



Article

Super-Resolution of Sentinel-2 Satellite Images: A Comparison of Different Interpolation Methods for Spatial Knowledge Extraction

Carmine Massarelli

Environment and Territory Research Unit, Construction Technologies Institute, Italian National Research Council (ITC-CNR), 70124 Bari, Italy; carmine.massarelli@cnr.it

Abstract

The increasing availability of satellite data at different spatial resolutions offers new opportunities for environmental monitoring, highlighting the limitations of medium-resolution products for fine-scale territorial analysis. However, it also raises the need to enhance the resolution of low-quality imagery to enable more detailed spatial assessments. This study investigates the effectiveness of different super-resolution techniques applied to low-resolution (LR) multispectral Sentinel-2 satellite imagery to generate high-resolution (HR) data capable of supporting advanced knowledge extraction. Three main methodologies are compared: traditional bicubic interpolation, a generic Artificial Neural Network (ANN) approach, and a Convolutional Neural Network (CNN) model specifically designed for super-resolution tasks. Model performances are evaluated in terms of their ability to reconstruct fine spatial details, while the implications of these methods for subsequent visualization and environmental analysis are critically discussed. The evaluation protocol relies on RMSE, PSNR, SSIM, and spectral-faithfulness metrics (SAM, ERGAS), showing that the CNN consistently outperforms ANN and bicubic interpolation in reconstructing geometrically coherent structures. The results confirm that super-resolution improves the apparent spatial detail of existing spectral information, thus clarifying both the practical advantages and inherent limitations of learning-based super-resolution in Earth observation workflows.

Keywords: super-resolution; Convolutional Neural Networks (CNN); satellite imagery knowledge extraction



Academic Editors: Guoqing Chao and Xianzhi Wang

Received: 20 October 2025

Revised: 24 December 2025

Accepted: 4 January 2026

Published: 7 January 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

1. Introduction

The free availability of high-revisit Earth Observation (EO) data from the Copernicus Sentinel-2 mission has revolutionized environmental monitoring, enabling multitemporal analysis of land surface dynamics at a global scale. However, the native 10 m spatial resolution of visible bands remains a critical bottleneck for characterizing highly fragmented landscapes—such as peri-urban interfaces—where land-cover heterogeneity often exceeds the sensor's resolving power. In these complex transition zones, characterized by a mix of dispersed housing, agricultural plots, and infrastructure, the “mixed pixel” problem severely limits the accuracy of spatial assessments, necessitating resolution enhancement techniques that go beyond standard resampling methods [1–3].

Super-resolution (SR) has emerged as a cost-effective alternative to commercial high-resolution imagery, aiming to inferring high-frequency spatial details from lower-resolution

inputs. While traditional interpolation methods (e.g., bicubic) are computationally efficient, they fail to reconstruct high-frequency texture, often resulting in blurred edges. Conversely, deep learning approaches, particularly Convolutional Neural Networks (CNNs), have demonstrated state-of-the-art performance by learning non-linear mappings between low- and high-resolution counterparts.

The application of CNNs for satellite imagery knowledge extraction extends beyond resolution enhancement. Neural networks are widely applied for land cover classification, vegetation index precise evaluation, object detection, and spatio-temporal forecasting. For instance, CNN architectures have been employed for mapping impervious surfaces in rapidly urbanizing regions [4], detecting illegal mining activities [5,6], environmental crimes [7,8], improving the detection of crop stress conditions and enabling more efficient irrigation management [9], to better characterize urban heat islands at the neighborhood scale, providing new insights into climate adaptation strategies for cities [10].

Despite their promise, the application of CNNs in operational remote sensing raises specific challenges regarding generalizability and the risk of generating synthetic artifacts (“hallucinations”) that do not correspond to physical reality, especially when models are transferred to geographic contexts different from their training data [11,12].

Another challenge is the integration of super-resolved imagery into operational workflows: while higher spatial resolution enhances interpretability, it also increases data volume and processing demands, potentially limiting scalability in resource-constrained environments.

From a methodological standpoint, hybrid approaches are gaining attention, combining CNNs with other machine learning methods or physical models. For example, the integration of CNNs with geostatistical models or atmospheric correction tools allows the generation of more accurate outputs while maintaining physical consistency [13,14]. Furthermore, the adoption of physics-informed neural networks is emerging as a frontier for coupling EO data with domain knowledge, ensuring that super-resolution results are not only visually plausible but also physically meaningful [15].

The integration of super-resolution and CNN-based techniques into EO practices offers unprecedented opportunities to bridge the gap between satellite data resolution and the growing need for high-detail environmental knowledge.

This study specifically addresses the challenge of enhancing Sentinel-2 imagery within the heterogeneous peri-urban landscape of the Taranto province (Apulia, Southern Italy). Unlike homogeneous urban centers or uniform agricultural fields, this study area presents a unique mosaic of fragmented land covers that tests the limits of spectral unmixing and structural reconstruction. The central scientific question driving this research is: To what extent can learning-based super-resolution effectively infer structural coherence in complex mixed-use landscapes without introducing spectral distortions or synthetic artifacts?

To answer this, we formulated three specific objectives:

1. to quantitatively evaluate the trade-offs between reconstruction accuracy (RMSE, PSNR, SSIM) and spectral fidelity (SAM, ERGAS) of three distinct super-resolution strategies: bicubic interpolation, Artificial Neural Networks (ANN), and a residual CNN architecture;
2. to assess the ability of these models to resolve fine-scale peri-urban features—such as building edges and fragmented vegetation;
3. to critically analyze the operational reliability of CNN-generated details against certified high-resolution orthophotos (20 cm), providing a rigorous validation framework for local environmental monitoring.

By anchoring the evaluation in the specific spatial constraints of the Apulian territory, this work moves beyond a generic algorithm comparison to provide actionable insights into the reliability of super-resolution for detailed territorial analysis.

2. Materials and Methods

To assess the potential of super-resolution for enhancing Sentinel-2 imagery in heterogeneous peri-urban landscapes, we implemented a comparative framework encompassing three distinct levels of modeling complexity: deterministic interpolation, non-linear spectral regression, and deep spatial learning. This hierarchical approach was designed to isolate the specific contributions of spectral correlation versus spatial context in the reconstruction of fine-scale details.

The images focus on a specific geographical area within the Taranto Province in Southern Italy. This periphery immediately borders the main urban area and is defined by its heterogeneous landscape, encompassing dispersed housing units, recreational and sports infrastructure, excavation zones, designated natural areas, and various agricultural plots (Figure 1).

