

# Learning-Based Restoration of Raw Satellite Imagery for Efficient Onboard Processing

Anonymous CVPR submission

Paper ID \*\*\*\*\*

## Abstract

001 *Satellite image restoration aims to improve image qual-*  
002 *ity by compensating degradations (e.g., noise and blur)*  
003 *introduced by the imaging system and acquisition condi-*  
004 *tions. As a fundamental preprocessing step, restoration di-*  
005 *rectly impacts both ground-based product generation and*  
006 *emerging onboard AI applications. Traditional restoration*  
007 *pipelines are based on sequential physical models compu-*  
008 *tationally and memory demanding involving multiple oper-*  
009 *ations and intermediate buffers, and poorly suited to re-*  
010 *source constrained onboard environments. We investigate*  
011 *whether a light and non-generative residual convolutional*  
012 *network (EDSR), trained exclusively on physics based sim-*  
013 *ulated satellite data, can match or surpass a traditional*  
014 *ground-processing restoration pipeline across multiple op-*  
015 *erating conditions ./ while meeting the efficiency require-*  
016 *ments of onboard deployment. A realistic simulation frame-*  
017 *work modeling ground sampling distance (GSD), modula-*  
018 *tion transfer function (MTF), and signal dependent noise*  
019 *is used to generate supervised training pairs without rely-*  
020 *ing on unavailable raw/L1 ground-truth data. By sampling*  
021 *degradations across a wide range of operating conditions,*  
022 *the model learns robustness to variability in blur and noise.*  
023 *Experiments conducted on simulated datasets, real Pleiades*  
024 *imagery, and object detection tasks demonstrate that the*  
025 *proposed approach achieve competitive performance with*  
026 *respect to the traditional pipeline in image quality while*  
027 *significantly reducing end-to-end processing time. Physical*  
028 *indicators confirm robust blur compensation across varying*  
029 *degradation levels. Moreover, restoration consistently im-*  
030 *proves downstream detection performance, highlighting its*  
031 *relevance for onboard AI pipelines. Finally, we demonstrate*  
032 *its effective deployment on embedded hardware, validating*  
033 *its onboard applicability.*

## 1. Introduction

034

High-resolution optical satellite imagery is acquired on-  
board at Raw level, where sensor optical limitations and  
acquisition conditions, including modulation transfer func-  
tion effects and noise, inherently degrade image quality.  
Image restoration aims to compensate for these degrada-  
tions by jointly addressing optical blur, sensor noise, and  
acquisition-related effects. It therefore constitutes a pre-  
requisite pre-processing step for any higher-level exploita-  
tion of satellite imagery. This is true not only for traditional  
ground-based workflows but also for emerging onboard pro-  
cessing pipelines, where early enhancement of image qual-  
ity can directly benefit downstream analytical tasks.

035

036

037

038

039

040

041

042

043

044

045

046

Traditional restoration approaches are typically based on  
sequential physical models, combining deconvolution and  
denoising stages [5, 10, 11]. While robust and well estab-  
lished, such pipelines, historically deployed on ground in-  
frastructures, are computationally demanding, often involv-  
ing iterative optimization procedures and multiple interme-  
diate buffers. This results in significant memory usage and  
processing latency, which severely limit their suitability for  
onboard implementation. Furthermore, their effectiveness  
relies on precise knowledge of acquisition parameters and  
imaging system characteristics. Yet, in-flight operational  
conditions induce dynamic variations in the physical image  
formation process (e.g., thermomechanical effects, micro-  
vibrations, pointing instability), leading to deviations from  
nominal models that are difficult to anticipate and faith-  
fully capture. Such variability challenges the robustness of  
purely model-based restoration methods.

047

048

049

050

051

052

053

054

055

056

057

058

059

060

061

062

063

The increasing demand for low-latency applications, in-  
cluding disaster response [ Mini biblio TODO], maritime  
surveillance[ Mini biblio TODO], and autonomous on-  
board decision-making, motivates a shift toward more effi-  
cient processing strategies. In this context, recent learning-  
based approaches enable joint optimization of blur compen-  
sation and noise reduction within a single model [2–4, 6]  
offering a promising alternative for ground and board.

064

065

066

067

068

069

070

071

In this work, we investigate whether a light non-

072

generative residual convolutional neural network, EDSR [13], trained exclusively on physically realistic simulated satellite data, can achieve restoration performance comparable to a traditional ground-processing pipeline when applied to both simulated and real Pleiades imagery.

Beyond image quality assessment, we evaluate the impact of such restoration on AI based remote sensing applications, specifically object detection. Then, we analyze the computational efficiency and architectural properties of EDSR in the context of onboard deployment, demonstrating its suitability under memory, latency, and power constraints typical of spaceborne systems.

