# Rethinking Satellite Image Restoration for Onboard AI: A Lightweight Learning-Based Approach

Anonymous CVPR submission

Paper ID 11

## Abstract

*Satellite image restoration aims to improve image quality by compensating for 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 based on sequential physical models are computationally intensive and slow, making them unsuitable for onboard environments. In this paper, we introduce ConvBEERS: a Convolutional Board-ready Embedded and Efficient Restoration model for Space to investigate whether a light and non-generative residual convolutional network, trained on simulated satellite data, can match or surpass a traditional ground-processing restoration pipeline across multiple operating conditions.*

*Experiments conducted on simulated datasets and real Pleiades-HR imagery demonstrate that the proposed approach achieves competitive image quality, with a +6.9dB PSNR improvement. Evaluation on a downstream object detection task demonstrates that restoration significantly improves performance, with up to +5.1% mAP@50. In addition, successful deployment on a Xilinx Versal VCK190 FPGA validates its practical feasibility for satellite onboard processing, with a $\sim$41$\times$ reduction in latency compared to the traditional pipeline. These results demonstrate the relevance of using lightweight CNNs to achieve competitive restoration quality while addressing real-world constraints in spaceborne systems.*

## 1. Introduction

High-resolution optical satellite imagery is acquired at the raw level onboard, where sensor optical limitations and acquisition conditions, including Modulation Transfer Function (MTF) effects and Signal-to-Noise Ratio (SNR), inherently impact image quality. Image restoration aims to improve image quality by jointly addressing optical blur, sensor noise, and acquisition-related effects. It therefore constitutes a prerequisite preprocessing 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 image quality enhancement can directly benefit downstream analytical tasks.

Traditional restoration approaches are typically based on sequential physical models that combine deconvolution and denoising stages [9, 29, 30]. While robust and well-established, these ground-based pipelines are computationally demanding, with iterative optimisation and multiple buffers. This results in significant memory usage and processing latency, which severely limit their suitability for onboard implementation.

The increasing demand for low-latency applications, including disaster response [15, 20], maritime surveillance [26, 37] and autonomous onboard decision-making [17], motivates a shift toward more efficient processing strategies. In this context, recent studies evaluate the feasibility of learning-based approaches for remote sensing applications [3, 4, 6, 11], offering a promising alternative for ground and board processing. However, limited access to raw imagery often limits their use for onboard applications [14, 36].

In this work, we investigate whether a light, non-generative residual convolutional neural network trained exclusively on physically realistic simulated satellite data can achieve restoration performance comparable to that of a traditional ground-processing pipeline when applied to both simulated and real imagery. To do so, we introduce ConvBEERS, a Convolutional Bord-ready Embedded and Efficient Restoration model for Space, trained. Our approach proposes a novel training strategy on physics-driven simulated images. Beyond image quality assessment, we also evaluate the impact of such restoration on AI-based remote sensing applications, specifically object detection. Finally, we analyse the onboard performance of our architecture by deploying it on a Xilinx Versal VCK190 FPGA, demonstrating its suitability under the memory, latency, and power constraints typical of spaceborne systems.

## 2. Background and related work

Remote sensing products are distributed according to *processing levels*, which reflect the degree of radiometric and geometric correction applied to the raw sensor measurements. The first available product is generally a Level-1 image, obtained after radiometric calibration, geometric correction, co-registration, and orthorectification of the raw acquisition [13]. Here, we concentrate on radiometric enhancement.

**Traditional restoration.** Classical radiometric restoration pipelines typically combine deconvolution and denoising stages. In very-high-resolution optical systems such as Pléiades, restoration relies on an instrument-driven deconvolution step followed by advanced denoising methods such as NL-Bayes [9, 29, 30]. These methods rely on explicit physical modelling of the imaging system (system point spread function (PSF) and noise statistics), enabling improved sharpness while controlling noise amplification. Although operationally validated [12, 29], such pipelines involve sequential processing steps and iterative operations. Consequently, they are primarily suited for ground-based processing and remain difficult to adapt to onboard or real-time deployment. Moreover, because they are tailored to specific optical system characteristics, generalisation to other systems is challenging.

**AI-based restoration.** In parallel, learning-based approaches have emerged for raw image restoration [5, 6, 11]. CNN-based architectures are commonly employed due to their efficient convolutional operations and ability to model local spatial structures [21, 45]. Their residual layers facilitate stable optimisation while maintaining moderate computational complexity. Generative strategies have also been explored. GAN-based methods optimise perceptual and adversarial objectives to enhance visual sharpness [31], but may synthesise high-frequency components and are prone to creating artefacts [7]. In remote sensing applications, where geometric accuracy and radiometric consistency are critical, such hallucinated structures can affect quantitative reliability. Transformer-based models and diffusion models have demonstrated competitive performance in image restoration by capturing long-range dependencies through self-attention mechanisms [10, 44, 50], but rely on large backbone architectures and computationally intensive inference procedures, leading to significant computational and memory demands that are incompatible with resource-constrained onboard deployment.