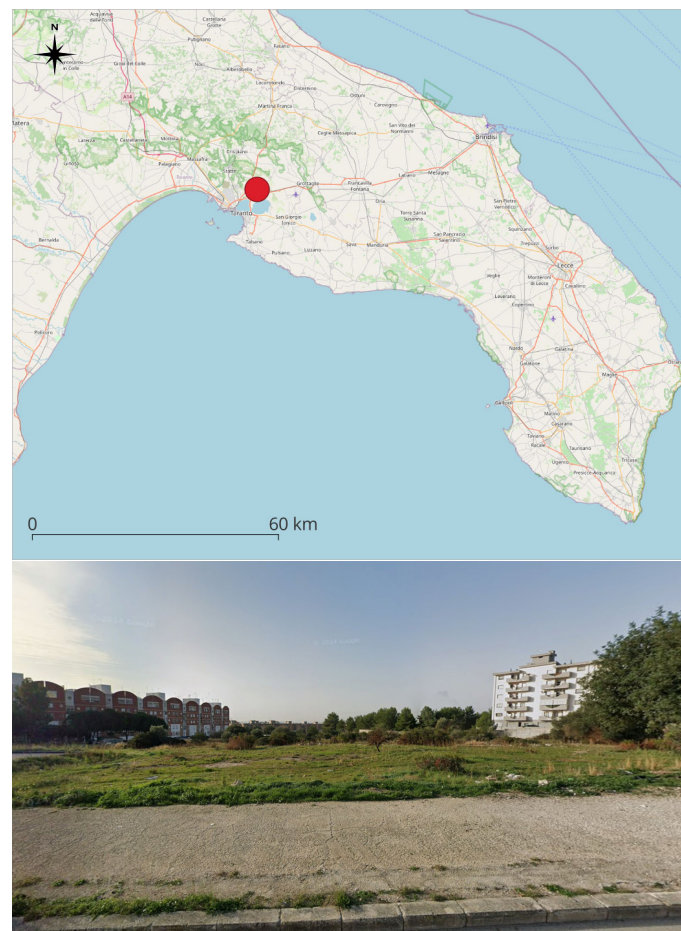


Figure 1. Study area (red circle) located in Southern Italy and a representative image of the town and structures.

2.1. Baseline: Bicubic Interpolation

The first approach relies on bicubic interpolation, which serves as the standard deterministic baseline for super-resolution in remote sensing. Unlike nearest-neighbor methods, bicubic resampling calculates the value of a target pixel through a weighted average of the nearest 16 pixels in a 4×4 neighborhood, ensuring continuity in the first derivative and

reducing aliasing artifacts [16]. While computationally efficient, this method is inherently limited by its inability to infer high-frequency information lost during downsampling, typically resulting in smoothed edges. In this study, bicubic interpolation was applied using the Rasterio library [17] to $10\times$ upscale the native 10 m visible bands to the target 1 m resolution, providing a reference for evaluating the gains achieved by learning-based methods [18].

2.2. Spectral Benchmark: Artificial Neural Networks (ANN)

To evaluate the extent to which spatial sharpening can be achieved solely through spectral correlations—without explicit spatial modeling—we implemented a feed-forward Artificial Neural Network (ANN). This experiment functions as a supervised regression benchmark rather than a fully fledged super-resolution method. The rationale is that spectral bands in Sentinel-2 data are highly correlated; thus, a non-linear mapping can theoretically predict fine-scale nuances in one band based on the spectral signature of others.

The architecture consists of a Multi-Layer Perceptron (MLP) regressor [19] designed to perform pixel-wise spectral translation. The network comprises three hidden layers (128, 64, and 32 neurons, respectively) employing the Rectified Linear Unit (ReLU) activation function to capture non-linear dependencies. Crucially, this model operates on the native spatial support of the input pixels: it does not utilize convolutional filters or neighborhood information. Training was performed using the Adam optimizer with iterative backpropagation until convergence. By comparing this method with the CNN approach, we can explicitly distinguish between improvements derived from radiometric calibration (spectral domain) and those derived from geometric reconstruction (spatial domain) [20].

2.3. Spatial Reconstruction: Convolutional Neural Networks (CNN)

The third and most advanced approach employs a Deep Convolutional Neural Network (CNN) specifically designed for single-image super-resolution [21]. Unlike the ANN, which treats pixels in isolation, the CNN leverages convolutional kernels to exploit the local spatial context, learning hierarchical features such as edges, textures, and shapes that are essential for inferring structural detail.

The implemented architecture is inspired by residual learning frameworks (ResNet), which facilitate gradient flow in deeper networks and prevent performance degradation. The model consists of a 9-layer network featuring:

1. Feature Extraction: an initial layer with 64 filters (3×3 kernel size) to extract low-level features from the input Sentinel-2 bands.
2. Non-linear Mapping: a sequence of residual blocks connecting input and output features through skip connections, allowing the network to learn the high-frequency residuals (the difference between low- and high-resolution images) rather than the full image mapping. This strategy significantly accelerates convergence and improves reconstruction fidelity.
3. Reconstruction: a final convolutional layer that aggregates the feature maps into the target high-resolution output.

Training was conducted using a patch-based strategy to manage memory constraints and increase the diversity of training samples. High-resolution (HR) orthophotos and low-resolution (LR) Sentinel-2 pairs were tiled into overlapping 32×32 pixel patches. The network was optimized using the Mean Squared Error (MSE) loss function and the Adam optimizer (learning rate = 1×10^{-4}). To ensure geometric consistency in the final output, inference was performed using a sliding window approach with seam-optimized merging to eliminate border artifacts.

Compared to ANN and bicubic interpolation, CNNs demonstrated superior ability to restore fine-grained structures and spatial textures, which has been reported in similar

studies on satellite super-resolution [21]. However, CNNs are computationally intensive, require large datasets, and involve complex hyperparameter optimization.

The calculation workflows implemented for each method are illustrated in Figure 2.