The remainder of this paper is organized as follows. Section 2 reviews related work in satellite image restoration and learning-based approaches. Section 3 describes the proposed methodology, including the physics-based simulation framework, the datasets, and the EDSR based restoration model. Section 4 presents experimental results on simulated and real Pleiades imagery, as well as the evaluation on object detection tasks and embedded performance analysis. Section 5 discusses the operational implications of the proposed approach, and Section 6 concludes the paper with perspectives for future research.

## 2. Related Work

## 3. Material and methods

### 3.1. Datasets

Learning-based image restoration requires paired data, where a simulated degraded image is mapped to a simulated reference image serving as training reference. However, in operational satellite imaging, only raw-level images are available, and no true ground-truth level-1 (L1) image exists. Consequently, supervised training cannot rely on real raw/L1 pairs without biasing the model toward reproducing a specific processing chain. To overcome this limitation, we construct a physics-based simulated dataset that approximates realistic satellite imaging conditions while maintaining full control over degradation parameters. This simulated dataset serves as the sole source of supervision for training. We consider three datasets in this study: (i) a physics-based simulated dataset used for supervised training, (ii) real Pleiades imagery for operational validation, and (iii) Maxar imagery for applicative evaluation.

#### 3.1.1. Spatial Resolution and GSD Degradation

Spatial blur is modeled through a parametric Modulation Transfer Function (MTF) combining optical and sensor effects:

$$\text{MTF}(f_x, f_y) = \exp(-\gamma f_r) \cdot \text{sinc}(f_x) \cdot \text{sinc}(f_y), \quad (1)$$

where

$$f_r = \sqrt{f_x^2 + f_y^2}$$

and  $\gamma$  is calibrated from a Nyquist value  $\text{MTF}_{Nyq}$ . The corresponding Point Spread Function (PSF) is obtained as

$$\text{PSF} = \mathcal{F}^{-1}\{\text{MTF}\}, \quad (2)$$

and each spectral band is convolved with the PSF.

To simulate a coarser Ground Sampling Distance, the blurred image is then downsampled by a factor  $r$  (oversampling ratio). Prior to subsampling, an anti-aliasing Gaussian filter is applied. The image is then sampled every  $r$  pixels:

$$I_{\text{GSD}}(x, y) = I_{\text{blur}}(x_0 + rx, y_0 + ry), \quad (3)$$

where  $(x_0, y_0)$  denotes a centered offset ensuring no spatial shift. This procedure simulates both resolution loss and pixel footprint enlargement.

#### 3.1.2. Signal-Dependent Noise Model

Radiometric degradation is modeled using signal-dependent Gaussian noise:

$$\sigma^2(L) = \alpha L + \beta, \quad (4)$$

where  $L$  is the luminance. The parameters  $\alpha$  and  $\beta$  are estimated from two reference luminance–SNR pairs  $(L_0, \text{SNR}_0)$  and  $(L_1, \text{SNR}_1)$ :

$$\sigma_i^2 = \left( \frac{L_i}{\text{SNR}_i} \right)^2, \quad \sigma_i^2 = \alpha L_i + \beta. \quad (5)$$

Noise is sampled as

$$\sigma(L) = \sqrt{\max(0, \alpha L + \beta)}, \quad (6)$$

and added independently to each pixel.

By varying  $\text{MTF}_{Nyq}$ , the downsampling factor  $r$ , and the reference SNR values, multiple degradation levels are generated, enabling systematic evaluation under controlled and realistic imaging conditions.

#### 3.1.3. Simulated Dataset

The simulated dataset constitutes the core of our training framework. The ability of a learning-based restoration model to effectively learn and generalize strongly depends on the richness, diversity, and reliability of the training data. To this end, we design a simulated dataset that aims to be representative of the wide range of landscapes and imaging conditions encountered in operational satellite imagery.

The simulated dataset is constructed from high-resolution aerial RGB imagery obtained from OpenAerialMap [15], with an initial GSD of 10 cm. This resolution provides sufficient spatial margin to realistically simulate satellite acquisition conditions by degrading the data toward the targeted GSD, MTF, and SNR operating ranges. The dataset composition covers a wide variety of landscape types (urban, suburban, rural, agricultural, forested, coastal,

Table 1. Summary of simulated and real datasets used in this work.