**Onboard restoration in remote sensing.** Ongoing work studies the feasibility of light-based restoration models [39], but despite the rapid progress of AI-based restoration methods, onboard image processing pipelines in remote sensing still largely replicate traditional ground-based approaches. A few experimental demonstrations have explored learning-based enhancement in orbit. Notably, the OPS-SAT mission reported a generative AI-based image denoising experiment deployed on board, demonstrating the feasibility of neural restoration under real-space constraints [28]. More recently, scalable neural pushbroom architectures for real-time denoising have been implemented on NVIDIA Jetson-class embedded hardware, highlighting the potential of lightweight CNN designs for edge inference in space environments [46]. Other recent approaches, such as PyRaws [41], primarily focus on restoring the geometric aspects of raw satellite data through coarse-coregistration.

**Positioning of this work.** Previous work have demonstrated the feasibility of deploying learning-based image enhancement onboard satellites. However, these works typically focus either on specific enhancement tasks or on proof-of-concept deployment, without systematic comparison to operational restoration pipelines or evaluation within a complete application chain. In contrast, this work investigates whether a lightweight, non-generative CNN trained exclusively on physics-based simulated data can match the performance of traditional model-based restoration, while satisfying FPGA-onboard constraints. Beyond image quality assessment, we evaluate its impact on an end-to-end use case, thus positioning restoration as a task-oriented preprocessing module for operational spaceborne AI systems.

## 3. Proposed method

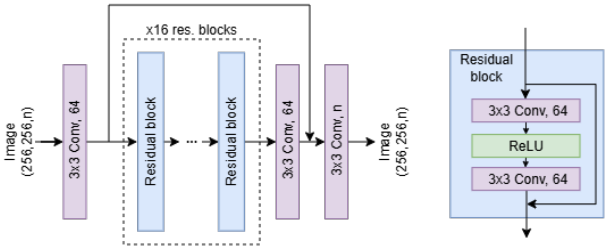



Figure 1. ConvBEERS architecture. It is composed of 3x3 size-preserving convolution layers (stride=1, padding=1) and 16 residual blocks without batch normalisation for a total of 1.2M trainable parameters.

### 3.1. A light residual network for restoration

Residual connections in convolutional neural networks greatly improve the flow of gradient, and allow for deeper architectures [22]. This was later successfully adapted into super-resolution models [31, 34]. Image super-resolution can be formulated as an inverse image restoration problem where the degradation operator consists of blurring and downsampling [47, 48].

With that in mind, we built upon the Enhanced Deep

Super-Resolution (EDSR) design [34], which uses residual layers without batch normalisation for the restoration task.

We propose ConvBEERS, a Convolutional Board-ready Embedded and Efficient Restoration model for Space, as a light, non-generative learning-based restoration technique. ConvBEERS is a fully convolutional architecture composed of a sequence of 16 residual blocks without batch normalisation and followed by a reconstruction layer, as shown in Fig.1. The absence of adversarial training and generative components ensures stable optimisation and avoids hallucination of artificial structures, which is critical for operational satellite imagery. The residual design encourages the model to learn how to correct the degradations with respect to the input image, and the absence of batch normalisation 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 generalisation capability of the model.

In addition, ConvBEERS exhibits strong portability properties for onboard deployment. As a fully convolutional network, it is resolution-agnostic and can process large images through tiling without architectural changes. The absence of batch normalisation simplifies inference and facilitates quantised or fixed-point implementations on embedded hardware. Moreover, the model does not require scene-dependent parameter tuning at inference time, enabling straightforward integration across sensors and evolving operating conditions.

### 3.2. Training objective

To balance spatial fidelity, perceptual realism, and spectral consistency, we employ a composite loss function combining pixel-domain, perceptual, and frequency-domain supervision:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{L1} + \lambda_p \mathcal{L}_{LPIPS} + \lambda_f \mathcal{L}_{FFT}, \tag{1}$$

where $\lambda_1$, $\lambda_p$, and $\lambda_f$ denote weighting coefficients. In this work, experiments were performed with $\lambda_1 = 1.0$, $\lambda_p = 0.5$ and $\lambda_f = 0.1$

**Pixel reconstruction loss..** We adopt the $\ell_1$ loss to enforce pixel-wise consistency:

$$\mathcal{L}_{L1} = \|x - \hat{x}\|_1, \tag{2}$$

which has been shown to better preserve sharp structures than $\ell_2$ in super-resolution and restoration tasks [34].

**Perceptual loss..** To encourage visually meaningful reconstructions, we incorporate LPIPS [25, 49], which measures similarity in deep feature space:

$$\mathcal{L}_{LPIPS} = \sum_l w_l \|\phi_l(x) - \phi_l(\hat{x})\|_2^2, \tag{3}$$

where $\phi_l(\cdot)$ denotes feature activations extracted from a pre-trained network. Perceptual supervision has proven effective for restoration and super-resolution, preserving semantic structure and fine details.

**Frequency-domain loss..** Since degradation processes primarily attenuate high-frequency components, we additionally enforce spectral consistency by minimising the $\ell_1$ distance between Fourier representations:

$$\mathcal{L}_{FFT} = \|\mathcal{F}(x) - \mathcal{F}(\hat{x})\|_1, \tag{4}$$

where $\mathcal{F}$ denotes the 2D Fourier transform. Frequency-domain supervision has been shown to improve reconstruction of fine textures and structural details in image restoration [24, 43].