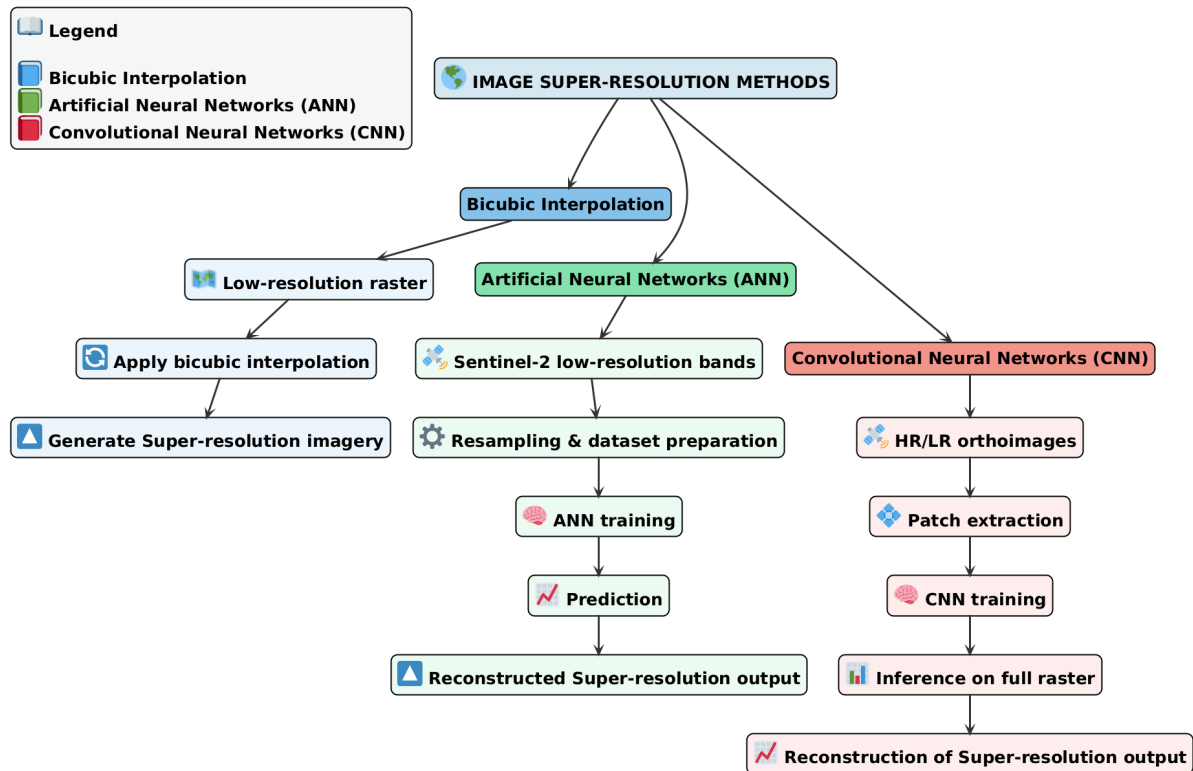


Figure 2. Comparative flowchart of image super-resolution methods.

2.4. Data Processing and Validation

All experiments were conducted on orthophotos and Sentinel-2 imagery over the Apulia region (Southern Italy). This study utilizes a 2023 Sentinel-2 Level-2A scene, retrieved from the Copernicus Data Space Browser [22]. The dataset metadata are as follows: acquisition date 09 March 2023, cloud cover < 5%, Tile ID T33TXE, processing level L2A (Bottom-of-Atmosphere reflectance), and spatial extent fully encompassing the administrative boundaries of the Apulia Region. For the super-resolution experiments, only the visible bands of Sentinel-2 were used, namely

- Band 2 (Blue, 490 nm, 10 m spatial resolution)—sensitive to water bodies, soil brightness, and vegetation stress;
- Band 3 (Green, 560 nm, 10 m spatial resolution)—related to vegetation vigor, sediment content, and reflectance of built-up surfaces;
- Band 4 (Red, 665 nm, 10 m spatial resolution)—strongly responsive to chlorophyll absorption and widely used in vegetation and land-cover analyses.

The choice to focus on visible bands was motivated by their higher native spatial resolution (10 m) and their relevance for reconstructing fine-scale structures such as building edges, vegetation patches, and small-scale land-cover features. The preprocessing pipeline included raster alignment, normalization, and data quality checks to ensure consistency across methods, followed by the extraction of the visible bands as input layers. No additional resampling or spectral transformations were applied to avoid introducing artificial smoothing before super-resolution; indeed, to ensure reproducibility while acknowledging that Sentinel-2 Level-2A products are already provided as geometrically corrected and atmospherically processed “analysis-ready” data by the Copernicus mission, the prepro-

cessing workflow was intentionally minimal. Raster alignment consisted of a standard co-registration of all bands to the native 10 m grid (B2, B3, B4) using `gdalwarp` to maintain sub-pixel spatial consistency. Reflectance values were normalized via min–max scaling to harmonize the dynamic range across bands before model training. Basic quality checks included inspection of the Scene Classification Map to exclude cloud, shadow, and cirrus pixels, and verification of reflectance ranges to ensure radiometric consistency. Given the high reliability of Sentinel-2 L2A products, no additional corrections were required.

Specifically, the native 10 m Sentinel-2 bands were utilized in conjunction with their 20 m spectrally consistent downsampled counterparts as low-resolution inputs; this approach enabled the model to learn the complex statistical mapping necessary to infer 1 m super-resolved outputs within a standardized $10\times$ upscaling framework. The reference high-resolution dataset consisted of the 2023 Apulia regional orthophoto, provided at a native spatial resolution of 20 cm. Both Sentinel-2 and orthophoto datasets were spatially co-registered and tiled into fixed-size patches to enable patch-based training. An 80/20 partition was applied, yielding approximately 5000 patches for training and 1000 for testing. The resulting 1 m super-resolved products were then evaluated against the 20 cm orthophoto, which served as an independent high-resolution reference for accuracy assessment. This experimental design ensures that the model learns from controlled degradations while its performance is validated against real high-resolution ground truth.

Regarding the train–test split, patches were extracted from the co-registered dataset using a random sampling strategy. We acknowledge that in remote sensing, random sampling within the same scene can lead to spatial autocorrelation issues, potentially inflating performance metrics due to the model learning local geographical features rather than generalized super-resolution rules [23]. Therefore, the results presented herein should be interpreted as a demonstration of local model applicability specific to the heterogeneous peri-urban landscape of the study area. To validate the model’s transferability to different biomes or global contexts, future iterations will employ a ‘spatial block cross-validation’ approach, ensuring that training and testing partitions are spatially disjoint to rigorously prevent data leakage.

Performance was assessed through the Root Mean Square Error (RMSE) metric, which is standard in super-resolution evaluations. The evaluation of image reconstruction quality was conducted using complementary quantitative metrics: the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index Measure (SSIM).