Dataset	MTF	SNR @ $L_0 / L_1$	# Patches
Sim-Degraded-Variable	3~7%	$50 \pm 40 / 110 \pm 40$	196 128 / 36
Sim-Degraded-Fixed	7%	50 / 110	196 128 / 36
Sim-Reference-Fixed	25%	80 / 170	196 128 / 36
Real-Raw-Pleiades	~7%	~50 / ~110	- / 19

SNR values (in dB) are defined at reference luminance levels  $L_0 = 25$  and  $L_1 = 100 \text{ W/m}^2/\text{sr}/\mu\text{m}$ . Number of patches is reported as *Train* / *Test*. Training patches are  $128 \times 128$  pixels, while test patches are  $1500 \times 1500$  pixels.

and mountainous areas), with an emphasis on urban environments characterized by high spatial frequencies, sharp edges, and dense structural details. Such scenes are particularly challenging for image restoration and constitute a robust evaluation of the model’s ability to recover fine details without introducing artifacts such as ringing or oversmoothing.

To emulate realistic satellite acquisition conditions, aerial RGB images are converted to panchromatic imagery and degraded through a physics-based sensor simulation pipeline modeling the key characteristics of the target system (GSD, MTF and SNR). This process generates paired simulated degraded and simulated reference images under controlled and reproducible degradation settings.

The resulting images are produced at a target GSD of 50 cm and stored as 12-bit panchromatic TIFF patches. Several dataset configurations are generated (see Tab. 4.1). Fixed-degradation datasets correspond to the nominal operating point representative of chosen typical acquisition conditions, denoted as *Sim-Degraded-Fixed* / *Sim-Reference-Fixed* configuration in Tab. 4.1. In addition, variable-degradation datasets are created by sampling MTF and SNR values within realistic ranges around this nominal configuration, enabling the model to learn robustness to variations in imaging conditions.

The training set consists of  $128 \times 128$  pixel patches (196 128 samples), while testing is performed on  $1500 \times 1500$  pixel images (36 scenes) in order to evaluate both restoration fidelity and the impact of patch-based tiling during inference.

Experimental analysis indicates that training with multiple degradation levels (variable configuration) yields superior robustness compared to training at a single nominal operating point. Exposure to diverse MTF and SNR conditions improves generalization to unseen acquisition settings. Although fine-tuning on real data can further enhance performance when such data are available, all results reported in this work rely exclusively on models trained on simulated data to assess intrinsic sim-to-real generalization capability.

### 3.1.4. Real Imagery: Pleiades-HR

For real-world validation, we use optical panchromatic imagery acquired by the Pleiades satellite constellation oper-

ated by the French Space Agency (Centre National d’Études Spatiales, CNES). Raw-level images are used as input to the restoration algorithms to ensure a fair and unbiased comparison between learning-based and traditional approaches. Their main characteristics are reported in Tab. 4.1. Pleiades L1 products processed by the CNES ground segment are available and are used as a reference baseline for comparison. However, these L1 images are deliberately not employed as supervision targets during training. Using raw/L1 image pairs for supervised learning would bias the model toward reproducing the output of the traditional processing chain rather than learning an independent restoration mapping. Our objective is not to mimic the legacy ground-processing pipeline, but to assess whether a learning-based approach can directly infer restoration parameters that achieve or surpass the desired image quality. By avoiding supervision on L1 products, we prevent the model from inheriting design assumptions, artifacts, or limitations specific to the traditional method, and instead encourage it to learn optimal restoration behavior driven by physically realistic data. As no ground-truth reference exists for real satellite imagery, evaluation on Pleiades data relies on a combination of physical image quality indicators, visual inspection, and direct comparison with the traditional CNES processing chain.

## 3.2. Deep Learning Restoration

### 3.2.1. Traditional Restoration Pipeline

As a reference baseline, we consider the classical image restoration pipeline deployed by CNES for Pleiades-HR ground processing [7, 10]. This approach follows a sequential processing scheme, combining optical deconvolution to compensate sensor blur with NL-bayes denoising to reduce noise while preserving structural details [11]. The method relies on explicit physical modeling of the imaging system and requires accurate knowledge of sensor characteristics and acquisition parameters. While robust and well validated in operational contexts, the pipeline involves multiple processing stages, intermediate buffers, and iterative operations, resulting in high computational complexity and limited flexibility when faced with variations in imaging conditions. These characteristics make the approach difficult to adapt to onboard or real-time deployment.

### 3.2.2. Learning-Based Restoration Model: EDSR

We adopt the Enhanced Deep Super-Resolution (EDSR) network [13] as a light and non-generative learning-based restoration model. EDSR is a fully convolutional residual architecture composed of a sequence of residual blocks without batch normalization and followed by a reconstruction module, as shown in Fig.1. In the proposed configuration, the architecture is adapted to use a scale factor of 1, enabling exclusively image restoration without spatial res-



figures/edsr.png

Figure 1. EDSR architecture [13]

olution enhancement. The absence of adversarial training and generative components ensures stable optimization and avoids hallucination of artificial structures, which is critical for operational satellite imagery. The residual design facilitates learning of high-frequency corrections, and the absence of batch normalization helps preserve feature magnitude and overall image consistency. Training is performed in a fully supervised manner using simulated image pairs (*Sim-Degraded-Variable / Sim-Reference-Fixed*) only, to assess the intrinsic generalization capability of the model.

