# HSREST: BENCHMARKING HYPERSPECTRAL SATELLITE IMAGE RESTORATION ACROSS TASKS, DATA REGIMES, AND REGIONS

**Anonymous authors**
Paper under double-blind review

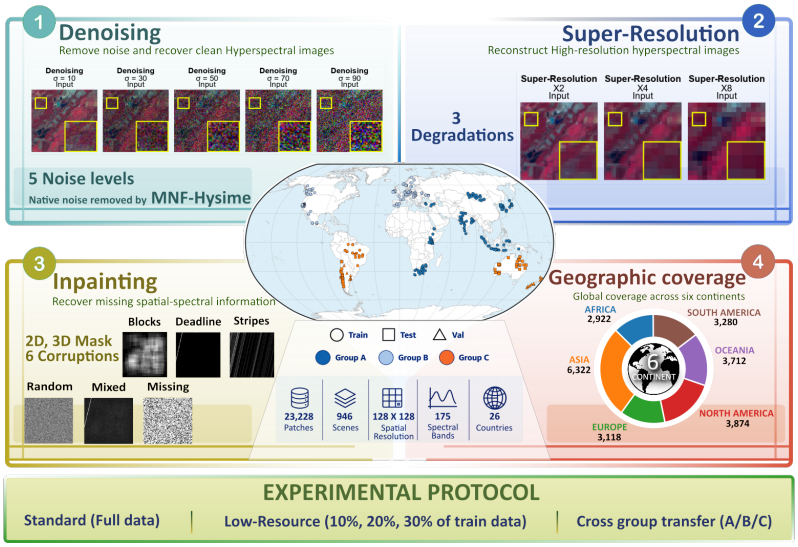



Figure 1: **Overview of HSRest:** HSRest contains 23,228 non-overlapping $128 \times 128$ hyperspectral patches from 946 EO-1 Hyperion scenes spanning 26 countries across six continents. It benchmarks three image restoration tasks - Gaussian denoising, spatial super-resolution, and structured inpainting, under full-data, low-resource (10%, 20%, 30%), and A/B/C cross-geographic group evaluation.

## ABSTRACT

Hyperspectral satellite images provide valuable observations of the Earth, but their restoration is difficult to evaluate consistently because existing studies differ in data selection, degradation protocols, training regimes, and evaluation metrics. We introduce **HSRest**, a benchmark for systematic evaluation of hyperspectral image (HSI) restoration, comprising 23,228 non-overlapping $128 \times 128$ patches extracted from 946 EO-1 Hyperion scenes across 26 countries and six continents. HSRest covers three representative restoration tasks, namely Gaussian denoising, spatial super-resolution, and structured inpainting, with multiple degradation levels, and organizes the same corpus under three complementary evaluation settings. Setting I provides country- and scene-disjoint train/validation/test partitions for standard evaluation. Setting II defines fixed 10%, 20%, and 30% training subsets for low-resource evaluation. Setting III partitions all scenes into three geographically distinct groups for evaluation under geographic distribution shifts. We benchmark recent restoration methods using complementary spatial, spectral, and efficiency measures. Across tasks, training budgets, metrics, and held-out regions, method rankings change substantially, showing that conclusions drawn from a single split, degradation level, or reconstruction metric may not reliably characterize restoration performance. HSRest releases fixed partitions, degradation protocols, metadata, evaluation code, and baseline configurations, providing a reproducible testbed for data-efficient and region-robust HSI restoration. HSRest Dataset Link

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

# 1 INTRODUCTION

Hyperspectral satellite imagery records the Earth's surface across hundreds of narrow spectral bands, providing material-sensitive information beyond RGB and conventional multispectral sensing. In practice, sensor noise, missing measurements, stripe artefacts, and limited spatial resolution can distort both spatial structure and spectral signatures. Image restoration is therefore an important preprocessing step for reliable hyperspectral analysis, and an effective restoration method must recover degraded observations while preserving both spatial fidelity and spectral structure.

