

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

Anonymous CVPR submission

Paper ID *****

Abstract

Satellite image restoration aims to improve image quality by compensating degradations (e.g., noise and blur) introduced by the imaging system and acquisition conditions. As a fundamental preprocessing step, restoration directly impacts both ground-based product generation and emerging onboard AI applications. Traditional restoration pipelines are based on sequential physical models computationally and memory demanding involving multiple operations and intermediate buffers, and poorly suited to resource constrained onboard environments. We investigate whether a light and non-generative residual convolutional network (EDSR), trained exclusively on physics based simulated satellite data, can match or surpass a traditional ground-processing restoration pipeline across multiple operating conditions while meeting the efficiency requirements of onboard deployment. A realistic simulation framework modeling ground sampling distance (GSD), modulation transfer function (MTF), and signal dependent noise is used to generate supervised training pairs without relying on unavailable raw/L1 ground-truth data. By sampling degradations across a wide range of operating conditions, the model learns robustness to variability in blur and noise. Experiments conducted on simulated datasets, real Pleiades imagery, and object detection tasks demonstrate that the proposed approach achieve competitive performance with respect to the traditional pipeline in image quality while significantly reducing end-to-end processing time. Physical indicators confirm robust blur compensation across varying degradation levels. Moreover, restoration consistently improves downstream detection performance, highlighting its relevance for onboard AI pipelines. Finally, we demonstrate its effective deployment on embedded hardware, validating its onboard applicability.

1. Introduction

034

High-resolution optical satellite imagery is acquired onboard at Raw level, where sensor optical limitations and acquisition conditions, including modulation transfer function effects and noise, inherently degrade image quality. Image restoration aims to compensate for these degradations by jointly addressing optical blur, sensor noise, and acquisition-related effects. It therefore constitutes a prerequisite pre-processing step for any higher-level exploitation of satellite imagery. This is true not only for traditional ground-based workflows but also for emerging onboard processing pipelines, where early enhancement of image quality 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 established, such pipelines, historically deployed on ground infrastructures, are computationally demanding, often involving iterative optimization procedures and multiple intermediate 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 faithfully 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, including disaster response [Mini biblio TODO], maritime surveillance[Mini biblio TODO], and autonomous onboard decision-making, motivates a shift toward more efficient processing strategies. In this context, recent learning-based approaches enable joint optimization of blur compensation 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
072

In this work, we investigate whether a light non-

073 generative residual convolutional neural network, EDSR
 074 [13], trained exclusively on physically realistic simulated
 075 satellite data, can achieve restoration performance compa-
 076 rable to a traditional ground-processing pipeline when ap-
 077 plied to both simulated and real Pleiades imagery.

078 Beyond image quality assessment, we evaluate the im-
 079 pact of such restoration on AI based remote sensing ap-
 080 plications, specifically object detection. Then, we analyze
 081 the computational efficiency and architectural properties of
 082 EDSR in the context of onboard deployment, demonstrating
 083 its suitability under memory, latency, and power constraints
 084 typical of spaceborne systems.

085 The remainder of this paper is organized as follows. Sec-
 086 tion 2 reviews related work in satellite image restoration
 087 and learning-based approaches. Section 3 describes the pro-
 088 posed methodology, including the physics-based simulation
 089 framework, the datasets, and the EDSR based restoration
 090 model. Section 4 presents experimental results on simu-
 091 lated and real Pleiades imagery, as well as the evaluation
 092 on object detection tasks and embedded performance anal-
 093 ysis. Section 5 discusses the operational implications of the
 094 proposed approach, and Section 6 concludes the paper with
 095 perspectives for future research.

096 2. Related Work

097 3. Material and methods

098 3.1. Datasets

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

115 3.1.1. Spatial Resolution and GSD Degradation

116 Spatial blur is modeled through a parametric Modulation
 117 Transfer Function (MTF) combining optical and sensor ef-
 118 fects:

$$119 MTF(f_x, f_y) = \exp(-\gamma f_r) \cdot \text{sinc}(f_x) \cdot \text{sinc}(f_y), \quad (1)$$

120 where

$$121 f_r = \sqrt{f_x^2 + f_y^2}$$

and γ is calibrated from a Nyquist value MTF_{Nyq} . The
 corresponding Point Spread Function (PSF) is obtained as

$$122 \text{PSF} = \mathcal{F}^{-1}\{\text{MTF}\}, \quad (2) \quad 124$$

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 (oversam-
 pling ratio). Prior to subsampling, an anti-aliasing Gaussian
 filter is applied. The image is then sampled every r pixels:

$$125 I_{\text{GSD}}(x, y) = I_{\text{blur}}(x_0 + rx, y_0 + ry), \quad (3) \quad 130$$

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

131 3.1.2. Signal-Dependent Noise Model

Radiometric degradation is modeled using signal-dependent
 135 Gaussian noise:

$$136 \sigma^2(L) = \alpha L + \beta, \quad (4) \quad 137$$

where L is the luminance. The parameters α and β
 are estimated from two reference luminance-SNR pairs
 (L_0, SNR_0) and (L_1, SNR_1) :

$$138 \sigma_i^2 = \left(\frac{L_i}{\text{SNR}_i} \right)^2, \quad \sigma_i^2 = \alpha L_i + \beta. \quad (5) \quad 141$$

Noise is sampled as

$$142 \sigma(L) = \sqrt{\max(0, \alpha L + \beta)}, \quad (6) \quad 143$$

and added independently to each pixel.