**TODO: ajouter ici un petit paragraphe pour l’aspect portabilité?**

### 3.3. Embedded Materials?

\*\*\*\*\* TODO \*\*\*\*\*

### 3.4. Metrics

Restoration performance is evaluated through two complementary perspectives: reference-based image similarity metrics and physically grounded sensor-level performance indicators, together with qualitative visual inspection.

#### 3.4.1. Full-Reference Image Quality Metrics

On simulated datasets, where simulated reference images are available, restoration performance is evaluated using classical full-reference image quality metrics, including Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). These metrics quantify pixel-level fidelity and structural consistency between restored images and their reference counterparts. Although widely adopted, these metrics may exhibit limited correlation with perceived visual quality, particularly in the presence of blur–noise trade-offs. To better capture perceptual and structural aspects of restoration, we additionally report learned perceptual metrics: LPIPS (Learned Perceptual Image Patch Similarity) and DISTS (Deep Image Structure and Texture Similarity). These feature-based metrics rely on deep neural rep-

resentations and have demonstrated improved sensitivity to texture reconstruction, edge preservation, and structural distortions, properties that are critical in high-resolution satellite imagery.

For real satellite data, where no reference image is available, full-reference metrics (PSNR, SSIM, LPIPS, and DISTS) cannot be computed.

#### 3.4.2. Physical Image Quality Metrics

Beyond similarity-based measures, we estimate physical image quality metrics directly related to imaging system performance. In particular, MTF and the SNR are measured on restored images and compared with their input counterparts.

The MTF is estimated using a slanted-edge method implemented in the *MTF Estimator* plugin for QGIS [8], computed across several multiple selected edge regions to obtain representative measurement. The SNR is evaluated over multiple homogeneous regions using the variance-based radiometric approach described in [14], following the signal-dependent noise model introduced in Sec. 3.1.2.

These indicators provide a physically interpretable assessment of blur compensation and noise amplification, ensuring consistency with operational imaging specifications. **For real satellite imagery, evaluation relies primarily on these sensor level indicators, complemented by qualitative visual inspection and comparison with the conventional processing chain.**

#### 3.4.3. Onboard benchmark

\*\*\*\*\* TODO \*\*\*\*\*

## 4. Experiments

### 4.1. Metric Sensitivity Analysis

To better interpret full reference image quality metrics (PSNR, SSIM, LPIPS, and DISTS), a dedicated analysis is conducted to evaluate their sensitivity to physical image quality metrics (MTF and SNR). This analysis is performed on 24 simulated datasets composed of 96 patches (1500 x 1500 pixels) generated at fixed SNR and MTF degradation levels around the nominal operating point, corresponding to the *Sim-Degraded-Fixed / Sim-Reference-Fixed* configuration (see ).

First, for each dataset, average MTF and SNR values are estimated and compared to their simulated reference values, confirming a strong consistency between measured and simulated sensor characteristics. Then, PSNR, SSIM, LPIPS, and DISTS are computed using the simulated reference images. Results indicate that classical and perceptual metrics exhibit markedly different sensitivities to blur and noise degradations. PSNR, LPIPS, and DISTS show a strong correlation with variations in optical blur, effectively capturing

changes in image sharpness as the MTF varies. SSIM appears less sensitive to blur variations. A major limitation is observed regarding noise sensitivity. Across all evaluated metrics, variations in SNR are only weakly reflected in metric values, indicating a limited ability of standard image quality metrics to capture noise-related degradations. This observation highlights the necessity of complementing conventional image similarity metrics with physically inspired indicators when evaluating satellite image restoration performance.

Table 2. Coefficient of determination ( $R^2$ ) for different metrics. mMTF and mSNR denote mean estimated values.

	mMTF	mSNR@L <sub>0</sub>	mSNR@L <sub>1</sub>
$R^2$	0.9765	0.7663	0.8204

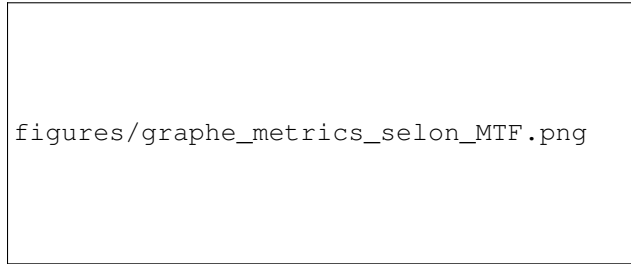


Figure 2. SSIM, LPIPS and DISTS metrics vs. MTF level at constant SNR.