Despite substantial progress in HSI restoration, evaluation remains fragmented. Classical datasets such as Indian Pines Baumgardner et al. (2015), Pavia University, Salinas, Botswana Scenes (2020), Houston Debes et al. (2014), and Chikusei Yokoya & Iwasaki (2016) were developed primarily for recognition and related remote-sensing tasks, while larger resources such as HySpecNet-11K Fuchs & Demir (2023), HyperGlobal-450K Wang et al. (2025), and SpectralEarth Braham et al. (2025) focus on compression, representation learning, and pretraining. These resources do not jointly specify controlled restoration degradations, common reference construction, scene- and country-disjoint splits, low-resource settings, and geographic-transfer evaluation. Moreover, exact source scenes, geographic locations, sample counts, and train/validation/test assignments are not always fully documented, making faithful reproduction of reported comparisons difficult.

We introduce **HSRest**, a benchmark for HSI restoration constructed from public EO-1 Hyperion Level-1Gst radiance products. Hyperion is used because its archive is publicly accessible and suitable for benchmark release; restricting the benchmark to a single sensor also keeps spectral response, ground-sampling distance, and radiometric characteristics fixed, allowing the effects of restoration task, training-data availability, and geographic variation to be studied without additional cross-sensor confounding. Our main contributions are:

- **A standardized hyperspectral restoration benchmark.** HSRest contains $23{,}228$ non-overlapping $128 \times 128$ patches from $946$ scenes spanning $26$ countries across six continents. A literature-established $175$-band configuration is used for benchmark evaluation, while the original $242$-band representation is retained for diagnostic analysis. The benchmark covers three representative restoration tasks: Gaussian denoising, spatial super-resolution at $\times 2$, $\times 4$, and $\times 8$, and structured inpainting. To account for native sensor noise in Hyperion Level-1Gst imagery, the denoising benchmark uses an MNF-HySIME-based procedure to construct noise-reduced reference cubes.
- **Three reproducible evaluation settings.** HSRest defines three complementary evaluation settings. Setting I provides a global country- and scene-disjoint train/validation/test split for standard full-data evaluation. Setting II uses fixed 10%, 20%, and 30% subsets of the same training split, with unchanged validation and test sets, to evaluate existing restoration methods under low-resource conditions. Setting III partitions the data into geographically distinct Group-A, Group-B, and Group-C groups for cross-group transfer without shared scenes or countries. Each sample retains its source-scene, geographic, patch, and split information, enabling exact reconstruction of the benchmark settings.
- **Condition-dependent benchmark analysis.** HSRest is a benchmarking study rather than a new restoration architecture. Existing restoration methods are evaluated using common spatial, spectral, and efficiency criteria. Our experiments show that relative method rankings can change across restoration tasks, training budgets, metric families, and held-out geographic regions. Thus, conclusions drawn from a single split, degradation level, or reconstruction metric may not fully characterize restoration robustness.

# 2 RELATED WORK

We organize the related work into three areas: hyperspectral datasets, restoration methods, and evaluation protocols. These areas provide the context for the design and evaluation of HSRest.

Traditional hyperspectral datasets such as Pavia University Scenes (2020), Indian Pines Baumgardner et al. (2015), Salinas Scenes (2020), Houston Debes et al. (2014), and Chikusei Yokoya & Iwasaki (2016) have been widely used for classification, segmentation, and spectral–spatial representation

Table 1: Comparison of HSRest with representative hyperspectral datasets and benchmarks. While several existing resources are substantially larger, HSRest is organized specifically for restoration benchmarking with fixed degradation protocols, low-resource training settings, and geographic-transfer evaluation.