### 3.3. Sensor simulation

Learning-based image restoration requires paired data, where a simulated degraded image is paired with a simulated reference image used for training. 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 sensor simulator that approximates realistic satellite imaging conditions while maintaining full control over the simulation parameters. From there, we use the sensor simulator to degrade Level-1 products, creating simulated Level-1 and simulated raw imagery while limiting the impact of the original Level-1 restoration on the image. By doing so, we ensure better control over our experiment parameters.

**Spatial resolution.** Spatial blur is modelled through a parametric MTF combining optical and sensor effects:

$$\mathrm{MTF}(f_x, f_y) = \exp(-\gamma f_r) \cdot \mathrm{sinc}(f_x) \cdot \mathrm{sinc}(f_y), \tag{5}$$

where

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

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

$$\mathrm{PSF} = \mathcal{F}^{-1}\{\mathrm{MTF}\}, \tag{6}$$

and each spectral band is convolved with the PSF.

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

$$I_{\text{GSD}}(x, y) = I_{\text{blur}}(x_0 + rx,\ y_0 + ry), \tag{7}$$

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

**Signal-dependent noise model.** Radiometric degradation is modelled using signal-dependent Gaussian noise centered on 0 and with standard deviation:

$$\sigma^2(L) = \alpha L + \beta, \tag{8}$$

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

$$\sigma_i^2 = \left(\frac{L_i}{\mathrm{SNR}_i}\right)^2, \qquad \sigma_i^2 = \alpha L_i + \beta. \tag{9}$$

Noise is sampled as

$$\sigma(L) = \sqrt{\alpha L + \beta}, \tag{10}$$

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. In our experiments, ConvBEERS is trained solely on simulated images.

# 4. Experimental Setup

## 4.1. OpenAerialMap dataset

Our dataset is constructed from high-resolution aerial RGB imagery obtained from OpenAerialMap [38], 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 comprises a wide variety of landscape types (urban, suburban, rural, agricultural, forested, coastal, and mountainous), with an emphasis on urban environments characterised 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 artefacts such as ringing or oversmoothing.

From there, these aerial RGB images are converted to panchromatic imagery and degraded using our physics-based sensor simulator pipeline (see Sec.3.3), which models the key characteristics of the target system (GSD, MTF, and SNR). This process generates paired simulated raw and reference images under controlled, reproducible degradation settings. Several configurations are produced at a target GSD of 50 cm and stored as 12-bit panchromatic TIFF patches. Their specificities are disclosed in Tab. 1.

Two types of datasets are constructed. The first type, referred to as *Sim-Degraded-Fixed / Sim-Reference-Fixed*, corresponds to a fixed-degradation setting defined at a nominal operating point representative of typical acquisition conditions. The second type consists of variable-degradation datasets, obtained by sampling MTF and SNR values within realistic ranges around this nominal configuration. Introducing such variability in the degradation process encourages the model to learn to handle diverse imaging conditions. It improves robustness while reducing the risk of overfitting to a specific system configuration.

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 $L_0 = 25$ W/m$^2$/sr/$\mu$m and $L_1 = 100$ W/m$^2$/sr/$\mu$m. Number of patches is reported as *Train / Test*. Training patches are $128 \times 128$ pixels, while test patches are $1500 \times 1500$ pixels.

## 4.2. Metrics

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

**Full-reference image quality metrics.**

When 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. Reported values are computed over the entire simulated test set and averaged across all test patches to ensure statistically representative evaluation. 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 representations and have demonstrated improved sensitivity to texture reconstruction, edge preservation, and structural distortions.

**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 im-

Table 2. Full-reference image quality metrics on OpenAerialDataset.

| Method | PSNR↑ | SSIM↑ | LPIPS↓ | DISTS↓ |
|---|---|---|---|---|
| Traditional | 20.6 | 0.843 | 0.216 | 0.220 |
| ConvBEERS | 27.5 | 0.932 | 0.068 | 0.077 |

plemented in the *MTF Estimator* plugin for QGIS [19]. For each image, measurements are performed on three manually selected edge regions, and the reported value is the average of these regions, ensuring a robust and representative estimate of spatial resolution. The SNR is evaluated over 13 homogeneous regions using the variance-based radiometric approach described by Liu et al. [35], following the signal-dependent noise model introduced in Sec. 3.3.

These indicators provide a physically interpretable assessment of blur compensation and noise amplification, ensuring consistency with operational imaging specifications.

