open Source. Significant findings include the following:

EEGAN performs better at preserving clean, clear edges while lowering noise and artifacts. Quantitative measures and visual quality both significantly improve when the edge-enhancement approach successfully improves the image contours. Furthermore, the approach is robust under unknown degradation situations, which makes it extremely useful for real-world satellite imaging SR tasks.

The study emphasizes how well the EEGAN framework handles the problems caused by SR in remote sensing imagery. The capacity of the suggested procedure to produce visually appealing outcomes while preserving distinct and finely detailed edges is highlighted. The research also addresses the possibility of using the EEGAN framework for picture restoration applications other than remote sensing.

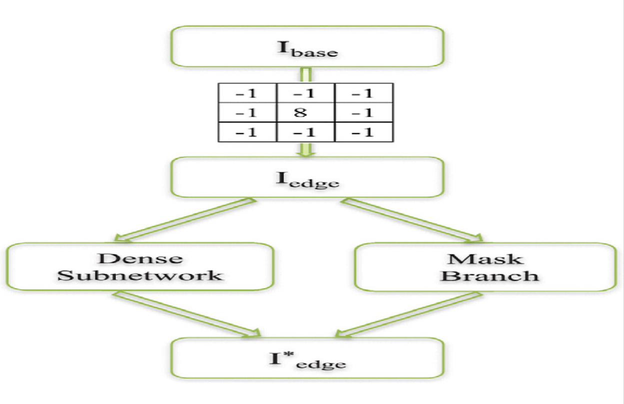

Figure 17. A description of the suggested EESN subnetwork (from: **Jiang et al., 2019**)

We have compiled the main characteristics, benefits, and drawbacks of the approaches covered in Sections 3.3.1 and 3.3.2 in **Table 2** to enable a clear comparison of them. In order to give readers a rapid and thorough grasp of each technique's suitability for remote sensing image enhancement tasks, this table attempts to highlight each technique's advantages and disadvantages. By combining this data, we hope to make it easier to choose the best approaches for particular use cases and promote more research into their possibilities in this area.

4. Integration of Deep Learning and GIS

Integrating a deep learning model for super-resolution image processing with GIS can significantly enhance satellite imagery by automatically improving the image quality (**Asif Raihan, 2023**). This integration can have numerous applications in fields like environmental monitoring, urban planning, agriculture, and more. Here's how such an integration might work.

The deep learning model used could be something like ESRGAN (Enhanced Super-Resolution Generative Adversarial Networks), SRCNN (Super-Resolution Convolutional Neural Network), or other advanced super-resolution techniques. The model would take lower-resolution satellite images as input and output high-resolution images by predicting finer details and enhancing the image quality (**Singla et al., 2022**).

The GIS platform would serve as the interface for inputting satellite images. Users can select regions of interest and specify the resolution they desire. Upon uploading or selecting a satellite image in the GIS platform, the deep learning model automatically processes the image in the background, enhancing its resolution. The enhanced image is then integrated back into the GIS environment, where it can be used for various spatial analyses. Higher resolution images provide more detailed data, improving the accuracy of analyses like land use classification, change detection, and feature extraction.

GIS is strongly reliant on these computational skills, and significant advances in computer power are rapidly opening up new opportunities, especially in the field of DL. Particularly in the areas of 3D modelling, map creation, and route calculation, the combination of GIS with DL has shown great promise as a useful tool (**Bogus-**

Table 2. Advantages and Disadvantages of CNN and GAN Models in Satellite Image Enhancement

Method	Advantages	Disadvantages
SRCNN	Simple architecture, computationally efficient	Limited ability to capture complex features
LGCNet	Captures local-global relationships effectively	Higher computational cost
PECNN	Gradual refinement improves fine details	Substantial training time required
DDRN	High reconstruction accuracy, efficient feature extraction	Complex architecture may increase the computational load
SRGAN	Generates realistic textures	Prone to artefacts at abrupt gradient changes
ESRGAN	Superior texture recovery and artifact reduction	High computational cost, complex training process
EEGAN	Preserves fine edge details	Sensitive to noise and requires significant preprocessing