| Dataset | Sensor | Patches / Scenes | Scene / Patch Size | Bands | Restoration Benchmark | Low-Resource Setting | Geographic Transfer |
|---|---|---|---|---|---|---|---|
| Indian Pines( Baumgardner et al. (2015)) | AVIRIS | 1 scene | $145 \times 145$ | 224 | No | No | No |
| KSC( Scenes (2020)) | AVIRIS | 1 scene | $512 \times 614$ | 224 | No | No | No |
| Salinas( Scenes (2020)) | AVIRIS | 1 scene | $512 \times 217$ | 224 | No | No | No |
| Pavia University( Scenes (2020)) | ROSIS | 1 scene | $610 \times 340$ | 103 | No | No | No |
| Pavia Center( Scenes (2020)) | ROSIS | 1 scene | $1096 \times 1096$ | 102 | No | No | No |
| Botswana( Scenes (2020)) | EO-1 Hyperion | 1 scene | $1476 \times 256$ | 242 | No | No | No |
| Houston University( Debes et al. (2014)) | ITRES-CASI | 1 scene | $349 \times 1905$ | 144 | No | No | No |
| SDFC( Sun et al. (2023)) | HAHS | 1 scene | $1201 \times 201$ | 63 | No | No | No |
| Chikusei( Yokoya & Iwasaki (2016)) | Hyperspec -VNIR-C | 1 scene | $2517 \times 2335$ | 128 | No | No | No |
| WHU-Hi( Zhong et al. (2020)) | Nano-Hyperspec | 3 scenes | $1217 \times 303$ | 270-274 | No | No | No |
| Hyperview( Nalepa et al. (2022)) | HySpex | $\sim$3k patches | Avg. $\sim 60 \times 60$ | 150 | No | No | No |
| Toulouse HS( Thoreau et al. (2024)) | HySpex | 9 scenes | $\sim 4000 \times 1000$ | 310 | No | No | No |
| HySpecNet-11K( Fuchs & Demir (2023)) | EnMAP | $\sim$11k patches | $128 \times 128$ | 202 | No | No | No |
| MSST( Scheibenreif et al. (2023)) | EnMAP | $\sim$20k patches | $64 \times 64$ | 200 | No | No | No |
| HSIHybrid( Wang et al. (2024)) | Multiple | $\sim$4M patches | $9 \times 9$ | Variable | No | No | No |
| HyperGlobal -450K ( Wang et al. (2025)) | EO-1 Hyperion +Gaofen-5B | $\sim$447k patches | $64 \times 64$ | 242 / 330 | No | No | No |
| SpectralEarth ( Braham et al. (2025)) | EnMAP | $\sim$538k patches | $128 \times 128$ | 202 | No | No | No |
| **HSRest (Ours)** | **EO-1 Hyperion** | **23,228 patches** | $\mathbf{128 \times 128}$ | **175 / 242** | **Yes** | **Yes** | **Yes** |

*Note.* A *restoration benchmark* specifies degraded inputs, task-specific reference targets, fixed splits, and evaluation scripts. *Low-resource* and *geographic transfer* mean those protocols are part of the original release. "No" means the original resource does not define that protocol.

learning. These datasets were not originally designed around restoration-specific references, degradation models, or standardized evaluation splits. More recent resources substantially increase scale and geographic coverage. HySpecNet-11K Fuchs & Demir (2023) provides $11{,}483$ non-overlapping hyperspectral patches and primarily targets learning-based compression, HyperGlobal-450K Wang et al. (2025) supports large-scale hyperspectral pretraining, and SpectralEarth Braham et al. (2025) provides large-scale pretraining data together with downstream evaluation for recognition-oriented tasks. These resources significantly advance hyperspectral learning, but their primary benchmark objectives differ from systematic restoration evaluation across degradation type, training-data regime, and geographic distribution shift.

Table 1 positions HSRest relative to representative hyperspectral datasets and benchmarks in terms of scale, restoration-specific evaluation, low-resource settings, and geographic-transfer evaluation. While several existing resources are substantially larger, HSRest is organized specifically for controlled restoration benchmarking across data regimes and geographic transfer.

Hyperspectral restoration has been studied using a broad range of methodological families. Earlier approaches exploit low-rank structure, sparse representations, tensor factorization, and image priors, while more recent methods use convolutional, recurrent, transformer-based, state-space, and diffusion architectures to model spatial and spectral dependencies. These methods have been applied to denoising, spatial super-resolution, and inpainting, but their experimental setups often differ in scene selection, spectral-band filtering, patch extraction, corruption severity, normalization, training-data size, and train–test construction. Consequently, reported results across restoration methods are not always directly comparable, particularly when exact samples and split assignments differ across studies.