By varying 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.

149 3.1.3. Simulated Dataset

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

The simulated dataset is constructed from high-
 155 resolution aerial RGB imagery obtained from Open-
 156 AerialMap [15], with an initial GSD of 10 cm. This reso-
 157 lution provides sufficient spatial margin to realistically sim-
 158 ulate satellite acquisition conditions by degrading the data
 159 toward the targeted GSD, MTF, and SNR operating ranges.
 The dataset composition covers a wide variety of landscape
 160 types (urban, suburban, rural, agricultural, forested, coastal,
 161 162 163 164

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 ± 40 / 110 ± 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 128×128 pixels, while test patches are 1500×1500 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 128×128 pixel patches (196 128 samples), while testing is performed on 1500×1500 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 Imageries: 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-



Figure 1. EDSR architecture [13]

258 olution enhancement. The absence of adversarial training
 259 and generative components ensures stable optimization and
 260 avoids hallucination of artificial structures, which is criti-
 261 cal for operational satellite imagery. The residual design
 262 facilitates learning of high-frequency corrections, and the
 263 absence of batch normalization helps preserve feature mag-
 264 nitude and overall image consistency. Training is performed
 265 in a fully supervised manner using simulated image pairs
 266 (*Sim-Degraded-Variable / Sim-Reference-Fixed*) only, to as-
 267 sess the intrinsic generalization capability of the model.

268 **TODO: ajouter ici un petit paragraphe pour l'aspect**
 269 **portabilité?**

270 3.3. Embedded Materials?

271 ***** TODO *****

272 3.4. Metrics

273 Restoration performance is evaluated through two com-
 274 plementary perspectives: reference-based image similarity
 275 metrics and physically grounded sensor-level performance
 276 indicators, together with qualitative visual inspection.

277 3.4.1. Full-Reference Image Quality Metrics

278 On simulated datasets, where simulated reference images
 279 are available, restoration performance is evaluated using
 280 classical full-reference image quality metrics, including
 281 Peak Signal-to-Noise Ratio (PSNR) and Structural Simi-
 282 larity Index (SSIM). These metrics quantify pixel-level fi-
 283 delity and structural consistency between restored images
 284 and their reference counterparts. Although widely adopted,
 285 these metrics may exhibit limited correlation with perceived
 286 visual quality, particularly in the presence of blur–noise
 287 trade-offs. To better capture perceptual and structural as-
 288 pects of restoration, we additionally report learned percep-
 289 tual metrics: LPIPS (Learned Perceptual Image Patch Simi-
 290 larity) and DISTS (Deep Image Structure and Texture Simi-
 291 larity). These feature-based metrics rely on deep neural rep-

292 resentations and have demonstrated improved sensitivity to
 293 texture reconstruction, edge preservation, and structural dis-
 294 tortions, properties that are critical in high-resolution satel-
 295 lite imagery.

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

299 3.4.2. Physical Image Quality Metrics

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

305 The MTF is estimated using a slanted-edge method im-
 306 plemented in the *MTF Estimator* plugin for QGIS [8], com-
 307 puted across several multiple selected edge regions to obtain
 308 representative measurement. The SNR is evaluated over
 309 multiple homogeneous regions using the variance-based ra-
 310 diometric approach described in [14], following the signal-
 311 dependent noise model introduced in Sec. 3.1.2.

312 These indicators provide a physically interpretable as-
 313 sessment of blur compensation and noise amplification, en-
 314 suring consistency with operational imaging specifications.
For real satellite imagery, evaluation relies primarily on
315 these sensor level indicators, complemented by qualita-
316 tive visual inspection and comparison with the conven-
317 tional processing chain.

318 3.4.3. Onboard benchmark

319 *****TODO*****

320 4. Experiments

321 4.1. Metric Sensitivity Analysis

322 To better interpret full reference image quality metrics
 323 (PSNR, SSIM, LPIPS, and DISTS), a dedicated analysis
 324 is conducted to evaluate their sensitivity to physical image
 325 quality metrics (MTF and SNR). This analysis is performed
 326 on 24 simulated datasets composed of 96 patches (1500 x
 327 1500 pixels) generated at fixed SNR and MTF degradations
 328 levels around the nominal operating point, corresponding to
 329 the *Sim-Degraded-Fixed / Sim-Reference-Fixed* configura-
 330 tion (see).