lawski et al., 2022). The system's analytical capabilities are improved when remotely sensed data is combined with other geographical variables inside a GIS framework (**Bilotta et al., 2023**). In several phases of the Artificial Intelligence (AI) workflow, GIS is essential. As an example, GIS plays can be used to provide geographic coordinates to photos that are detected using DL algorithms. Furthermore, an algorithm can be used in GIS to recognize and examine photographs. For the storage, retrieval, analysis, and visualization of data, Artificial Neural Networks (ANNs) require a compatible environment. One possible remedy has been suggested: the combination of GIS and ANN (**Liao et al., 2023**). According to **Ferchichi et al. (2022)**, the integration of GIS and DL has great promise for a variety of applications, such as the classification of RS imagery and attribute data analysis. It may be possible to improve environmental mapping and the identification and retrieval of objects inside an integrated database by integrating a GIS database with sophisticated DL models (**Chen, 2022**). Since neural network classifiers are independent of previous statistical models for input data, they can operate as broad pattern recognition systems and show flexibility in integrating various data types. As a result, they can be used to incorporate GIS information into the classification of remote sensing photos (**Fırat et al., 2023**). There are numerous previously published studies that used GIS data in their investigations. The work by **Benediktsson et al. (1990)**, which used topographic data in conjunction with Landsat Multi-spectral Scanner (MSS) data for the goal of mapping ground cover, serves as an example of this. Landscape height information obtained from a digital terrain model (DTM) was used as an additional input for a multi-layer perceptron neural network in a different study carried out by the Joint Research Committee in 1991. Two sets of SPOT (a series of French high-resolution optical imaging Earth observation satellites) High-Resolution Visible (HRV) photos taken at various times were used to train this neural network. Classifying land cover in satellite photos was the project's goal. In this case, the overall accuracy of the categorization process was significantly improved by integrating DTM data into a GIS framework. Additionally, the neural network's training time was decreased significantly as a result of this integration, essentially half the duration. The application of GIS data to neural networks is a relatively new field with little actual research to support it. However, the minimal amount of data suggests that this strategy could end up being a profitable one (**Huang et al., 2022**). For real-time applications, the system could be set up to process images on-the-fly, allowing users to see the enhanced results almost immediately after image acquisition (**Sarker, 2021**).

Geospatial data and particularly remote sensing data are known in the study of urban areas for the purpose of gathering comprehensive data about and monitoring the changing aspects of different constituents of urban areas (**Ouchra et al., 2022**). This is important for urban planning, monitoring, and supporting decision-makers with insightful information for making informed decisions about the different facets and systems within urban areas (**Wu et al., 2020**). GIS is an essential tool in enabling the simulation and illustration of different aspects of urban life (**Sinjari & Kosovrasti, 2015**). They make use of the spatial data available, which are sets of data with varying attributes and characteristics, allowing for the exploration and analysis of the relationships between different aspects of urban life (**Kareem Jebur, 2021; Sinjari et al., 2015**). These systems are widely used due to their ability to integrate, analyze, map, and depict spatial data for the purpose of monitoring and guiding the development and change of urban areas over time (**Yu et al., 2023**).

5. Evaluation Metrics for Image Enhancement

The evaluation of the performance of a model is an important aspect of deep learning models. It refers to how well a model generates an output from an input as per the task defined. The evaluation includes measuring the performance of the model in a meaningful way relative to the data distribution, how to optimize the performance, and how it must measure the model's accuracy. The performance of any model first must be reviewed against the ground truth. Based on the application of the model developed and their architecture, these applications have pre-defined evaluation metrics, and they range from simple measures such as accuracy to more comprehensive measures such as recall and precision including Mean Square Error, Structural Similarity Index, Peak Signal-to-Noise Ratio, and their function for satellite image processing (**Mohammed et al., 2023**).

All the evaluation metrics like mean squared errors, structural similarity indices, peak signal-to-noise ratio, and corresponding functions are developed to signal overlap among different images or signal files. Currently, there are no such specialized or tailored evaluation metrics developed exclusively to evaluate the generated enhanced pair of images due to image enhancement or super-resolution tasks as there is no explicit ground truth image (**Sara et al., 2019**). As there is no universal metric developed to assess the quality of the generated images like denoisegans, super-resolving GANs, and so on, one has to make the final decision based on these general-purpose image similarity indices to select the model leading to this major challenge and lack of standard evaluation metric give rise to inconsistent results quoted in the literature leading to Web Object Creation (WOG) approach. In this study, a detailed summation of evaluation metrics used to appraise the image quality enhancement techniques using deep learning advanced or tailored specifically Graphical Information System operations tailored for satellite images are reviewed (**Sara et**

al., 2019; Wang et al., 2024; Wenlong et al., 2021). The derivation of these symbolic algebra is also provided. To the best of the author's knowledge, this is the first review of its kind in literature. The outcomes from this study help scientists select the appropriate model for enhancing a given type of satellite image **(Sara et al., 2019)**.

5.1. Peak Signal-to-Noise Ratio (PSNR)

The signal-to-noise ratio (SNR) or the peak signal-to-noise ratio (PSNR) is a widely used algorithm for computing the quality metric in an image processing system. The higher the value of SNR, the better the image quality. Document Image Binarization Contest (DIBCO) 2010 used it to measure the effectiveness of the proposed algorithm. It is often applied for denoising, transforming, or enhancing satellite images and functions (**Petrovic et al., 2016**). Its method is straightforward and common, which is the ratio of the peak value of a signal to the effective noise. The normalized form of the signal-to-noise ratio is expressed in decibels as a relatively comprehensible measure (**Noor Azam et al., 2022**). The conventional transformation is used to convert the ratio from amplitude to decibels. PSNR is a simple, yet effective, method that can be used. Various types of encoders deal with testing into different transmissions, lossy coding, decompressed video, and images with a marked SNR. It is often employed as an objective rating for graphic processing and Computer-Generated Imagery (CGI) production effects to assess graphic fidelity (**Horé et al., 2010**).