Evaluation protocols also vary considerably. Pixel-level reconstruction is commonly assessed using peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and root mean square error (RMSE), while hyperspectral restoration additionally requires explicit assessment of spectral fidelity. Metrics such as spectral angle mapper (SAM), spectral information divergence (SID), ERGAS, and spectral correlation provide complementary information about spectral and relative reconstruction quality Yuhas et al. (1992); Chang & Du (2000). In addition, exact source scenes, geographic locations, sample identities, and train/validation/test assignments are not always fully reported, which can make faithful reproduction of published comparisons difficult.

HSRest complements these lines of work by providing a common benchmark in which existing restoration methods are evaluated on the same data, controlled degradations, fixed sample assignments,

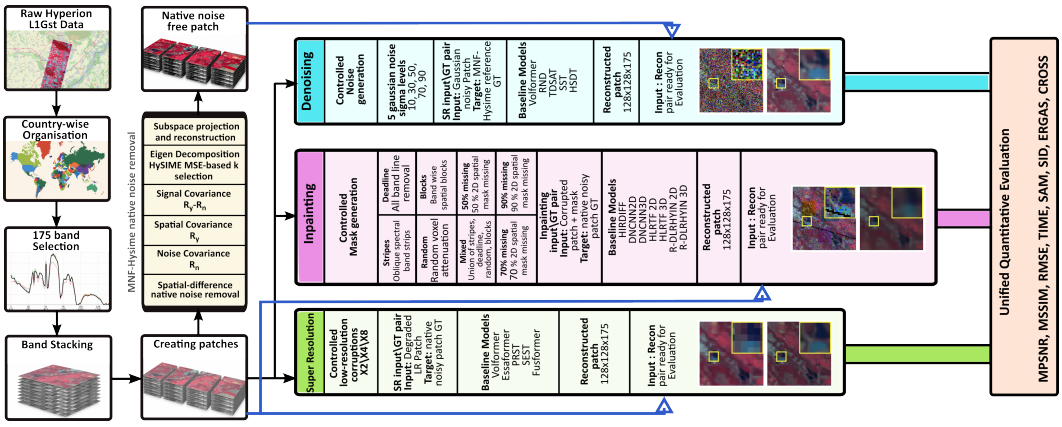



Figure 2: Complete pipeline from data creation and model training to quantitative assessment

and spatial–spectral criteria see Figure 2. It further extends evaluation beyond a single full-data setting through explicit low-resource and geographic-transfer protocols, enabling restoration behaviour to be compared across tasks, data regimes, and regions.

## 3 HSREST DATASET

HSRest is a benchmark for hyperspectral satellite image restoration built from publicly available EO-1 Hyperion Level-1Gst radiance products (242 bands, 0.357-2.576 $\mu$m, $\approx 10$ nm spectral sampling, and 30 m ground-sampling distance). The benchmark contains 23,228 non-overlapping $128 \times 128$ patches extracted from 946 scenes spanning 26 countries across six continents. Restricting HSRest to a single sensor keeps spectral response, ground-sampling distance, and radiometric characteristics fixed while restoration task, training-data availability, and geographic region vary.

### 3.1 PATCH EXTRACTION AND SPECTRAL CURATION

Non-overlapping $128 \times 128$ patches are extracted with stride 128. Incomplete boundary windows and patches containing only invalid or zero-valued spatial support are discarded. Scene selection is guided by seven Anderson-inspired surface categories Anderson (1976): water, cropland, forest, grassland/shrubland, bare soil, built-up areas, and wetlands. These categories are used only to support sampling and metadata organization and are never treated as prediction targets. False-colour composites generated from bands 43, 23, and 12 are used for visual inspection of candidate regions. Representative hyperspectral samples and task-specific degradations are illustrated in Figure 3.