331 First, for each dataset, average MTF and SNR values are
 332 estimated and compared to their simulated reference values,
 333 confirming a strong consistency between measured and sim-
 334 ulated sensor characteristics. Then, PSNR, SSIM, LPIPS,
 335 and DISTS are computed using the simulated reference im-
 336 ages. Results indicate that classical and perceptual met-
 337 rics exhibit markedly different sensitivities to blur and noise
 338 degradations. PSNR, LPIPS, and DISTS show a strong cor-
 339 relation with variations in optical blur, effectively capturing

341 changes in image sharpness as the MTF varies. SSIM ap-
 342 pears less sensitive to blur variations. A major limitation
 343 is observed regarding noise sensitivity. Across all eval-
 344 uated metrics, variations in SNR are only weakly reflected in
 345 metric values, indicating a limited ability of standard image
 346 quality metrics to capture noise-related degradations. This
 347 observation highlights the necessity of complementing con-
 348 ventional image similarity metrics with physically inspired
 349 indicators when evaluating satellite image restoration per-
 350 formance.

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

	mMTF	mSNR@L ₀	mSNR@L ₁
R^2	0.9765	0.7663	0.8204

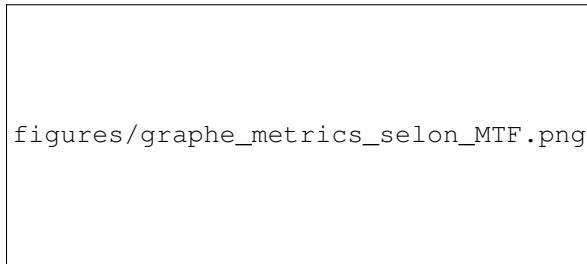


Figure 2. SSIM, LPIPS and DISTs metrics vs. MTF level at constant SNR.

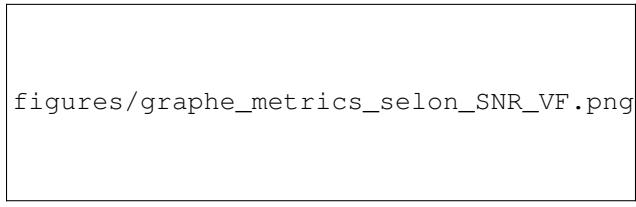


Figure 3. PSNR metric vs. SNR level at constant MTF.

351 4.2. Benchmark on Simulated Restoration Task

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

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

Table 3. Quantitative performance comparison on simulated datasets.

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

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.

395 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_{7%}, MTF_{5%}, and MTF_{3%}, respectively. For each input image, the MTF is estimated both

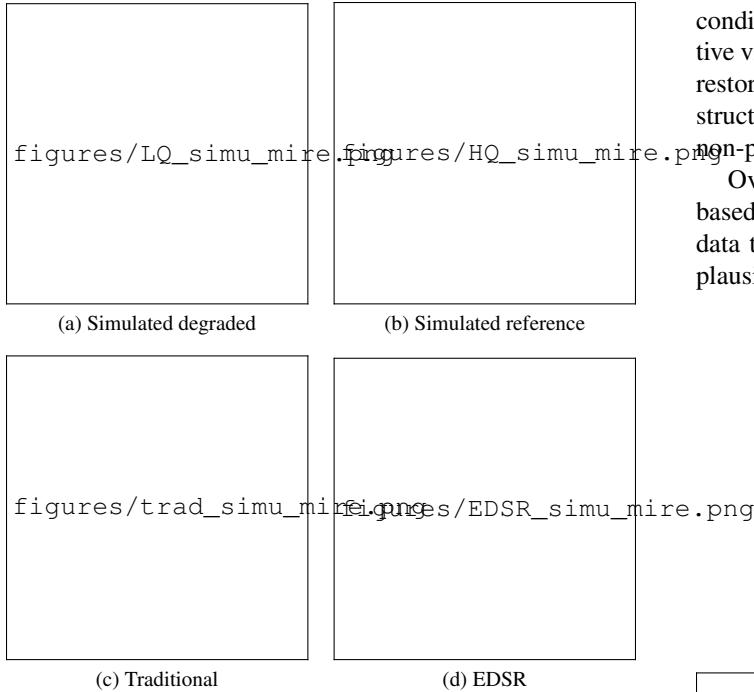


Figure 4. Image restoration on simulated degraded image with embedded mire.

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)	Δ MTF
MTF _{7%}	0.0741	0.2017	+0.1276
MTF _{5%}	0.0558	0.1555	+0.0997
MTF _{3%}	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 ($MTF_{3\%}$), EDSR provides a significant improvement in spatial resolution.