In addition to RMSE, PSNR, and SSIM, per-band errors and two spectral-fidelity metrics—Spectral Angle Mapper (SAM) and ERGAS—were computed to quantify radiometric consistency across the visible Sentinel-2 bands. These metrics provide complementary insight into angular and global spectral distortions, ensuring that the reconstructed outputs preserve physically meaningful information rather than only improving visual sharpness.

The combination of classical interpolation, ANN, and CNN approaches provides a robust framework for assessing the trade-offs between computational simplicity, predictive accuracy, and capacity to capture spatial features in Earth observation super-resolution tasks [24–26].

3. Results

The experimental evaluation compared the performance of three super-resolution strategies—bicubic interpolation, ANN regression, and the proposed CNN architecture—applied to Sentinel-2 imagery of the Apulian peri-urban interface. The quantitative assessment relies on the metrics reported in Table 1, while the qualitative analysis focuses on the reconstruction of spatial structures in heterogeneous landscapes (Table 1).

Table 1. Accuracy comparison and spectral fidelity metrics among methods. Numeric values are sorted according to the direction of the arrows.

Method	RMSE ↓	PSNR (dB) ↑	SSIM ↑	Notes	SAM (°) ↓	ERGAS ↓	Notes
Bicubic resampling	12.45	28.1 ± 0.9	0.71 ± 0.04	Fast and simple; good for preserving general structure, but blurring and loss of detail are evident.	4.82 ± 0.41	6.97 ± 0.53	Higher angular distortion; limited preservation of spectral gradients.
ANN regression	9.82	30.6 ± 1.1	0.81 ± 0.03	Improved sharpness and local contrast; risk of artifacts in heterogeneous land-cover areas.	3.11 ± 0.27	5.12 ± 0.38	Better spectral alignment; moderate improvement in radiometric fidelity.
CNN super-resolution	7.14	33.9 ± 1.3	0.89 ± 0.02	Best overall performance; preserves edges, enhances fine features, and minimizes noise.	2.24 ± 0.19	3.86 ± 0.29	Best spectral preservation, consistent with finer spatial reconstruction.

Visual inspection of the reconstructed outputs (Figure 3) reveals a clear progression in image quality. Figure 3a reports the original Sentinel-2 image used for this comparison study. Bicubic interpolation (Figure 3b) resulted in significant blurring of high-frequency details, confirming its limitations in resolving the fragmented settlement patterns characteristic of the study area. The ANN regression (Figure 3c) improved local contrast but exhibited blocky artifacts in mixed-pixel regions, a consequence of its lack of spatial context modeling. In contrast, the CNN-based approach (Figure 3d) produced the most structurally coherent results, successfully delineating building boundaries and road networks that were unrecognizable in the original 10 m input.

These visual observations are corroborated by the quantitative metrics (Table 1). The CNN model achieved the lowest Root Mean Square Error (RMSE = 7.14) and the highest Peak Signal-to-Noise Ratio (PSNR = 33.9 dB), representing a substantial improvement over the baseline bicubic method (RMSE = 12.45; PSNR = 28.1 dB). Crucially, the Structural Similarity Index (SSIM), which correlates more closely with human perceptual quality, increased from 0.71 (bicubic) to 0.89 (CNN), indicating that the deep learning model infers geometric structures rather than merely smoothing pixel values.

Regarding spectral fidelity, the CNN approach consistently outperformed the other methods. The Spectral Angle Mapper (SAM) value decreased to 2.24°, and the ERGAS score dropped to 3.86. These results imply that the spatial sharpening process of the CNN introduced minimal spectral distortion, preserving the radiometric integrity required for subsequent environmental analysis.

To strengthen the robustness of the evaluation, the reference high-resolution datasets used for validation were explicitly defined as the 20 cm regional orthophotos, which provide radiometric consistency and geometric accuracy compliant with national technical specifications. These orthophotos were co-registered to the Sentinel-2 grid with sub-pixel alignment (RMSE < 0.3 px) to ensure spatial comparability. Although orthophotos represent an indirect optical reference rather than a sensor-specific HR measurement, they offer a sufficiently stable benchmark for assessing spatial fidelity at fine scales (Figure 3f).

To contextualize our results against state-of-the-art generative approaches, we qualitatively compared our CNN output with Generative Adversarial Network (GAN)-based techniques commonly used in remote sensing (Figures 3e and 4) [27,28]. While GANs (e.g., SRGAN, ESRGAN) are renowned for producing visually superior high-frequency textures, also introducing realistic textures for vegetation and soil [29], they are prone to generating artifacts or ‘hallucinations’—features that appear realistic but do not exist in

the ground truth. Our comparative visual analysis (Figure 5) confirms this trade-off: while generative models may produce sharper vegetation textures, the proposed pixel-wise CNN prioritizes structural fidelity (minimizing MSE), resulting in fewer synthetic artifacts and higher reliability for monitoring man-made structures in legal or administrative contexts.