The primary evaluation configuration follows the literature-established 175-band selection of Awad et al. Awad et al. (2018): bands 8-57, 77-120, 127-166, and 181-221. The complementary 67 bands (1-7, 58-76, 121-126, 167-180, and 222-242) are predominantly noisy, low-response, or strongly affected by atmospheric absorption and are retained for diagnostic analysis. Detailed source-scene metadata and acquisition information are provided in Tables 5 - 7 in the supplementary material, together with additional details on scene ingestion, spectral stacking, geospatial metadata preservation, and null-pixel filtering.

### 3.2 BENCHMARK ORGANIZATION

The same patch identities are reused across denoising, spatial super-resolution, and structured inpainting. HSRest defines three complementary evaluation settings.

**Setting I.** A global country- and scene-disjoint train/validation/test split containing 16,213, 2,369, and 4,646 patches, respectively.

**Setting II.** Fixed 10%, 20%, and 30% subsets of the Setting I training partition containing 1,615, 3,243, and 4,863 patches, respectively, while the validation and test sets remain unchanged. These predefined subsets are reused across all compared restoration methods to enable controlled low-

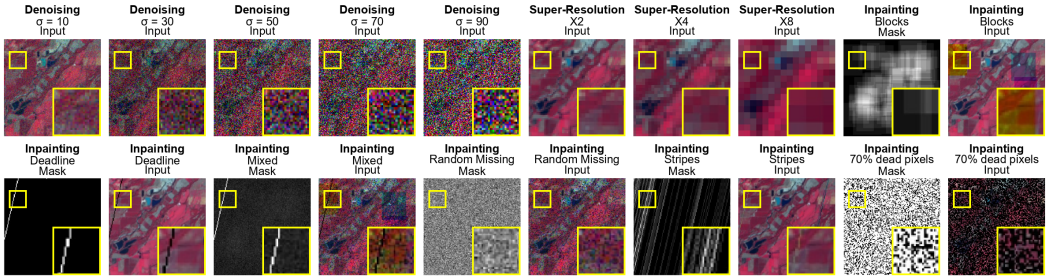



Figure 3: Denoising, Inpainting and Super-Resolution images provided in the HSRest benchmark dataset.

Table 2: Best denoising results across full-data, low-resource, and geographic-transfer protocols. Model abbreviations: V-VolFormer, R-RND, T-TDSAT, H-HSDT, and S-SST.