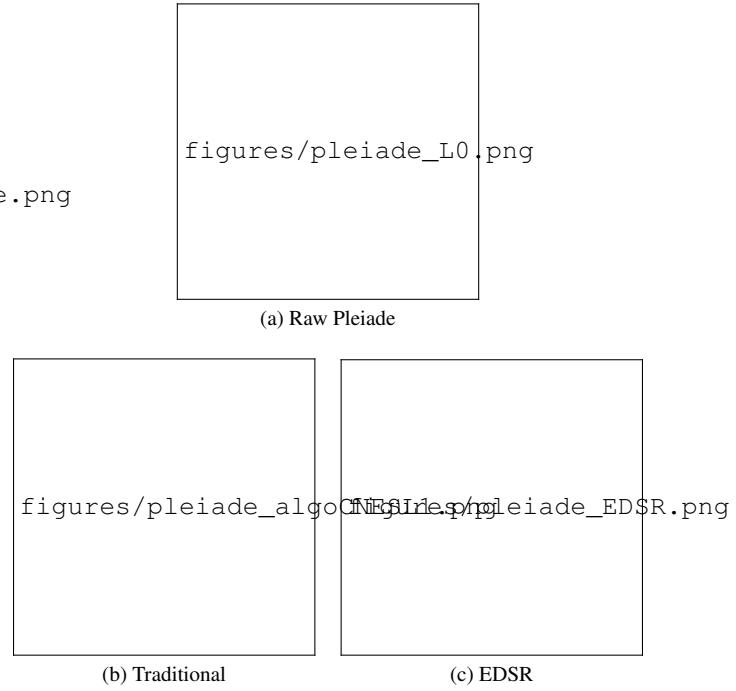
These results indicate that EDSR trained on variable degradation conditions does not overfit to a single nominal operating point, but instead learns a robust restoration behavior that generalizes to unseen blur levels. This robustness is a key property for operational satellite imagery, where imaging 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.



Pléiades © CNES 2022-2023, Distributed by CNES PWH, all rights reserved. Commercial use is prohibited.

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 800×800 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 x 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%

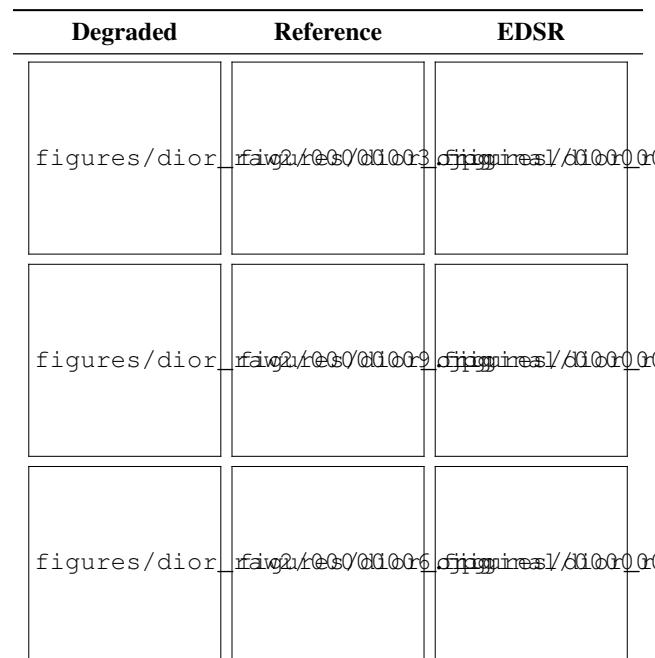


Figure 6. Qualitative comparison of restored RGB images with EDSR

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.

489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504

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 DIST) 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 pre-processing 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 Kaniadakis. 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.

- Thèse de doctorat dirigée par Oberlin, Thomas et Chabert, Marie Informatique et Télécommunications Toulouse, ISAE 2025. 1
- [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.
 - [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. 1
 - [5] Mikael Carlavan. *Optimization of the compression/restoration chain for satellite images*. Theses, Université Nice Sophia Antipolis, 2013. 1
 - [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. 1
 - [7] Sébastien Fourest Christophe Latry, Gwendoline Blancher. Chaîne de fusion p+xs des images pléiades-hr, 2013. 3
 - [8] Jorge Gil. Fidmtf: An mtf estimator (qgis plugin). https://plugins.qgis.org/plugins/mtf_estimator/. 4
 - [9] Rahima Khanam and Muhammad Hussain. Yolov11: An overview of the key architectural enhancements, 2024. 6
 - [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. 1, 3
 - [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>
 - [12] Ke Li, Gang Wan, Gong Cheng, Lijiu 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. 6
 - [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, 4
 - [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
 - [15] OpenAerialMap. Openaerialmap. <https://openaerialmap.org>. 2