5.2. Structural Similarity Index (SSI)

Structural Similarity Index (SSI) is a widely used metric that considers the changes in structural information in the original and enhanced images. To do this, the SSI index is first computed and then used as a factor to enhance GS satellites' input images. SSI is calculated as the product of three functions: luminance, contrast, and structure (**Chebbi et al., 2014**). However, these three functions can be specifically tuned to contain abundant local structural information. Therefore, the SSI has the ability to weigh up structures not just in the whole domain, but also locally based on visually hidden human visual perception. Using human visual psychological factors to simulate the pixel processing process boosts the enhancement effects (**Peng et al., 2020; Renieblas et al., 2017**).

In general, three model assumptions are used to slim down the number of solutions. These assumptions are: (I) the pixel values in the dark channel H* have a Gaussian distribution; (II) the condition p(H(ti)|Ei, A (ti), T (ti), θi) is a Gaussian distribution; (III) the variance σ^2(H(ti)) concerning the A and T are independent of the width of dark channel edges in H. These assumptions allow the semidark channel estimation equation to derive the enhancement model E (ti) for its pixel values. The enhanced satellite input image is then reconstructed with selected estimation, rendering the enhancement color-consistent with human vision. In the experiments, well-known classic bio visual perception theories are used as golden standard visual models – to parameterize and learn unbiased filters to predict the low-frequency visual black-gray model for enhancement. Promising results based on the GS satellite input from the WorldView-3 dataset are reported (**Renieblas et al., 2017**).

5.3. Mean Squared Error (MSE)

Mean squared error (MSE) is the average of the squared errors or deviations over all pairs of true and predicted values in a dataset. Specifically, for the pair of true and predicted values from an image, the square of their difference is calculated (**Over et al., 2021**).

$$MSE = \frac{\sum (y_i - \hat{y_i})}{n} \tag{11}$$

Where:

y_i is the i[th] observed value.

$\hat{y}_i$ is the corresponding predicted value.

n = the number of observations.

Then, the mean of these different squared values (one for each pair of true and predicted pixels in the image) results in MSE. An MSE value closer to zero denotes a more accurate model. Particularly, zero MSE suggests a 100% accurate model. This error function penalizes a large error. Hence, larger MSE values point to increased disparities between true and predicted labels. In the context of deep learning, lower MSE values imply better regression models for a wide array of datasets as well as applications such as advanced computer vision (**Zhao et al., 2024**).

6. Future Directions and Emerging Trends and Applications

Due to the promising results of satellite image enhancement using the DL model, researchers have been setting their focus to explore other exciting and interesting avenues. These unique perspectives are likely to transform a nascent field into an exhilarating domain of in-depth research (**Zhang et al., 2022**). This article reviews the current critical pressing questions and discusses the direction of future efforts, including a wide variety of directions that could be pursued (**Wang et al., 2018**).

It would be interesting to know other innovative and compelling approaches for enhancing satellite imagery. The current approaches of performing pan-sharpening, i.e. super resolution for satellite images, are necessary (**Wang et al., 2018**), yet still need to be upgraded with cutting-edge techniques to improve satellite image enhancement.

In conclusion, DL could learn a sophisticated transformation from low-quality to high-quality satellite images while concurrently improving our geographical maps (**Liu et al., 2021; Wang et al., 2018**). This promising technique could bolster broader satellite image assessments, leading to results with more accurate images and thus vital improvements. Moreover, the outcome of the intended model should also be assessed with the aid of more semantic depth radiometric correction and true validation (**Zhang et al., 2022**). Additional annotations to train the model would augment pan-sharpening effectively, making it more accurate and suitable for wider applications (**Liu et al., 2021; Zhang et al., 2022**). All in all, combining DL integrated with GIS represents an exciting opportunity for the development and advancement of high-resolution remote sensing in more in-depth and substantial fashions to benefit various fields. With the progression of this technology, beneficial applications for urban remote sensing will drastically widen, significantly improving everyday life (**Abdalla, 2024**). Deep learning has the potential to strengthen GIS-related applications, including real-time mapping, historical map validation, image resolution enhancement, imagery analysis, disease detection, traffic prediction, question-and-answer GIS platforms, object detection in large-scale maps, species detection and count, forecasting of municipal services, map navigation, and many other diverse real-world GIS problems (**Kiwelekar et al., 2020**).