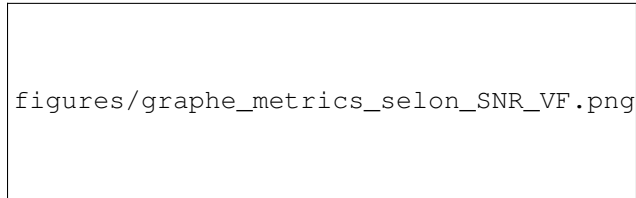


Figure 3. PSNR metric vs. SNR level at constant MTF.

## 4.2. Benchmark on Simulated Restoration Task

A comparative evaluation of the two restoration methods introduced in Sec. 3.2, is conducted on the simulated datasets. EDSR is configured with an upscaling factor of 1 and 16 residual blocks, focusing exclusively on image restoration. The model is trained in a fully supervised manner for 30 epochs on the simulated dataset which includes multiple degradation levels : *Sim-Degraded-Fixed* / *Sim-Reference-Fixed* (see Sec. 3.1.3).

Both approaches are evaluated on a fixed-degradation simulated test set (*Sim-Degraded-Fixed* / *Sim-Reference-Fixed*) using the metrics described in Sec. 3.4. As reported in Tab. 3, EDSR outperforms the traditional pipeline, with

Table 3. Quantitative performance comparison on simulated datasets.

Method	MTF@Nyq	SNR @ L <sub>0</sub> / L <sub>1</sub>	PSNR↑	SSIM↑	LPIPS↓	DISTS↓
End-to-End time						
Traditional	N/A	65.33 / 138.51	20.6	0.843	0.216	0.220
87.3 s						
EDSR	0.20	66.27 / 156.07	27.5	0.932	0.068	0.077
39.8 s						

MTF is averaged over 3 selected edge regions, SNR over 13 homogeneous patches (1500 × 1500 pixels), and full-reference image quality metrics over 36 patches (1500 × 1500 pixels). End-to-End time includes tiling, network inference, and reconstruction, measured on a 197 232 640 pixels sub-swath image.

a +6.9 dB PSNR gain, higher SSIM, and a markedly lower LPIPS score. The restored MTF reaches 0.20 while preserving comparable SNR levels, confirming effective sharpness blur compensation without significant noise amplification. MTF estimation is not reported for the traditional pipeline since the slanted-edge method did not converge reliably due to nonlinear ringing effects introduced by the deconvolution and denoising stages.

End-to-end processing time is measured to assess computational efficiency under realistic operational conditions. The traditional pipeline is inherently CPU bound and not designed for GPU acceleration, whereas EDSR naturally benefits from parallel execution on GPU architectures. Although the hardware platforms differ, each method is evaluated under a realist operational configuration, ensuring a fair comparison in terms of practical deployment feasibility. In our experiments, the traditional pipeline is executed on CPU, while EDSR is evaluated on an NVIDIA GeForce RTX 4090 GPU. Under these conditions, EDSR reduces the total end-to-end processing time, including tiling, network inference, and reconstruction, by nearly a factor of two compared to the traditional pipeline.

In addition to quantitative metrics, a qualitative visual analysis is performed, focusing on edge sharpness, texture preservation, residual noise, and the absence of non-physical artifacts. A synthetic resolution pattern embedded in simulated scenes (Fig. 4) provides a controlled and repeatable assessment of high-frequency reconstruction. Visual observations align with the quantitative results, confirming an effective balance between smoothing and sharpness restoration, with good robustness to noise.

## 4.3. Robustness to Blur Degradation

The robustness of the EDSR model to blur degradation is further investigated under increasingly severe imaging conditions. The model is trained on *Sim-Degraded-Variable* / *Sim-Reference-Fixed* datasets (see Tab. 4.1) and evaluated on simulated degraded images generated at fixed MTF levels. Three blur levels are considered: low, medium, and high, corresponding to MTF<sub>7%</sub>, MTF<sub>5%</sub>, and MTF<sub>3%</sub>, respectively. For each input image, the MTF is estimated both