# 5. Experimental results and embedded system performance analysis

## 5.1. Benchmark on OpenAerialMap

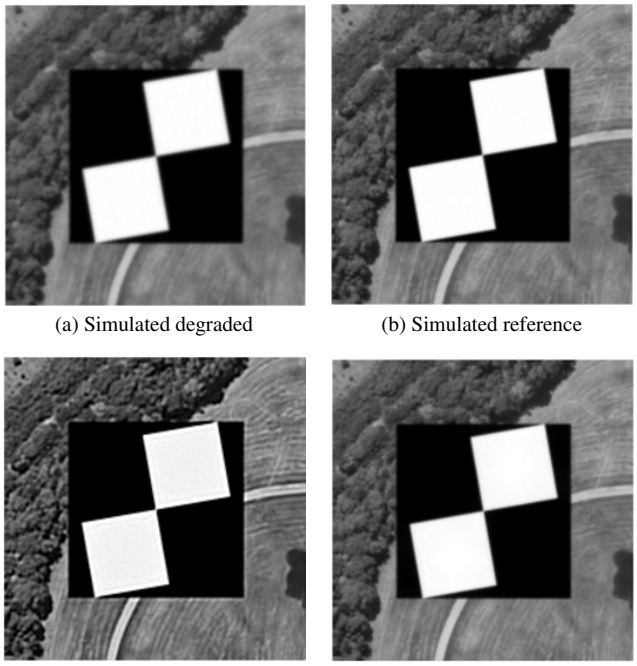
We evaluate ConvBEERS on the OpenAerialMap dataset and compare its performance to the Pléiades-HR restoration, which serves as a baseline [12, 29]. The model is trained for 30 epochs on the training subset of *Sim-Degraded-Variable / Sim-Reference-Fixed*, with the degraded set serving as input, and the reference set serving as target (see Sec. 3.3). Both approaches are evaluated on the fixed-degradation simulated test set (*Sim-Degraded-Fixed / Sim-Reference-Fixed*) using the metrics described in Sec. 4.2.

Results are reported in Tab. 2 and Tab. 3. We see that ConvBEERS outperforms the traditional pipeline, with a +6.9 dB PSNR gain, higher SSIM, and a markedly lower LPIPS score. The restored MTF reaches 0.20 while preserving comparable SNR, 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.

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 artefacts. A synthetic resolution pattern embedded in simulated scenes (Fig. 2) 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.

Table 3. Physical quality metric and end-to-end time on OpenAerialDataset.

| Method | MTF@Nyq | SNR @ $L_0/L_1$ |
|---|---|---|
| Traditional | N/A | 65.33 / 138.51 |
| ConvBEERS | 0.20 | 66.27 / 156.07 |



(a) Simulated degraded (b) Simulated reference

(c) Traditional (d) ConvBEERS

Figure 2. Image restoration on a simulated degraded image with embedded mire.

## 5.2. Robustness to blur degradation

The robustness of ConvBEERS 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. 1) and evaluated on simulated degraded images generated at fixed MTF levels. Three blur levels are considered: low, medium, and high, corresponding to $\mathrm{MTF}_{7\%}$, $\mathrm{MTF}_{5\%}$, and $\mathrm{MTF}_{3\%}$, respectively. For each input image, the MTF is estimated both before restoration (on degraded input) and after restoration by ConvBEERS. 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 as blur severity increases. Even for the most degraded input ($\mathrm{MTF}_{3\%}$), ConvBEERS provides a significant improvement in spatial resolution.

These results indicate that ConvBEERS trained on variable degradation conditions does not overfit to a single nominal operating point, but instead learns a robust restoration behaviour that generalises to unseen blur levels. This ro-

Table 4. MTF before and after restoration with ConvBEERS for different blur levels on simulated degraded data.

| Deg. Level | MTF (Deg.) | MTF (Rest.) | $\Delta$MTF |
|---|---|---|---|
| $\text{MTF}_{7\%}$ | 0.0741 | 0.2017 | +0.1276 |
| $\text{MTF}_{5\%}$ | 0.0558 | 0.1555 | +0.0997 |
| $\text{MTF}_{3\%}$ | 0.0336 | 0.1383 | +0.1047 |

bustness is a key property for operational satellite imagery, where imaging conditions may deviate from nominal specifications due to acquisition geometry, temporal variations, or sensor ageing.

### 5.3. Experiments on Real Pléiades Data

For real-world validation, we use optical panchromatic imagery acquired by the Pléiades satellite constellation operated by the French Space Agency (Centre National d'Études Spatiales, CNES), which provided access to raw products. In this experiment, we aim to evaluate the generalisation capability of ConvBEERS. Thus, the Pléiades images are not used to train or fine-tune the model.

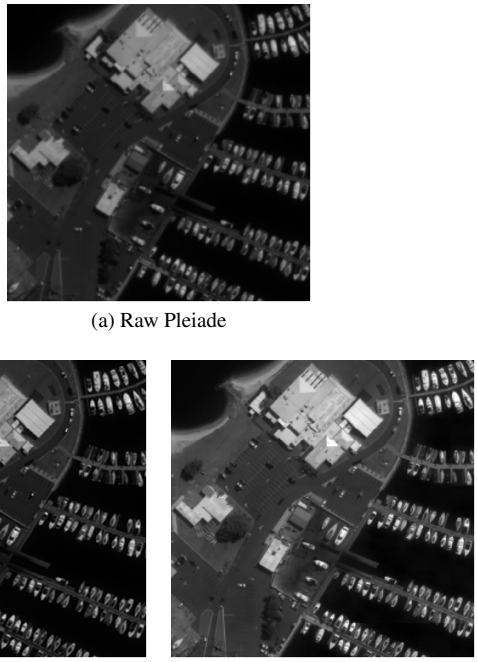
Evaluation is conducted through qualitative visual inspection and comparison with the conventional restoration pipeline. In addition, a blind visual assessment was conducted in which observers compared the outputs of both methods without knowledge of the processing approach. Both methods produce visually close results, with comparable levels of sharpness and noise control. However, in a blind evaluation, ConvBEERS-restored images were judged as more visually comfortable, with improved edge definition and no evident non-physical artefacts or excessive ringing.