Deep learning can enhance GIS science, handle online GIS data effectively, and give users useful information. Deep learning models can be taught and evaluated to extract significant patterns and insights from big datasets, including crowdsourced GIS data, satellite photography, and aerial imagery. These models use input data, such as geographic datasets and high-resolution imagery, and training data, which usually consists of labelled samples to teach the model particular tasks, such as item detection or land use categorization. The accuracy and generalization capacity of the model are then assessed using test data. Deep learning integration is crucial to realizing the national spatial data infrastructure's vision and helps build the spatial data infrastructures accessible via Google Earth, Wikimapia, Google Maps, Bing Maps, TomTom, OpenStreetMap, and other platforms (**Bill et al., 2022**). To further integrate satellite and aerial images with GIS and progress both domains at the same time, the GIS and deep learning communities should collaborate (**Williamson et al., 2007**). Working concurrently with deep learning, we can make great strides in advancing the state of GIS practice, thereby also enhancing the lives and decision-making skills of ordinary users (**Hosen et al., 2023**).

7. Conclusions

The survey showed that deep learning models are more reliable for enhancing specialized information than commonly used image enhancement tools in GIS for multi-temporal/sensor satellite images, except for very high resolution (VHR) images Particularly, due to the inability of the developed model to extract large context information, there were certain problems in the application to VHR images. Therefore, it is expected that studies will be conducted to create a more flexible structure to use higher content information and context, especially in VHR images. It is expected that the inclusion of higher-level abstraction information in the network structure will allow the model to take a more comprehensive approach to the context it is in and perform better.

In addition, the desired image quality parameters will be the priority. If there is no satisfactory realization of enhanced results, deep learning models will not be very preferred because of the high computational costs of these methods. Since the deep learning concept has come to the fore as a very popular and beautiful concept of the day and everyone follows and applies it in its field, studies to be conducted on this subject will appear with different methods, approaches, data, and results. Recent conclusions are expected to become a shortcut and guide for further deep-learning image enhancement applications. Another important finding of the work is the determination of the necessity for spectral index generation and enhancement based on these indices, independent of the selected transfer learning architecture. This flexibility allowed a generalist model to be developed, and different deep-based learning models tailored specifically to the object were compared. It is considered as a significant point as it is beneficial in terms of providing speed, consistency, and model complexity reduction based on the application of the developed model, and hence the very high computing cost and infrastructure required in the background.

The current review on using deep learning approaches in enhancing satellite image resolutions combined with GIS models showcased the efficiency of using deep learning in solving such spatial expertise. It discussed various combinations of deep learning methods to resolve the current inadequacy in obtaining GIS data and proposed a potential combination of deep learning models. As the advancements in studying the earth phenomenon are rapidly increasing, the usage of GIS data is incrementing. Alongside with that, the increase of data from satellite images from UAVs and from other sources may increase the demand for high-resolution GIS-based analysis, which includes a lot of ground truth information. The collaboration of deep learning with other GIS models has great potential in the optimization of GIS and the model scenarios.

To ensure the applications are more successful, studies are increasing their attention to the training methodology and problems that hinder the model's application. Publicly shared code and models can also improve the models' potential for application in real-world scenarios. Additionally, with an increase in the models' use, it may be carefully considered setting an ecological framework that may include, but not just focus on, increasing the prediction of data or the precision of data resolution quality with less truth value and examination. The potential use of deep learning in geospatial data is enormous, and hand in hand, GIS and deep learning models will reshape future geospatial data globally.

Acknowledgement

The authors would like to extend their sincere appreciation to Assiut University, Egypt.

Dalia A. Hussein extremely grateful to **Mohamed E. Ibrahim**; it would not have been possible without his support and nurturing; he helped provide expertise that she lacked.

8. References

Abdalla, R. (Ed.). (2024). *Geographic Information Systems*. IntechOpen. https://doi.org/10.5772/intechopen.111053

Asif Raihan. (2023). A Comprehensive Review of the Recent Advancement in Integrating Deep Learning with Geographic Information Systems. *Research Briefs on Information and Communication Technology Evolution* , *9*(SE-Articles), 98–115. https://doi.org/10.56801/rebicte.v9i.160

Belgiu, M., & Stein, A. (2019). Spatiotemporal Image Fusion in Remote Sensing. In *Remote Sensing* (Vol. 11, Issue 7). https://doi.org/10.3390/rs11070818

Benediktsson, J. A., Swain, P. H., & Ersoy, O. K. (1990). Neural network approaches versus statistical methods in classification of multisource remote sensing data. *Vancouver, Canada, July 10-14, 1989) IEEE Transactions on Geoscience and Remote Sensing*.

Bill, R., Blankenbach, J., Breunig, M., Haunert, J.-H., Heipke, C., Herle, S., Maas, H.-G., Mayer, H., Meng, L., & Rottensteiner, F. (2022). Geospatial information research: state of the art, case studies and future perspectives. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, *90*(4), 349–389.