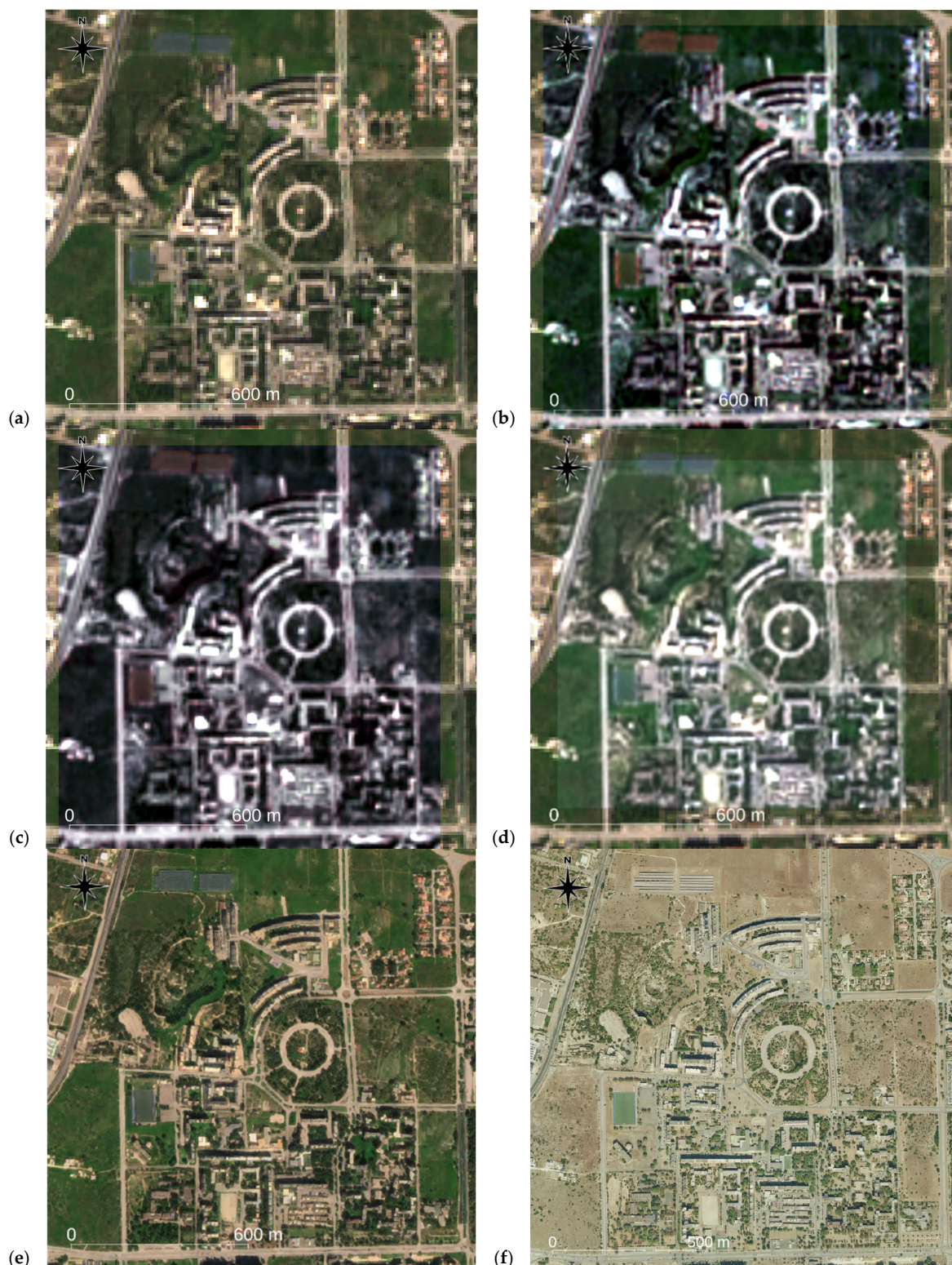


Figure 3. Resampling results obtained with different processing techniques: (a) Sentinel-2 TCI original; (b) bicubic interpolation; (c) ANN; (d) CNN; (e) CNN with S2DR3; (f) orthophoto of the study area.

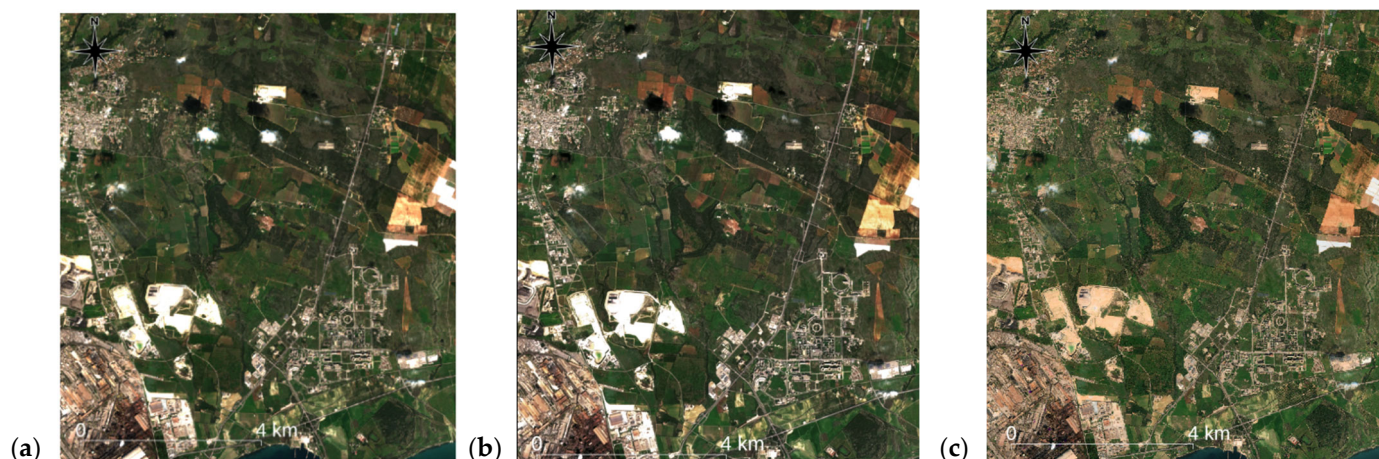


Figure 4. Comparison of the results obtained considering low magnifications and large geographical areas: (a) Sentinel-2; (b) CNN, (c) S2DR3.