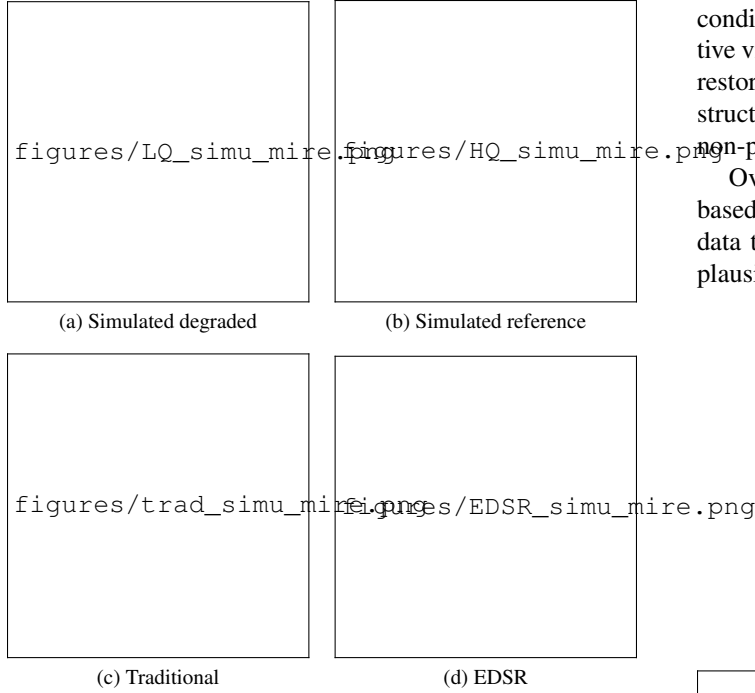


Figure 4. Image restoration on simulated degraded image with embedded mire.

Table 4. MTF before and after restoration with EDSR for different blur levels on simulated degraded data. **TODO ajouter les valeur HQ/LQ ref**

Deg. Level	MTF(Degraded)	MTF(Restored)	$\Delta$ MTF
MTF <sub>7%</sub>	0.0741	0.2017	+0.1276
MTF <sub>5%</sub>	0.0558	0.1555	+0.0997
MTF <sub>3%</sub>	0.0336	0.1383	+0.1047

before restoration (on degraded input) and after restoration using EDSR. Results show that the restored MTF is systematically higher than the input MTF across all degradation levels. Importantly, the gain in MTF remains relatively stable despite increasing blur severity. Even for the most degraded input (MTF<sub>3%</sub>), EDSR provides a significant improvement in spatial resolution.

These results indicate that EDSR trained on variable degradation conditions does not overfit to a single nominal operating point, but instead learns a robust restoration behavior that generalizes to unseen blur levels. This robustness is a key property for operational satellite imagery, where imaging conditions may deviate from nominal specifications due to acquisition geometry, temporal variations, or sensor aging.

#### 4.4. Experiments on Real Pleiades Data

The trained model is further evaluated on real Pleiades images to assess its behavior under operational acquisition

conditions. Evaluation is performed only through qualitative visual inspection and comparison with the conventional restoration pipeline. Visual inspection indicates clearer structural details, improved edge definition, and no evident non-physical artifacts or excessive ringing.

Overall, these observations indicate that the learning-based restoration model generalizes well from simulated data to real Pleiades imagery, while preserving physically plausible image characteristics.

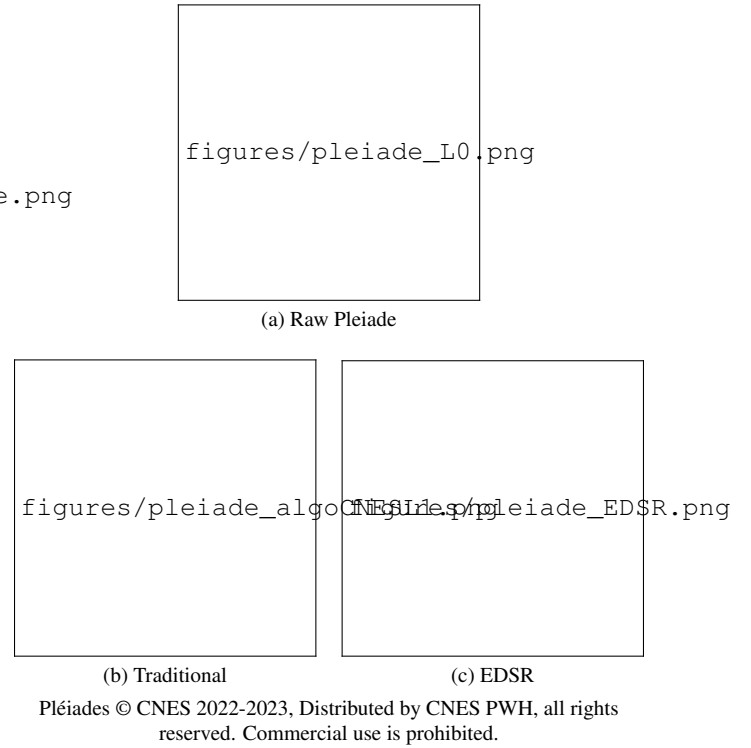


Figure 5. Comparison of raw Pleiades image and super-resolved results using the traditional algorithm and EDSR.