Bilotta, G., Genovese, E., Citroni, R., Cotroneo, F., Meduri, G. M., & Barrile, V. (2023). Integration of an innovative atmospheric forecasting simulator and remote sensing data into a geographical information system in the frame of agriculture 4.0 concept. *AgriEngineering*, *5*(3), 1280–1301.

Boguslawski, P., Zlatanova, S., Gotlib, D., Wyszomirski, M., Gnat, M., & Grzempowski, P. (2022). 3D building interior modelling for navigation in emergency response applications. *International Journal of Applied Earth Observation and Geoinformation*, *114*, 103066.

Chatterjee, P., Joshi, N., Kang, S. B., & Matsushita, Y. (2011). *Noise suppression in lowlight images through joint denoising and demosaicing*. https://doi.org/10.1109/CVPR.2011.5995371

Chebbi, E., Faouzi, B., & Amiri, H. (2014). An improvement of structural similarity index for image quality assessment. *Journal of Computer Science*, *10*, 353–360. https://doi.org/10.3844/jcssp.2014.353.360

Chen, C. (2007). *Signal and image processing for remote sensing*. CRC/Taylor & Francis.

Chen, S., & Guo, W. (2023). Auto-encoders in deep learning – a review with new perspectives. *Mathematics*, *11*(8), 1777.

Chen, Z. (2022). *Automated Geoscience with Robotics and Machine Learning: A New Hammer of Rock Detection, Mapping, and Dynamics Analysis*. Arizona State University.

Dibs, H., Jaber, H. S., & Al-Ansari, N. (2023). Multi-Fusion algorithms for Detecting Land Surface Pattern Changes Using Multi-High Spatial Resolution Images and Remote Sensing Analysis. *Emerging Science Journal*, *7*(4), 1215–1231. https://doi.org/10.28991/ESJ-2023-07-04-013

Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(2), 295–307.

Dumoulin, V., Belghazi, I., Poole, B., Mastropietro, O., Lamb, A., Arjovsky, M., & Courville, A. (2016). Adversarially learned inference. *ArXiv Preprint ArXiv:1606.00704*.

Ferchichi, A., Abbes, A. Ben, Barra, V., & Farah, I. R. (2022). Forecasting vegetation indices from spatio-temporal remotely sensed data using deep learning-based approaches: A systematic literature review. *Ecological Informatics*, *68*, 101552.

Fisher, J. R. B., Acosta, E. A., Dennedy-Frank, P. J., Kroeger, T., & Boucher, T. M. (2018). Impact of satellite imagery spatial resolution on land use classification accuracy and modeled water quality. *Remote Sensing in Ecology and Conservation*, *4*(2), 137–149. https://doi.org/https://doi.org/10.1002/rse2.61

Fırat, H., Asker, M. E., Bayındır, M. İ., & Hanbay, D. (2023). Hybrid 3D/2D complete inception module and convolutional neural network for hyperspectral remote sensing image classification. *Neural Processing Letters*, *55*(2), 1087–1130.

Fu, X., Wang, J., Zeng, D., Huang, Y., & Ding, X. (2015). Remote Sensing Image Enhancement Using Regularized-Histogram Equalization and DCT. *IEEE Geoscience and Remote Sensing Letters*, *12*(11), 2301–2305. https://doi.org/10.1109/LGRS.2015.2473164

Gao, L., Hong, D., Yao, J., Zhang, B., Gamba, P., & Chanussot, J. (2020). Spectral superresolution of multispectral imagery with joint sparse and low-rank learning. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(3), 2269–2280.

Goodfellow, I. (2016). *Deep learning*. MIT press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139–144.

Haeusser, P., Mordvintsev, A., & Cremers, D. (2017). Learning by association--A versatile semi-supervised training method for neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 89–98.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507.

Hong, S., Shin, D.-B., Park, B., & So, D. (2016). Development of prototype algorithms for quantitative precipitation nowcasts from AMI onboard the GEO-KOMPSAT-2A satellite. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(12), 7149–7156.

Horé, A., & Ziou, D. (2010). Image Quality Metrics: PSNR vs. SSIM. *2010 20th International Conference on Pattern Recognition*, 2366–2369. https://doi.org/10.1109/ICPR.2010.579

Hosen, B., Rahaman, M., Kumar, S., Sagar, L., & Akhtar, M. N. (2023). Leveraging Artificial Intelligence and Big Data for Advanced Spatial Analytics and Decision Support Systems in Geography. *Malaysian Applied Geography*, *1*(2), 62–67. https://doi.org/10.26480/magg.02.2023.62.67

Huang, H., Huang, J., Feng, Q., Liu, J., Li, X., Wang, X., & Niu, Q. (2022). Developing a dual-stream deep-learning neural network model for improving county-level winter wheat yield estimates in China. *Remote Sensing*, *14*(20), 5280.