Overall, these observations indicate that the learning-based restoration model generalises well from simulated data to real Pléiades imagery, while preserving physically plausible image characteristics and offering perceptually favourable rendering.

### 5.4. Impact on an application scenario: object detection on the DIOR dataset

To demonstrate the added value of learning-based restoration in a fully integrated processing pipeline, we evaluate the impact of ConvBEERS on an object detection task using the DIOR dataset [33]. This well-known dataset in the Earth observation community comprises 23,463 optical remote sensing images covering 20 classes. Object detection is performed using the YOLOv11n model [27]. The nano version, combined with a 256x256 downsampling, ensures that resulting performance reflects realistic edge-system constraints.

As displayed in Fig. 4, the original DIOR images are degraded using the simulation framework described in Sec. 4.1, generating pairs of simulated raw and simulated



(a) Raw Pleiade

(b) Traditional (c) ConvBEERS

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

Figure 3. Comparison of raw Pléiades image and restored results using the traditional algorithm and ConvBEERS.

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% | $30 \pm 5$ / $70 \pm 5$ | 18 000 / 2000 / 3463 |
| Deg-High | 0.1–0.5% | $10 \pm 5$ / $50 \pm 5$ | 18 000 / 2000 / 3463 |

$L_0 = 25$ W/m$^2$/sr/$\mu$m and $L_1 = 100$ W/m$^2$/sr/$\mu$m. #Patches is reported as *train / val / test*. Patches are dowsampled to $256 \times 256$ pixels.

reference images for training. Two test sets with different degradation levels are then created: *Deg-Med*, representative of typical satellite sensor degradation, and *Deg-High*, which applies more severe degradation to further challenge the restoration process. The corresponding degradation parameters are detailed in Tab. 5. Then, images are restored using ConvBEERS. Finally, YOLOv11n is trained on all three configuration (reference, degraded and restored). The trained model are used to output detection metrics comparing the baseline (reference images) with both degraded and restored images.

Tab. 6 displays the results for all three datasets. First, we can observe that the model achieves a strong baseline performance, with 75.90% mAP@50 and 54.10% mAP@90. These values indicate that the training process is under control, with no major flaws.

Regarding the *Deg-Med* experiment, we observe that degradation slightly impacted the detection model's perfor-

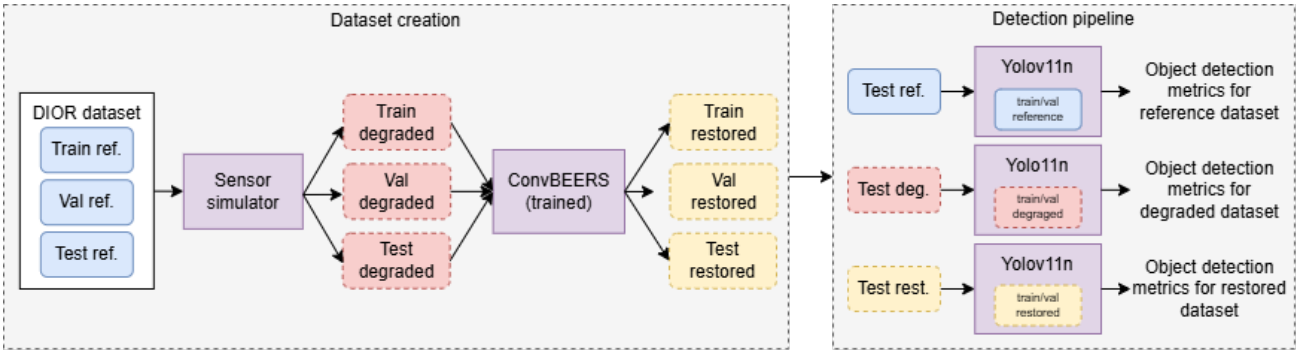



Figure 4. **Object detection evaluation pipeline**. The simulated degraded and restored datasets are created from the DIOR dataset using both our sensor simulation framework and ConvBEERS. Then, three YOLOv11n models are trained independently on (1) reference images, (2) degraded images, and (3) restored images using the training and validation sets. Their performance is compared on the test sets to evaluate the impact of restoration on a detection task.

mance, resulting in decreases of 2.6% for both mAP@50 and mAP@90. However, the restored version *Restored images* almost fully recovered the original performance.

Regarding the *Deg-High* experiment, the degradation effect 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 show that even when severe degradation heavily impacts detector performance, using our proposed light restoration model significantly improves performance.

Table 6. Object detection comparison between degraded and restored DIOR datasets with Yolo11n on 256 x 256 pixel images.