#### 4.5. Impact on Application Scenario: Object Detection on DIOR dataset

To demonstrate the versatility of EDSR, we conducted experiments on DIOR dataset [12]. This well known dataset in the Earth observation community is composed of 23 463 optical remote sensing images covering 20 classes. An object detection task is performed using the YOLOv11n model [9]. The nano version is chosen to ensure maximum performance on edge devices. Moreover, the images are down-scale to 256 x 256 pixels during training and inference to maximise inference speed.

The original images are degraded using the simulation described in Sec.3.1. The parameters used for the simulation are described in Tab.5. From there, degraded images are restored using our learning-based restoration as described in Sec.3.2.2. Thus, two datasets are available

Table 5. Summary of simulated **and real** DIOR datasets used in this work.

Dataset	MTF	SNR @ $L_0 / L_1$	#Patches
Deg-Med	1–2%	$50 \pm 5 / 60 \pm 5$	18 000 / 2000
Deg-High	0.1–0.5%	$40 \pm 5 / 50 \pm 5$	18 000 / 2000

Number of patches is reported as *Train / Test*. Patches are  $800 \times 800$  pixels.

for each type of degradation: a degraded dataset and its corresponding EDSR restored version, that emulates the characteristics of the original DIOR dataset. Two degradation settings are defined: *Deg-Med* (medium) and *Deg-High* (high). While *Deg-Med* is similar to a degradation that could be seen with a satellite sensor, we decided to push the degradation further to test the restoration capabilities of our learning-based restoration model specifically for object detection use case. In addition, performance on the simulated datasets are compared with the performance on original DIOR dataset. This constitutes our baseline for our experiments.

Tab.6 displays the results for all three datasets. First, we can observe that the model achieve a strong baseline performance, with 75.90% mAP@50 and 54.10% mAP@90. These values shows that the training process is under control, without major flaws.

Regarding the *Deg-Med* experiment, we see that the degradation impacted, even lightly so, the performance of the detection model, with a decrease of 2.6% for both mAP@50 and mAP@90. However, the restore version *Restored images* almost fully recovered the original performance.

Regarding the *Deg-High* experiment, the impact of the degradation is stronger. The raw images suffer a pronounced performance drop, especially in mAP@50 (–8.60%) and precision (–22.80%). It indicates that detection is significantly compromised under such heavy degradation. In contrast, the restored images dramatically recover performance, bringing mAP@50 (72.37%) and precision (85.03%) closer to reference levels.

Overall, the results shows that even when severe degradations heavily impact the detector performance, using our proposed light restoration model improves performance significantly.

#### 4.6. Embedded Deployment

– Performance speed **TODO** –

### 5. Discussions

From an operational perspective, the results highlight the relevance of lightweight learning-based restoration models for both ground-based and onboard processing. Although traditional restoration pipelines remain robust, their sequen-

Table 6. Object detection comparison between degraded and restored DIOR datasets with Yolo11n on  $256 \times 256$  pixels images.

Dataset	Set	mAP50	mAP90	Prec.	Rec.
Original	N/A	75.90%	54.10%	86.60%	69.30%
Deg-Med	Rest.	75.52%	53.88%	87.47%	68.32%
	Deg.	73.33%	51.5%	85.00%	66.93%
Deg-High	Rest.	72.37%	50.57%	85.03%	66.36%
	Deg.	67.30%	46.29%	63.80%	60.70%

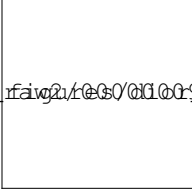
Degraded	Reference	EDSR
		
		
		

Figure 6. Qualitative comparison of restored RGB images with EDSR

tial structure and computational cost limit their suitability for resource-constrained environments. In contrast, the proposed EDSR based approach achieves comparable image quality, while significantly reducing processing time, making it better aligned with near-real-time operational constraints.

Beyond computational efficiency, the robustness of EDSR to variations in blur and noise levels constitutes a key advantage in practical scenarios, where imaging conditions may deviate from nominal specifications due to acquisition geometry, temporal effects, or sensor aging. These properties suggests that learning-based restoration can serve as a viable alternative to traditional pipelines, not only as a ground-based acceleration strategy but also as a candidate for onboard preprocessing, enabling earlier and more flexible exploitation of satellite imagery.