Jensen, J. R. (2009). *Remote sensing of the environment: An earth resource perspective 2/e*. Pearson Education India.

Jiang, K., Wang, Z., Yi, P., & Jiang, J. (2018). A progressively enhanced network for video satellite imagery superresolution. *IEEE Signal Processing Letters*, *25*(11), 1630–1634.

Jiang, K., Wang, Z., Yi, P., Jiang, J., Xiao, J., & Yao, Y. (2018). Deep distillation recursive network for remote sensing imagery super-resolution. *Remote Sensing*, *10*(11), 1700.

Jiang, K., Wang, Z., Yi, P., Wang, G., Lu, T., & Jiang, J. (2019). Edge-enhanced GAN for remote sensing image superresolution. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(8), 5799–5812.

Jing, C.-W., Huang, Z.-X., & Ling, Z.-Y. (2022). An image super-resolution reconstruction method based on PEGAN. *IEEE Access*, *11*, 102550–102561.

Kareem Jebur, Ahmed. (2021). Uses and Applications of Geographic Information Systems. *Saudi Journal of Civil Engineering*, *5*(2), 18–25. https://doi.org/10.36348/sjce.2021.v05i02.001

Karras, T. (2019). A Style-Based Generator Architecture for Generative Adversarial Networks. *ArXiv Preprint ArXiv:1812.04948*.

Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *173*, 24–49. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2020.12.010

Kim, J., Lee, J. K., & Lee, K. M. (2016a). Accurate image super-resolution using very deep convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1646–1654.

Kim, J., Lee, J. K., & Lee, K. M. (2016b). Deeply-recursive convolutional network for image super-resolution. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1637–1645.

Kingma, D. P. (2013). Auto-encoding variational bayes. *ArXiv Preprint ArXiv:1312.6114*.

Kiwelekar, A. W., Mahamunkar, G. S., Netak, L. D., & Nikam, V. B. (2020). Deep learning techniques for geospatial data analysis. *Machine Learning Paradigms: Advances in Deep Learning-Based Technological Applications*, 63–81.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*.

Lai, W.-S., Huang, J.-B., Ahuja, N., & Yang, M.-H. (2017). Deep laplacian pyramid networks for fast and accurate super-resolution. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 624–632.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A. P., Tejani, A., Totz, J., Wang, Z., & others. (2017). \href{https://ieeexplore.ieee.org/abstract/document/8099502}{Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network}. *Cvpr*, *2*(3), 4. http://openaccess.thecvf.com/content_cvpr_2017/papers/Ledig_Photo-Realistic_Single_Image_CVPR_2017_paper.pdf

Lee, J.-S. (1980). Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *2*, 165–168.

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.

Liao, H., He, Y., Wu, X., Wu, Z., & Bausys, R. (2023). Reimagining multi-criterion decision making by data-driven methods based on machine learning: A literature review. *Information Fusion*, 101970.

Liu, R., Yang, X., Xu, C., Li, L., & Zeng, X. (2021). *Landslide susceptibility mapping based on convolutional neural network and conventional machine learning methods*.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2019.04.015

Masi, G., Cozzolino, D., Verdoliva, L., & Scarpa, G. (2016). Pansharpening by Convolutional Neural Networks. In *Remote Sensing* (Vol. 8, Issue 7). https://doi.org/10.3390/rs8070594

Mohammed, A., & Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University - Computer and Information Sciences*, *35*(2), 757–774. https://doi.org/https://doi.org/10.1016/j.jksuci.2023.01.014

Noor Azam, M. H., Ridzuan, F., & Sayuti, N. (2022). A New Method to Estimate Peak Signal to Noise Ratio for Least Significant Bit Modification Audio Steganography. *Pertanika Journal of Science and Technology*, *30*, 497–511. https://doi.org/10.47836/pjst.30.1.27

Noshiri, N., Beck, M. A., Bidinosti, C. P., & Henry, C. J. (2023). A comprehensive review of 3D convolutional neural network-based classification techniques of diseased and defective crops using non-UAV-based hyperspectral images. *Smart Agricultural Technology*, *5*, 100316. https://doi.org/https://doi.org/10.1016/j.atech.2023.100316

Ouchra, H., Belangour, A., & Erraissi, A. (2022). A comprehensive study of using remote sensing and geographical information systems for urban planning. *Internetworking Indonesia Journal*, *14*, 15.

Oveis, A. H., Giusti, E., Ghio, S., & Martorella, M. (2021). A survey on the applications of convolutional neural networks for synthetic aperture radar: Recent advances. *IEEE Aerospace and Electronic Systems Magazine*, *37*(5), 18–42.