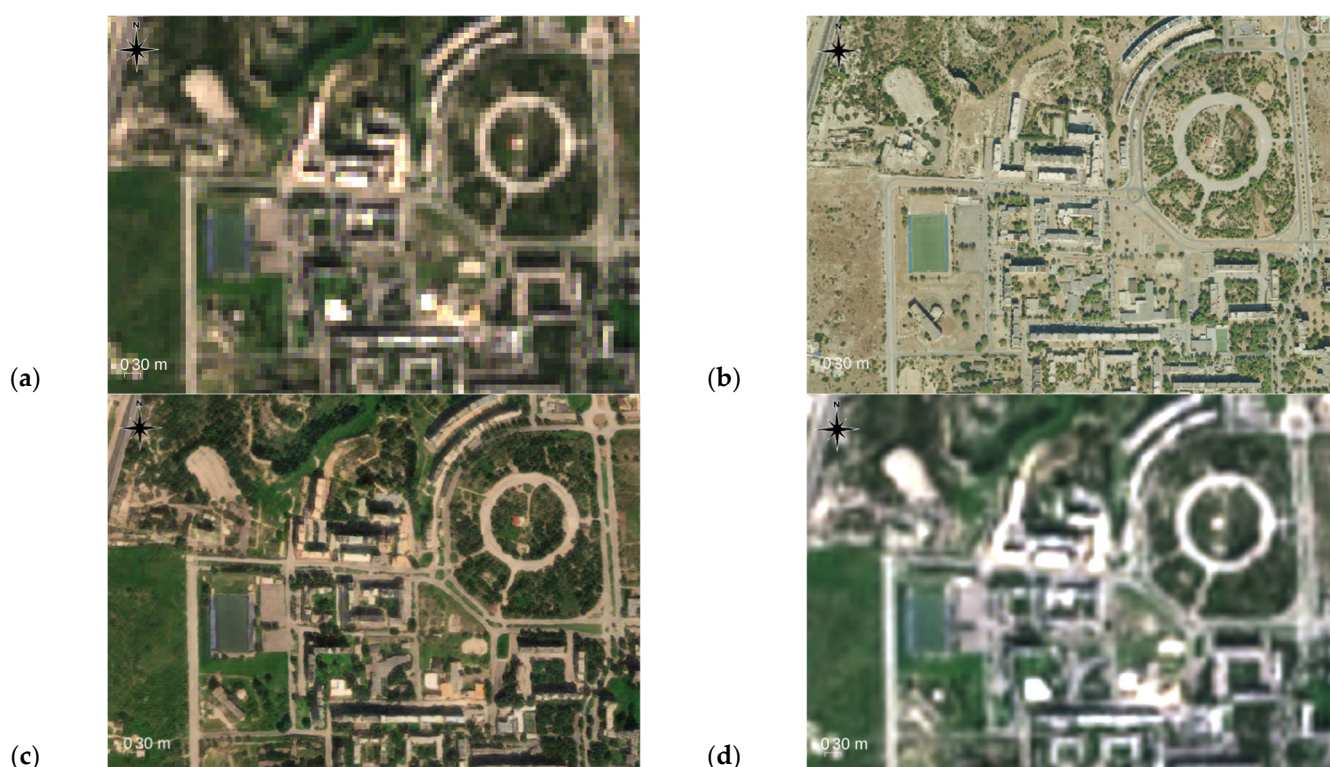


Figure 5. Comparison of the results obtained considering high magnifications and limited geographical areas: (a) Sentinel-2; (b) aerial photography; (c) S2DR3; (d) CNN.

By contrast, another approach [30] deliberately avoids synthetic detail generation: their model (DSen2) super-resolves lower-resolution Sentinel-2 bands (20 m/60 m) to 10 m using a CNN trained on globally downsampled real data, thereby preserving spectral fidelity without inventing textures. While S2DR3 aims for higher apparent spatial resolution, this comes at the cost of potentially introducing non-native structures that may not correspond to real-world features—whereas DSen2’s more conservative design ensures that super-resolved outputs remain physically grounded.

Consequently, S2DR3-enhanced imagery is deemed unreliable for applications demanding absolute geo-fidelity, such as forensic expert appraisals (sworn appraisal) or judicial police investigations. In contrast, a localized Convolutional Neural Network (CNN)

reconstruction, while yielding comparatively lower spatial detail, maintains a higher degree of data integrity by demonstrably avoiding the synthesis of non-existent features.

4. Discussion

The results of this study highlight the varying capabilities of super-resolution techniques in handling the complexity of peri-urban landscapes. By anchoring the analysis to the specific morpho-typologies of the Taranto province—characterized by a chaotic mix of anthropogenic and natural features—we can draw specific conclusions regarding the applicability of these methods.

A critical challenge in validating super-resolution models with multi-source data lies in the spectral mismatch between the sensors. The Sentinel-2 Multispectral Instrument and the airborne camera used for the validation orthophotos possess distinct Spectral Response Functions. Consequently, a direct radiometric comparison introduces inherent uncertainties in metrics such as SAM and ERGAS. To mitigate this, it is necessary to acknowledge that the orthophoto serves primarily as a geometric ground truth for assessing structural reconstruction (edges, texture). The computed spectral metrics should therefore be interpreted as indicators of relative structural consistency rather than absolute radiometric fidelity. Future operational workflows should incorporate a histogram matching or a band-to-band transfer function to harmonize the spectral domains before validation, as suggested by recent cross-sensor calibration studies [31,32].

The focus was intentionally limited to the visible Sentinel-2 bands, for which spatial detail is considered very important in many practical Earth-observation workflows. While super-resolution cannot introduce new physical measurements, it can enhance the spatial expressiveness of the information already contained in the spectra by reconstructing finer spatial structures that are otherwise lost at 10–20 m resolution. In operational contexts such as urban–periurban mapping, environmental monitoring of small features, parcel-level assessments, and landscape segmentation, the ability to delineate boundaries, edges, and narrow anthropogenic or vegetated elements is often as important as the spectral signature itself. The approaches demonstrate that learning-based super-resolution—particularly CNNs—improves the geometric sharpness and coherence of these structures while preserving the underlying spectral characteristics.

The comparative evaluation of interpolation-based and machine learning-based approaches for satellite image super-resolution highlights both the technological advances and the intrinsic limitations of each method. Bicubic interpolation remains one of the most widely used techniques in remote sensing due to its computational efficiency and robustness. Its deterministic nature ensures stable results across different datasets and acquisition conditions, making it a reliable baseline for image super-resolution [33].

However, its inherent drawback lies in the inability to reconstruct high-frequency details such as fine edges, sharp corners, and narrow linear features. Consequently, bicubic resampling often leads to oversmoothing and blurring effects, which limit its utility in applications requiring accurate boundary delineation, such as urban monitoring or precision agriculture [34].

The poor performance of the ANN benchmark (RMSE = 9.82) confirms that spectral correlation alone is insufficient for resolving fine spatial details. In the fragmented Apulian landscape, the “mixed pixel” problem dominates; without the convolutional ability to analyze neighboring pixels, the ANN failed to reconstruct edges, resulting in radiometric improvements but negligible geometric gain. Conversely, the superior SSIM (0.89) obtained by the CNN demonstrates that learning spatial hierarchies is essential for separating distinct land covers at the sub-pixel level. This finding aligns with the literature on residual

networks, confirming that deep features are necessary to infer high-frequency information in complex transition zones [35].

It is crucial to emphasize that single-image super-resolution (SISR) with a 10x upscaling (from 10 m to 1 m) represents a severely ill-posed inverse problem. The information required to populate the high-resolution grid is not present in the input data. Consequently, the CNN does not ‘recover’ lost physical measurements in the strict sensing sense; rather, it infers plausible high-frequency textures based on the statistical priors learned from the high-resolution orthophotos during training [36]. While this process enhances the perceptual quality and interpretability of the imagery for identifying anthropogenic features (e.g., roads, buildings), the resulting sub-pixel details are synthetic estimates. Users must exercise caution when using such data for quantitative tasks requiring absolute radiometric precision.