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

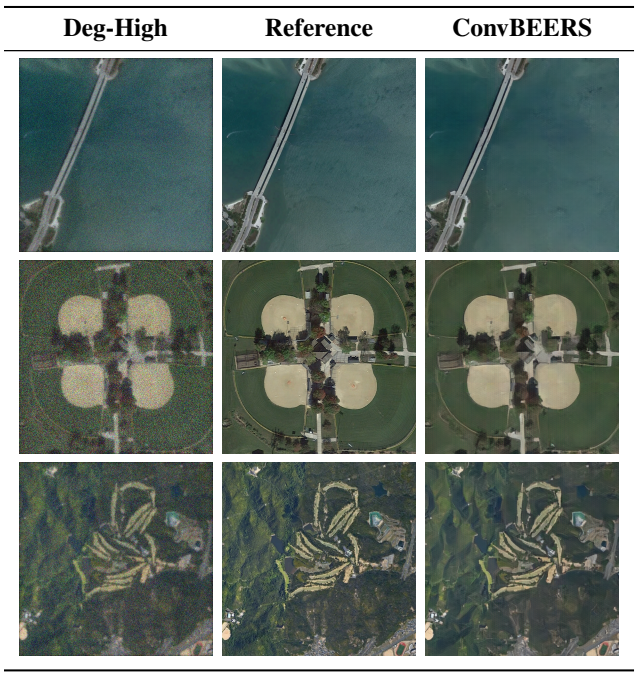



Figure 5. Qualitative comparison of restored RGB images with ConvBEERS

## 5.5. Embedded Deployment

As a final experiment, we benchmark ConvBEERS against a traditional Pléiades-inspired restoration pipeline [29] on the Versal Xilinx VCK190 FPGA platform. The Versal has gained traction in the space community, thanks to its radiation-tolerant specifications [16, 40], and state-of-the-art performances [8, 18, 23]. In this work, we use a Deep Learning Processor Unit (DPU) [1, 32], designed to accelerate CNN models. ConvBEERS is evaluated on the DPU accelerator. The baseline pipeline is executed only on the CPU because its algorithmic structure is not compatible with the DPU's operator constraints [1]. Since FPGA deployment of AI models typically requires quantisation, ConvBEERS is converted to INT8 precision using post-training quantisation [42]. Tab. 7 reports the results of a quantised ConvBEERS compared to its floating-point version as well as the baseline restoration. Firstly, the difference between the outputs of the quantised and floating-point ConvBEERS remains limited, with a Mean Absolute Error of 1.39%. This result ensures that the restored im-

Table 7. Quantization performance for a 13000 × 9000 panchromatic swath image.

| Model | Prec. | Size (Mb) | PSNR (dB) | SSIM |
|---|---|---|---|---|
| Baseline | Float32 | 1.97 | 28.37 | 0.600 |
| ConvBEERS | Float32 | 4.78 | 28.58 | 0.887 |
| ConvBEERS | Int8 | 2.01 | 26.77 | 0.848 |

The MAE between ConvBEERS Float32 and Int8 outputs is 0.0139, with a Std of 0.0413.

Table 8. Benchmark on Versal AI Edge VCK190 FPGA for a 13600 × 9016 panchromatic swath image. FPS are given for 640x640 patches

| Method | Latency (s)↓ | FPS↑ | Power (W)↓ | Pxls/s/W↑ |
|---|---|---|---|---|
| Baseline (CPU) | 1724 | 0.2 | 22.5 | 3161 |
| ConvBEERS (DPU) | 42 | 7.2 | 29.6 | 98965 |

age is similar before and after quantisation. We observe a modest decrease in PSNR (from 28.6 dB to 26.7 7dB). This degradation is expected for a restoration task. Nevertheless, the model still shows strong restoration performance, as reflected by an SSIM value close to 88.5.

Tab. 8 reports the deployment results on the Versal VCK190. The traditional baseline is slow on the Versal CPU, with an inference time of ∼28 min for the whole swath. In contrast, ConvBEERS takes advantage of the DPU accelerator to process the swath in 42 s, achieving up to 7.2 FPS on 640 × 640 patches. Even though the power consumption increases with the use of the DPU (from 22.5 W to 29.6 W), results show that this approach is compatible with onboard real-time processing.

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

Beyond the comparison with traditional pipelines, ConvBEERS's robustness to variations in blur and noise levels is a key advantage in practical scenarios, where imaging conditions may deviate from nominal specifications due to acquisition geometry, temporal effects, or sensor ageing. Learning-based restoration enables earlier and more flexible exploitation of satellite imagery.

When combined with onboard object-detection tasks, ConvBEERS significantly improved the AI model's performance. These results, coupled with ConvBEERS' ability to preserve the physical plausibility of the image, reinforce the suitability of learning-based restoration as a task-oriented preprocessing step for onboard AI applications in remote sensing.

Nevertheless, the ConvBEERS model currently only tackles radiometric degradation, lacking geometric enhancement capabilities. Thus, in cases of geometrical degradation, such as a significant band misalignment, coupling ConvBEERS with another algorithm might be needed.

Overall, the reasonable overhead in power, coupled with the growing performance of satellite's computing capability, converge towards the use of ConvBEERS-like architecture for future spaceborn processing.