Over, T., & Foks, S. (2021). Mean Squared Error, Deconstructed. *Journal of Advances in Modeling Earth Systems*, *13*. https://doi.org/10.1029/2021MS002681

Peng, J., Shi, C., Laugeman, E., Hu, W., Zhang, Z., Mutic, S., & Cai, B. (2020). Implementation of the structural SIMilarity (SSIM) index as a quantitative evaluation tool for dose distribution error detection. *Medical Physics*, *47*. https://doi.org/10.1002/mp.14010

Petrovic, V., Pavlović, B., Andrić, M., & Bondzulic, B. (2016). Performance of peak signal-to-noise ratio quality assessment in video streaming with packet losses. *Electronics Letters*, *52*. https://doi.org/10.1049/el.2015.3784

Pohl, C., & Van Genderen, J. L. (1998). Review article multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, *19*(5), 823–854.

Radford, A. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *ArXiv Preprint ArXiv:1511.06434*.

Rakotonirina, N. C., & Rasoanaivo, A. (2020). ESRGAN+: Further improving enhanced super-resolution generative adversarial network. *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3637–3641.

Renieblas, G. P., Nogués, A. T., González, A. M., Gómez-Leon, N., & Del Castillo, E. G. (2017). Structural similarity index family for image quality assessment in radiological images. *Journal of Medical Imaging (Bellingham, Wash.)*, *4*(3), 35501. https://doi.org/10.1117/1.JMI.4.3.035501

Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., & Kennedy, R. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training gans. *Advances in Neural Information Processing Systems*, *29*.

Santurri, L., Aiazzi, B., Baronti, S., & Carlà, R. (2012). Influence of spatial resolution on pan-sharpening results. *2012 IEEE International Geoscience and Remote Sensing Symposium*, 5446–5449. https://doi.org/10.1109/IGARSS.2012.6352374

Sara, U., Akter, M., & Uddin, M. S. (2019). Image quality assessment through FSIM, SSIM, MSE and PSNR – a comparative study. *Journal of Computer and Communications*, *7*(3), 8–18.

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, *2*(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Sharma, L., Sengupta, S., & Kumar, B. (2021). An Improved Technique for Enhancement of Satellite Image. *Journal of Physics: Conference Series*, *1714*(1), 12051. https://doi.org/10.1088/1742-6596/1714/1/012051

Sharshov, I., Shoshina, K., Vasendina, I., & Aleshko, R. (2024). *Enhancing super-resolution in remote sensing : Integrating GIS data with CNN-based SRGAN models for improved image reconstruction*. *05006*.

Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., Rueckert, D., & Wang, Z. (2016). Real-time single image and video super-resolution using an efficient subpixel convolutional neural network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1874–1883.

Singla, K., Pandey, R., & Ghanekar, U. (2022). A review on Single Image Super Resolution techniques using generative adversarial network. *Optik*, *266*, 169607. https://doi.org/https://doi.org/10.1016/j.ijleo.2022.169607

Sinjari, S., & Kosovrasti, A. (2015). Geographic Information Systems (GIS) in Urban Planning. *European Journal of Interdisciplinary Studies*, *1*, 85. https://doi.org/10.26417/ejis.v1i1.p85-92

Tong, T., Li, G., Liu, X., & Gao, Q. (2017a). *Image super-resolution using dense skip connections. IEEE Int Conf Comp Vis. 2017; 2017: 4809–17*.

Tong, T., Li, G., Liu, X., & Gao, Q. (2017b). Image super-resolution using dense skip connections. *Proceedings of the IEEE International Conference on Computer Vision*, 4799–4807.

Tsatsaris, A., Kalogeropoulos, K., Stathopoulos, N., Louka, P., Tsanakas, K., Tsesmelis, D. E., Krassanakis, V., Petropoulos, G. P., Pappas, V., & Chalkias, C. (2021). Geoinformation technologies in support of environmental hazards monitoring under climate change: An extensive review. *ISPRS International Journal of Geo-Information*, *10*(2), 94.

Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P.-A. (2008). Extracting and composing robust features with denoising autoencoders. *Proceedings of the 25th International Conference on Machine Learning*, 1096–1103.

Wang, X., Yu, K., Wu, S., Gu, J., & Liu, Y. (2018). *ESRGAN : Enhanced Super-Resolution Generative Adversarial Networks*. 1–16.

Wang, Y., Zhang, Q., Wang, G.-G., & Cheng, H. (2024). The application of evolutionary computation in generative adversarial networks (GANs): a systematic literature survey. *Artificial Intelligence Review*, *57*(7), 182. https://doi.org/10.1007/s10462-024-10818-y

Wenlong, Z., Yihao, L., Dong, C., & Qiao, Y. (2021). RankSRGAN: Generative Adversarial Networks with Ranker for Image Super-Resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PP*, 1. https://doi.org/10.1109/TPAMI.2021.3096327

Williamson, I., Rajabifard, A., & Binns, A. (2007). The role of spatial data infrastructures in establishing an enabling platform for decision making in Australia. *Research and Theory in Advancing Spatial Data Infrastructure Concepts*, 121–132.