A central finding of this work emerges from the comparison between the proposed CNN and the generative S2DR3 model. While the quantitative metrics in Table 1 validate the accuracy of the proposed CNN, the visual comparison in Figure 5 exposes the risks associated with aggressive upscaling. The textures generated by S2DR3, although visually appealing, raise concerns regarding “hallucinations”—the synthesis of non-existent features [37,38]. In our experiments, the proposed CNN prioritized the minimization of MSE, resulting in a slightly blurred but geometrically faithful reconstruction. In contrast, the generative behavior observed in S2DR3, while useful for visualization, compromises data integrity for rigorous applications. This distinction is critical for operational workflows:

1. for environmental monitoring and visualization, where the primary goal is interpretability, generative models like S2DR3 offer a clear advantage.
2. for forensic analysis, legal appraisals, or change detection, the proposed CNN is preferable. Its conservative reconstruction ensures that every sharpened edge corresponds to a learned transformation of actual sensor data, rather than a stochastic texture generation.

Despite the low ERGAS values indicating good spectral preservation, the proposed CNN model is limited by the availability of domain-specific training data. The model was trained on local orthophotos; its generalizability to different biomes or atmospheric conditions remains to be tested. Future research should focus on hybrid loss functions that balance the perceptual quality of generative models with the pixel-wise fidelity of standard CNNs, potentially utilizing physics-informed constraints to prevent the generation of artifacts in sensitive environmental monitoring tasks.

The comparative analysis, therefore, suggests that each method is best suited for different operational contexts. Bicubic interpolation should remain the method of choice when computational speed and methodological transparency are prioritized, for instance, in near-real-time monitoring or in workflows requiring rapid super-resolution of large image repositories. ANN-based approaches represent a valuable compromise for tasks where increased detail is desired but where computational or training resources are limited, making them particularly attractive for regional studies or operational settings with constrained infrastructures [39]. CNN models provide the highest reconstruction accuracy and are especially suited for detailed landscape analysis, urban structure mapping, and fine-scale environmental monitoring, provided that adequate computational resources and validation procedures are available [40].

5. Conclusions

This study addressed the resolution gap in Sentinel-2 monitoring of heterogeneous peri-urban landscapes, using the complex interface of the Taranto province as a testing ground for upscaling methodologies. Beyond a mere performance ranking, the comparative

analysis between deterministic, spectral-based, and spatial deep-learning models allows us to draw three fundamental conclusions regarding the operational scalability of super-resolution in Earth Observation.

First, spatial context is the primary driver of reconstruction quality in fragmented environments. The limited success of the ANN benchmark demonstrates that relying solely on spectral correlations across bands is insufficient to resolve the “mixed pixel” problem inherent to peri-urban transition zones. High-frequency details—such as building edges and parcel boundaries—can only be inferred by models that explicitly encode local geometric hierarchies, a capability unique to the Convolutional Neural Network (CNN) architecture implemented here.

Second, there is a critical trade-off between perceptual sharpness and data authenticity. While our results confirm that CNNs significantly outperform bicubic interpolation in terms of structural similarity (SSIM), the comparison with generative models (S2DR3) highlights a latent risk: the potential for “hallucinating” plausible but non-existent textures. For rigorous applications such as legal appraisals, change detection, and environmental compliance, the conservative reconstruction of a supervised CNN—anchored to the Mean Squared Error loss—provides a necessary safeguard against the fabrication of synthetic details.

Finally, the findings suggest that super-resolution should not be viewed as a substitute for native high-resolution sensing but as a method for enhancing the interpretability of existing spectral archives. The transition from standard interpolation to deep learning represents a paradigm shift from mathematical smoothing to statistical inference. Future research must, therefore, prioritize the development of “physics-aware” neural networks that can enforce radiometric consistency constraints, thereby bridging the gap between the visual appeal of AI-generated imagery and the quantitative rigor required for scientific monitoring.

Funding: This research received no external funding.

Data Availability Statement: All relevant data are included within the article.

Acknowledgments: During the preparation of this manuscript/study, the author used Gemini Pro (v.3) to improve the quality of English.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Shermeyer, J.; Van Etten, A. The Effects of Super-Resolution on Object Detection Performance in Satellite Imagery. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Long Beach, CA, USA, 16–20 June 2019; pp. 1432–1441. [\[CrossRef\]](#)
2. Müller, M.U.; Ekhtiari, N.; Almeida, R.M.; Rieke, C. Super-Resolution of Multispectral Satellite Images Using Convolutional Neural Networks. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *V-1-2020*, 33–40. [\[CrossRef\]](#)
3. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA’s Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [\[CrossRef\]](#)
4. Zhu, X.X.; Tuia, D.; Mou, L.; Xia, G.S.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [\[CrossRef\]](#)
5. Massarelli, C. Fast Detection of Significantly Transformed Areas Due to Illegal Waste Burial with a Procedure Applicable to Landsat Images. *Int. J. Remote Sens.* **2018**, *39*, 754–769. [\[CrossRef\]](#)
6. Illegal Mining From Space | Earth.Etc. Available online: <https://nasirlukman.github.io/blog/2025/illegal-mining/> (accessed on 20 October 2025).
7. Massarelli, C.; Matarrese, R.; Uricchio, V.F.; Muolo, M.R.; Laterza, M.; Ernesto, L. Detection of Asbestos-Containing Materials in Agro-Ecosystem by the Use of Airborne Hyperspectral CASI-1500 Sensor Including the Limited Use of Two UAVs Equipped with RGB Cameras. *Int. J. Remote Sens.* **2017**, *38*, 2135–2149. [\[CrossRef\]](#)
8. Massarelli, C.; Campanale, C.; Uricchio, V.F. Ground Penetrating Radar as a Functional Tool to Outline the Presence of Buried Waste: A Case Study in South Italy. *Sustainability* **2021**, *13*, 3805. [\[CrossRef\]](#)