## 7. Conclusion

In this paper, we valuated the relevance of learning-based methods for onboard preprocessing of raw satellite imagery. The residual layers of the EDSR model are adapted into our proposed restoration model, ConvBEERS. 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 OpenAerialMap images and real Raw/CNES-L1 pairs indicate that ConvBEERS-restored images exhibit characteristics comparable to those obtained with traditional restoration pipelines. When evaluated across varying levels of blur degradation, ConvBEERS demonstrated strong robustness, yielding stable restored MTF values across degradation scenarios. Furthermore, ConvBEERS was assessed as a lightweight preprocessing module for AI-based object detection. Experimental results show that detection performance consistently improves, confirming the benefit of restoration for onboard AI tasks. Finally, onboard deployment confirms its suitability for spaceborne applications, demonstrating significant acceleration while maintaining restoration quality.

In future work, the ConvBEERS architecture could be improved with remote sensing imagery knowledge. Indeed, because remote sensing data differ from natural images, incorporating physical priors [2] may prevent the emergence of non-physical geometric artifacts. Another promising direction would be to integrate restoration directly within a task-dependent AI model, simplifying the onboard processing pipeline by eliminating the need for two independent models. Finally, while this work focuses on radiometric restoration of raw images, it would be very interesting to couple with learning-based geometric restoration techniques, such as the one proposed by Del Prete et al. in PyRawS [41]. The result would be a fully operational restoration model that can work with any raw image.

# References


[1] Deep learning processor unit (dpu) — vitis ai. Xilinx Documentation, 2026. https://xilinx.github.io/Vitis-AI/. 7

[2] Chayan Banerjee, Kien Nguyen, Clinton Fookes, and George Karniadakis. Physics-informed computer vision: A review and perspectives, 2024. 8

[3] 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

[4] 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. 1

[5] Maud Biquard, Marie Chabert, Florence Genin, Christophe Latry, and Thomas Oberlin. Variational bayes image restoration with compressive autoencoders. *IEEE Transactions on Image Processing*, 34:2896–2909, 2025. 2

[6] 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, 2

[7] Yochai Blau and Tomer Michaeli. The perception-distortion tradeoff. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6228–6237, 2018. 2

[8] Jacob Brown, Colton Yates, Jeffrey Goeders, and Michael Wirthlin. Rareplanes detection using yolov5 on the versal adaptive soc. In *2025 IEEE Aerospace Conference*, pages 1–9, 2025. 7

[9] Mikaël Carlavan. *Optimization of the compression/restoration chain for satellite images*. Theses, Université Nice Sophia Antipolis, 2013. 1, 2

[10] Xiangyu Chen, Xintao Wang, Wenlong Zhang, Xiangtao Kong, Yu Qiao, Jiantao Zhou, and Chao Dong. Hat: Hybrid attention transformer for image restoration, 2025. 2

[11] 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, 2

[12] Sebastien Fourest Christophe Latry, Gwendoline Blancher. Chaine de fusion p+xs des images pléiades-hr, 2013. 2, 5

[13] ESA Copernicus. Senti wiki: S2 products, 2025. https://sentiwiki.copernicus.eu/web/s2-products, Accessed on February, 2026. 2

[14] Roberto Del Prete, Gabriele Meoni, Manuel Salvoldi, Domenico Barretta, Maria Daniela Graziano, Nicolas Longépé, and Alfredo Renga. Enhanced maritime monitoring via onboard processing of raw multi-spectral imagery by deep learning. In *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, pages 1713–1717, 2024. 1

[15] Lorenzo Diana, Jia Xu, and Luca Fanucci. Oil spill identification from sar images for low power embedded systems using cnn. *Remote Sensing*, 13(18), 2021. 1

[16] Arnaud Dufour, Jérôme Carron, François Pierron, Matthieu Fongral, David Dangla, Guillaume Bascoul, Françoise Bezerra, Julien Mekki, Florence Malou, and Pierre Maillard. 7nm finfet technology heavy ion sel evaluation using xilinx versal as case study. In *2021 21th European Conference on Radiation and Its Effects on Components and Systems (RADECS)*, pages 1–6, 2021. 7

[17] Fabien Gallet, Christophe Marabotto, and Thomas Chambon. Exploring ai-based satellite pose estimation: from novel synthetic dataset to in-depth performance evaluation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 6770–6778, 2024. 1

[18] Luis M. Garcés-Socarrás, Raudel Cuiman, Flor Ortiz, Juan A. Vásquez-Peralvo, Jorge L. González-Rios, Mouhamad Chehaitly, Arkadii Kazanskii, Sahar Malmir, Amirhossein Nik, Jan Thoemel, Sumit Kumar, Marcele Kuhfuss, Swetha Varadajulu, Eva Lagunas, Juan C. M. Duncan, Jorge Querol, and Symeon Chatzinotas. Onboard machine learning for satellite edge computing: The spaice project use case. In *2025 European Data Handling & Data Processing Conference (EDHPC)*, pages 1–8, 2025. 7

[19] Jorge Gil. Fidmtf: An mtf estimator (qgis plugin). https://plugins.qgis.org/plugins/mtf_estimator/. 5