Wu, H., Gui, Z., & Yang, Z. (2020). Geospatial big data for urban planning and urban management. *Geo-Spatial Infor-*

mation Science, *23*, 273–274. https://doi.org/10.1080/10095020.2020.1854981

Xu, X., Hospedales, T. M., & Gong, S. (2016). Multi-task zero-shot action recognition with prioritised data augmentation. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, 343–359.

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, *9*(4), 611–629. https://doi.org/10.1007/s13244-018-0639-9

Yu, D., & Fang, C. (2023). Urban Remote Sensing with Spatial Big Data: A Review and Renewed Perspective of Urban Studies in Recent Decades. In *Remote Sensing* (Vol. 15, Issue 5). https://doi.org/10.3390/rs15051307

Zhang, J., Shao, M., Wan, Y., Meng, L., Cao, X., & Wang, S. (2024). Boundary-aware Spatial and Frequency Dual-domain Transformer for Remote Sensing Urban Images Segmentation. *IEEE Transactions on Geoscience and Remote Sensing*.

Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, *4*(2), 22–40.

Zhang, Y., Yu, W., & Zhu, D. (2022). Terrain feature-aware deep learning network for digital elevation model superresolution. *ISPRS Journal of Photogrammetry and Remote Sensing*, *189*, 143–162.

Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, *57*(4), 99. https://doi.org/10.1007/s10462-024-10721-6

SAŽETAK

Poboljšanje satelitske slike korištenjem dubokoga učenja i GIS integracije, sveobuhvatan pregled

U radu se donosi opsežan pregled 32 studije (20 časopisa, 11 zbornika i jedno poglavlje u knjizi) objavljenih od 2016. do 2023. u područjima dubokoga učenja (DL), poboljšanja slike, slike u superrezoluciji i geografskoga informacijskog sustava (GIS) usredotočujući se na integraciju DL metodologija s GIS-om radi poboljšanja kvalitete satelitskih slika. Pregled sažima pozadinu, načela, kvalitetu poboljšanja, brzinu i prednosti ovih tehnologija uspoređujući njihovu izvedbu na temelju metrike kao što su vršni omjer signala i šuma (PSNR), srednja kvadratna pogreška (MSE), korijen srednje kvadratne pogreške (RMSE), mjerenje indeksa strukturne sličnosti (SSIM) i vrijeme izračuna. Tehnologije satelitskoga daljinskog opažanja, koje su omogućile učinkovit način prikupljanja prostornih informacija od NASA-ina lansiranja Landsata 1 1972., nedavno su napredovale kako bi omogućile prikupljanje satelitskih (HRS) slika visoke rezolucije (≤ 30 cm). Međutim, čimbenici kao što su atmosferske smetnje, zasjenjenje i nedovoljna iskorištenost kapaciteta senzora često smanjuju kvalitetu slike. Kako bi se to riješilo, satelitske slike zahtijevaju poboljšanje, a DL se pokazao kao moćan alat zbog svoje sposobnosti modeliranja složenih odnosa i točnoga oporavka slika superrazlučivosti. Iako su DL i neuronske mreže pokazale znatan uspjeh u poboljšanju prirodne slike, njihova primjena na satelitske slike predstavlja jedinstven izazov. Ovi izazovi uključuju nedovoljno uzimanje u obzir različitih karakteristika satelitskih slika, kao što su različite prostorne rezolucije, šum senzora i spektralna raznolikost te oslanjanje na pretpostavke modeliranja koje se možda neće uskladiti sa složenošću satelitskih podataka. To naglašava potrebu za daljnjim istraživanjem naprednih DL pristupa posebno skrojenih za ovo područje.

Ključne riječi:
duboko učenje, GIS, neuronske mreže, satelitske slike, poboljšanje slike, superrezolucija

Author's contribution

Dalia A. Hussein (a candidate for a PhD of Science degree in Mining Engineering/Geodesy Surveying): conceptualization, literature review, investigation, synthesis of findings, writing – original draft, project administration, and writing – review & editing. **Mohamed A. Yousef** (an Associate Professor Emeritus in Mining Engineering/Geodesy Surveying): data curation, critical analysis of literature, thematic framework development, and writing – review & editing. **Hassan A. Abd el Hak** (Emeritus professor in Mining engineering/ surface Mining): supervision, validation, refinement of research scope, identification of key literature gaps, and writing – review & editing. **Yasser G. Mostafa** (a Professor in Civil Engineering/Engineering Surveying): assessment of remote sensing sources, validation of scientific content, visualization, and writing – review & editing.
All authors have read and approved the final version of the manuscript.