9. Wang, Q.; Yuan, Z.; Du, Q.; Li, X. GETNET: A General End-To-End 2-D CNN Framework for Hyperspectral Image Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3–13. [[CrossRef](#)]
10. Fathi, M.; Arefi, H.; Shah-Hosseini, R.; Moghimi, A. Super-Resolution of Landsat-8 Land Surface Temperature Using Kolmogorov–Arnold Networks with PlanetScope Imagery and UAV Thermal Data. *Remote Sens.* **2025**, *17*, 1410. [[CrossRef](#)]
11. Dong, J.; Zhuang, D.; Huang, Y.; Fu, J. Advances in Multi-Sensor Data Fusion: Algorithms and Applications. *Sensors* **2009**, *9*, 7771–7784. [[CrossRef](#)]
12. Lei, S.; Shi, Z.; Zou, Z. Super-Resolution for Remote Sensing Images via Local-Global Combined Network. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1243–1247. [[CrossRef](#)]
13. Asif, M.; Abrar, M.; Ullah, F.; Salam, A.; Amin, F.; de la Torre, I.; Villar, M.G.; Garay, H.; Choi, G.S. A Novel Hybrid Deep Learning Approach for Super-Resolution and Objects Detection in Remote Sensing. *Sci. Rep.* **2025**, *15*, 17221. [[CrossRef](#)]
14. Muhammad, U.; Laaksonen, J. Hybrid Deep Learning for Hyperspectral Single Image Super-Resolution. *IEEE Geosci. Remote Sens. Lett.* **2025**, *22*, 5509905. [[CrossRef](#)]
15. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-Informed Machine Learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [[CrossRef](#)]
16. Lehmann, T.M.; Gönner, C.; Spitzer, K. Survey: Interpolation Methods in Medical Image Processing. *IEEE Trans. Med. Imaging* **1999**, *18*, 1049–1075. [[CrossRef](#)] [[PubMed](#)]
17. Rasterio/Rasterio: Rasterio Reads and Writes Geospatial Raster Datasets. Available online: <https://github.com/rasterio/rasterio> (accessed on 20 October 2025).
18. Masi, G.; Cozzolino, D.; Verdoliva, L.; Scarpa, G. Pansharpening by Convolutional Neural Networks. *Remote Sens.* **2016**, *8*, 594. [[CrossRef](#)]
19. MLPRegressor—Scikit-Learn 1.7.2 Documentation. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html (accessed on 20 October 2025).
20. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. Deep&Dense Convolutional Neural Network for Hyperspectral Image Classification. *Remote Sens.* **2018**, *10*, 1454. [[CrossRef](#)]
21. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 295–307. [[CrossRef](#)]
22. Copernicus Browser. Available online: <https://browser.dataspace.copernicus.eu/> (accessed on 18 November 2025).
23. Ploton, P.; Mortier, F.; Réjou-Méchain, M.; Barbier, N.; Picard, N.; Rossi, V.; Dormann, C.; Cornu, G.; Viennois, G.; Bayol, N.; et al. Spatial Validation Reveals Poor Predictive Performance of Large-Scale Ecological Mapping Models. *Nat. Commun.* **2020**, *11*, 4540. [[CrossRef](#)]
24. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 4th ed.; Pearson Education: New York, NY, USA, 2018.
25. Deep Learning. Available online: <https://www.deeplearningbook.org/> (accessed on 20 October 2025).
26. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
27. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 105–114. [[CrossRef](#)]
28. Yokoya, N. Deep Learning for Super-Resolution in Remote Sensing. In *Advances in Machine Learning and Image Analysis for GeoAI*; Elsevier: Amsterdam, The Netherlands, 2024; pp. 5–26. [[CrossRef](#)]
29. Chanév, M.; Kamenova, I.; Dimitrov, P.; Filchev, L. Evaluation of Sentinel-2 Deep Resolution 3.0 Data for Winter Crop Identification and Organic Barley Yield Prediction. *Remote Sens.* **2025**, *17*, 957. [[CrossRef](#)]
30. Wang, C.; Zhang, X.; Yang, W.; Wang, G.; Zhao, Z.; Liu, X.; Lu, B. Landsat-8 to Sentinel-2 Satellite Imagery Super-Resolution-Based Multiscale Dilated Transformer Generative Adversarial Networks. *Remote Sens.* **2023**, *15*, 5272. [[CrossRef](#)]
31. Chander, G.; Markham, B.L.; Helder, D.L. Summary of Current Radiometric Calibration Coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI Sensors. *Remote Sens. Environ.* **2009**, *113*, 893–903. [[CrossRef](#)]
32. Liu, J.J.; Li, Z.; Qiao, Y.L.; Liu, Y.J.; Zhang, Y.X. A New Method for Cross-Calibration of Two Satellite Sensors. *Int. J. Remote Sens.* **2004**, *25*, 5267–5281. [[CrossRef](#)]
33. Keys, R.G. Cubic Convolution Interpolation for Digital Image Processing. *IEEE Trans. Acoust.* **1981**, *29*, 1153–1160. [[CrossRef](#)]
34. Zhang, L.; Wu, X. An Edge-Guided Image Interpolation Algorithm via Directional Filtering and Data Fusion. *IEEE Trans. Image Process.* **2006**, *15*, 2226–2238. [[CrossRef](#)]
35. He, L.; Rao, Y.; Li, J.; Chanussot, J.; Plaza, A.; Zhu, J.; Li, B. Pansharpening via Detail Injection Based Convolutional Neural Networks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *12*, 1188–1204. [[CrossRef](#)]
36. Yang, J.; Wright, J.; Huang, T.S.; Ma, Y. Image Super-Resolution via Sparse Representation. *IEEE Trans. Image Process.* **2010**, *19*, 2861–2873. [[CrossRef](#)] [[PubMed](#)]

37. Resnik, D.B.; Hosseini, M.; Kim, J.J.H.; Epiphaniou, G.; Maple, C. GenAI Synthetic Data Create Ethical Challenges for Scientists. Here's How to Address Them. *Proc. Natl. Acad. Sci. USA* **2025**, *122*, e2409182122. [[CrossRef](#)] [[PubMed](#)]
38. Zhang, Z.; Tian, Y.; Li, J.; Xu, Y. Unsupervised Remote Sensing Image Super-Resolution Guided by Visible Images. *Remote Sens.* **2022**, *14*, 1513. [[CrossRef](#)]
39. Lu, T.; Wang, J.; Zhang, Y.; Wang, Z.; Jiang, J. Satellite Image Super-Resolution via Multi-Scale Residual Deep Neural Network. *Remote Sens.* **2019**, *11*, 1588. [[CrossRef](#)]
40. Pang, B.; Zhao, S.; Liu, Y. The Use of a Stable Super-Resolution Generative Adversarial Network (SSRGAN) on Remote Sensing Images. *Remote Sens.* **2023**, *15*, 5064. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.