Finally, when coupled with onboard object detection tasks, EDSR revealed to significantly improve the performance of the AI model. In these application-driven scenarios, exact reconstruction of the ground truth is not strictly required, in contrast to traditional ground-based restoration objectives. Instead, the objective is to enhance task-relevant features while maintaining physical plausibility. The observed improvements therefore reinforce the suitability of learning-based restoration as a task-oriented preprocessing step for onboard AI applications in remote sensing.

## 6. Conclusion

In this paper, a series of experiments was conducted to evaluate the relevance of learning-based methods for onboard preprocessing of raw satellite imagery. A customized version of the Enhanced Deep Super-Resolution (EDSR) architecture was adapted for image restoration. An extensive dataset was constructed, comprising both real Raw / CNES-L1 pairs and simulated degraded/reference pairs.

The proposed model was first evaluated using full-reference image quality metrics (SSIM, PSNR, LPIPS, and DISTS) and physical image quality metrics (SNR, MTF). Results on simulated and real Raw / CNES-L1 pairs indicate that EDSR-restored images exhibit characteristics comparable to those obtained through traditional restoration pipelines. When evaluated under varying blur degradation levels, EDSR demonstrated strong robustness, yielding stable restored MTF values across degradation scenarios.

Furthermore, EDSR was assessed as a lightweight preprocessing module for AI-based object detection. Experimental results show that detection performance consistently improves when restoration is applied prior to inference, confirming the benefit of restoration for onboard AI tasks.

In future work, the EDSR architecture could be more specialized with remote sensing imagery. Indeed, because remote sensing data differ from natural images, incorporating physical priors into the restoration framework [1] may enhance performance while preventing the emergence of non-physical geometric artifacts.

Another promising direction would be to integrate restoration directly within a task-dependent AI model. Such an approach could simplify the onboard processing pipeline by removing the need of having two independent models.

Finally, because EDSR is natively a super-resolution model, it could be interesting to evaluate the impact of super-resolution on onboard AI-tasks.

## References

- [1] Chayan Banerjee, Kien Nguyen, Clinton Fookes, and George Karniadakis. Physics-informed computer vision: A review and perspectives, 2024. 8
- [2] Maud Biquard. *Régularisation avec des réseaux de neurones pour la restauration d'images satellites*. PhD thesis, 2025. 556
- [3] Maud Biquard, Marie Chabert, Florence Genin, Christophe Latry, and Thomas Oberlin. Régularisation par modèles pseudo-génératifs pour la restauration d'images. In *Colloque Gretsi 2023*, Grenoble, France, 2023. Association Gretsi. 557
- [4] Maud Biquard, Marie Chabert, Florence Genin, Christophe Latry, and Thomas Oberlin. Deep priors for satellite image restoration with accurate uncertainties. *IEEE Transactions on Geoscience and Remote Sensing*, 63:1–16, 2025. 558
- [5] Mikaël Carlván. *Optimization of the compression/restoration chain for satellite images*. Theses, Université Nice Sophia Antipolis, 2013. 559
- [6] F. Chouteau, Laurent Gabet, Renaud Fraisse, T. Bonfort, Bouchra Harnoufi, V. Greiner, M. Goff, M. Ortner, and V. Paveau. Joint super-resolution and image restoration for pléiades neo imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B1-2022:9–15, 2022. 560
- [7] Sebastien Fourest Christophe Latry, Gwendoline Blancher. Chaîne de fusion p+xs des images pléiades-hr, 2013. 561
- [8] Jorge Gil. Fidmtf: An mtf estimator (qgis plugin). [https://plugins.qgis.org/plugins/mtf\\_estimator/](https://plugins.qgis.org/plugins/mtf_estimator/). 562
- [9] Rahima Khanam and Muhammad Hussain. Yolov11: An overview of the key architectural enhancements, 2024. 563
- [10] C. Latry, S. Fourest, and C. Thiebaut. Restoration technique for pleiades-hr panchromatic images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B1:555–560, 2012. 564
- [11] Marc Lebrun, Antoni Buades, and Jean-Michel Morel. Implementation of the "Non-Local Bayes" (NL-Bayes) Image Denoising Algorithm. *Image Processing On Line*, 3:1–42, 2013. <https://doi.org/10.5201/ipol.2013.16.1.3>. 565
- [12] Ke Li, Gang Wan, Gong Cheng, Liqu Meng, and Junwei Han. Object detection in optical remote sensing images: A survey and a new benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159:296–307, 2020. 566
- [13] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. *CoRR*, abs/1707.02921, 2017. 2, 3, 567
- [14] J.G. Liu, L.F. Zhang, and Q.X. Tong. Estimation of signal to noise ratio of remote sensing images. *Journal of Basic Science and Engineering*, 4:360–365, 1999. 4 568
- [15] OpenAerialMap. Openaerialmap. <https://openaerialmap.org.2> 569