[20] Thomas Goudemant, Benjamin Francesconi, Michelle Aubrun, Erwann Kervennic, Ingrid Grenet, Yves Bobichon, and Marjorie Bellizzi. Onboard anomaly detection for marine environmental protection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 7918–7931, 2024. 1

[21] Juntao Guan, Rui Lai, and Ai Xiong. Wavelet deep neural network for stripe noise removal. *IEEE Access*, 7:44544–44554, 2019. 2

[22] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. 2

[23] Younis Ibrahim, Li Chen, and Tian Haonan. Deep learning-based ship detection on fpgas. In *2022 14th International Conference on Computational Intelligence and Communication Networks (CICN)*, pages 454–459, 2022. 7

[24] Liming Jiang, Bo Dai, Wayne Wu, and Chen Change Loy. Focal frequency loss for image reconstruction and synthesis, 2021. 3

[25] Younghyun Jo, Sejong Yang, and Seon Joo Kim. Investigating loss functions for extreme super-resolution. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1705–1712, 2020. 3

[26] Eleftheria Kalogirou, George Melillos, Diofantos G. Hadjimitsis, Egbert Schwarz, and Stefan Wiehle. Maritime surveillance in Cyprus using Sentinel-1 SAR. In *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2023*, page 127280N, 2023. 1

[27] Rahima Khanam and Muhammad Hussain. Yolov11: An overview of the key architectural enhancements, 2024. 6

[28] Georges Labrèche, Cesar Guzman, and Sam Bammens. Generative ai... in space! adversarial networks to denoise images onboard the ops-sat-1 spacecraft. In *2024 IEEE Aerospace Conference*, pages 1–17, 2024. 2

[29] 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, 2, 5, 7

[30] 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, 2

[31] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4681–4690, 2017. 2

[32] Guoyu Li, Pengbo Zheng, Jian Weng, and Enshan Yang. Dpuv4e: High-throughput dpu architecture design for cnn on versal acap, 2025. 7

[33] Ke Li, Gang Wan, Gong Cheng, Liqiu 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

[34] 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

[35] 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. 5

[36] Gabriele Meoni, Roberto Del Prete, Federico Serva, Alix De Beusscher, Olivier Colin, and Nicolas Longépé. Unlocking the use of raw multispectral earth observation imagery for onboard artificial intelligence. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17:12521–12537, 2024. 1

[37] Janne Mäyrä, Elina A. Virtanen, Ari-Pekka Jokinen, Joni Koskikala, Sakari Väkevä, and Jenni Attila. Mapping recreational marine traffic from sentinel-2 imagery using yolo object detection models. *Remote Sensing of Environment*, 326: 114791, 2025. 1

[38] OpenAerialMap. Openaerialmap. https://openaerialmap.org. 4

[39] Esha Pahwa, Achleshwar Luthra, and Pratik Narang. Lvrnet: Lightweight image restoration for aerial images under low visibility, 2023. 2

[40] Noah Perryman, Christopher Wilson, and Alan George. Evaluation of xilinx versal architecture for next-gen edge computing in space. In *2023 IEEE Aerospace Conference*, pages 1–11, 2023. 7

[41] Roberto Del Prete, Manuel Salvoldi, Domenico Barretta, Nicolas Longépé, Gabriele Meoni, Arnon Karnieli, Maria Daniela Graziano, and Alfredo Renga. Enhancing maritime situational awareness through end-to-end onboard raw data analysis, 2024. 2, 8

[42] Gaurav Upadhyay, C Siva Prasad, V Akash, Ayush Kumar Singh, R Monish Sai, Srijan Ghosal, Subrat Kar, S H Shantala, J Thakur LalitKrushna, S Srividhya, and Sunil Alias Balwantrao. Enhancing cnn inference time and reducing latency on edge and resource-constrained systems through quantization. In *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)*, pages 1–6, 2024. 7

[43] Chenyang Wang, Junjun Jiang, Zhiwei Zhong, and Xianming Liu. Spatial-frequency mutual learning for face super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 22356–22366, 2023. 3

[44] Hua Wei, Kun Gao, Jing Wang, Qiuyan Tang, Xiongxin Tang, and Fanjiang Xu. Optical imaging degradation simulation and transformer-based image restoration for remote sensing. *IEEE Geoscience and Remote Sensing Letters*, 21: 1–5, 2024. 2

[45] Wenjia Xu, Guangluan Xu, Yang Wang, Xian Sun, Daoyu Lin, and Yirong Wu. Deep memory connected neural network for optical remote sensing image restoration. *Remote Sensing*, 10(12), 2018. 2

[46] Ziyao Yi, Davide Piccinini, Diego Valsesia, Tiziano Bianchi, and Enrico Magli. Scalable neural pushbroom architectures for real-time denoising of hyperspectral images onboard satellites, 2026. 2

[47] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017. 2

[48] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep cnn denoiser prior for image restoration, 2017. 2

[49] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric, 2018. 3

[50] Yuanzhi Zhu, Kai Zhang, Jingyun Liang, Jiezhang Cao, Bihan Wen, Radu Timofte, and Luc Van Gool. Denoising diffusion models for plug-and-play image restoration, 2023. 2