| Protocol | $\sigma = 10$ | | | | $\sigma = 30$ | | | | $\sigma = 50$ | | | | $\sigma = 70$ | | | | $\sigma = 90$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Spatial Metrics** | | | | | | | | | | | | | | | | | | | | |
| | **MPSNR** ↑ | **MSSIM** ↑ | **RMSE** ↓ | **Time** ↓ | **MPSNR** ↑ | **MSSIM** ↑ | **RMSE** ↓ | **Time** ↓ | **MPSNR** ↑ | **MSSIM** ↑ | **RMSE** ↓ | **Time** ↓ | **MPSNR** ↑ | **MSSIM** ↑ | **RMSE** ↓ | **Time** ↓ | **MPSNR** ↑ | **MSSIM** ↑ | **RMSE** ↓ | **Time** ↓ |
| **Full Data** | | | | | | | | | | | | | | | | | | | | |
| **Full Data** | 50.794 (V) | 0.987 (V) | 0.005 (V) | 0.094 (S) | 42.895 (V) | 0.940 (V) | 0.010 (R/T) | 0.101 (S) | 41.197 (R) | 0.922 (V) | 0.013 (R) | 0.179 (S) | 40.826 (V) | 0.918 (V) | 0.016 (V) | 0.152 (S) | 40.108 (V) | 0.897 (V) | 0.016 (V) | 0.128 (S) |
| **Low-Resource Fine-Tuning** | | | | | | | | | | | | | | | | | | | | |
| **10% of Train** | 43.628 (H) | 0.909 (H) | 0.009 (H) | 0.138 (S) | 35.963 (R) | 0.782 (V) | 0.021 (R) | 0.134 (S) | 35.389 (V) | 0.758 (V) | 0.023 (V) | 0.070 (S) | 34.861 (V) | 0.751 (V) | 0.026 (V) | 0.214 (H) | 40.494 (V) | 0.897 (V) | 0.016 (V) | 0.143 (S) |
| **20% of Train** | 41.779 (H) | 0.923 (V) | 0.011 (H) | 0.214 (H) | 41.105 (V) | 0.898 (V) | 0.014 (V) | 0.130 (S) | 40.573 (V) | 0.886 (V) | 0.016 (V) | 0.070 (S) | 39.989 (V) | 0.877 (V) | 0.018 (V) | 0.214 (H) | 39.605 (V) | 0.869 (V) | 0.019 (V) | 0.140 (S) |
| **30% of Train** | 43.310 (R) | 0.928 (V) | 0.009 (R) | 0.141 (S) | 41.473 (V) | 0.920 (V) | 0.013 (V) | 0.139 (S) | 41.296 (V) | 0.917 (V) | 0.014 (V) | 0.071 (S) | 40.494 (V) | 0.897 (V) | 0.016 (V) | 0.213 (H) | 40.484 (V) | 0.894 (V) | 0.015 (V) | 0.141 (S) |
| **Groupwise Training** | | | | | | | | | | | | | | | | | | | | |
| **Train A → Test BC** | 47.387 (V) | 0.962 (V) | 0.007 (V) | 0.096 (S) | 45.160 (V) | 0.948 (V) | 0.009 (V) | 0.099 (S) | 43.723 (V) | 0.931 (V) | 0.012 (V) | 0.114 (S) | 43.261 (V) | 0.923 (V) | 0.013 (V) | 0.129 (S) | 43.015 (V) | 0.914 (V) | 0.014 (V) | 0.121 (S) |
| **Train B → Test CA** | 46.298 (V) | 0.956 (V) | 0.008 (V/R) | 0.096 (S) | 44.642 (V) | 0.935 (V) | 0.010 (V) | 0.102 (S) | 43.502 (V) | 0.921 (V) | 0.012 (V) | 0.189 (S) | 42.816 (V) | 0.909 (V) | 0.014 (V) | 0.154 (S) | 42.618 (V) | 0.904 (V) | 0.015 (V) | 0.123 (S) |
| **Train C → Test AB** | 48.719 (V) | 0.971 (V) | 0.006 (V) | 0.099 (S) | 45.531 (V) | 0.948 (V) | 0.010 (V) | 0.095 (S) | 44.919 (V) | 0.939 (V) | 0.011 (V) | 0.188 (S) | 43.878 (V) | 0.920 (V) | 0.013 (V) | 0.204 (S) | 43.869 (V) | 0.914 (V) | 0.013 (V) | 0.130 (S) |
| **Spectral Metrics** | | | | | | | | | | | | | | | | | | | | |
| | **SAM** ↓ | **SID** ↓ | **ERGAS** ↓ | **CC** ↑ | **SAM** ↓ | **SID** ↓ | **ERGAS** ↓ | **CC** ↑ | **SAM** ↓ | **SID** ↓ | **ERGAS** ↓ | **CC** ↑ | **SAM** ↓ | **SID** ↓ | **ERGAS** ↓ | **CC** ↑ | **SAM** ↓ | **SID** ↓ | **ERGAS** ↓ | **CC** ↑ |
| **Full Data** | | | | | | | | | | | | | | | | | | | | |
| **Full Data** | 0.121 (H) | 0.002 (V) | 34.963 (V) | 0.983 (H) | 0.296 (H) | 0.002 (V) | 43.433 (V) | 0.852 (H) | 0.319 (H) | 0.004 (V) | 47.104 (V) | 0.884 (H) | 0.436 (H) | 0.005 (V) | 37.837 (V) | 0.804 (H) | 0.482 (H) | 0.006 (V) | 53.656 (V) | 0.795 (H) |
| **Low-Resource Fine-Tuning** | | | | | | | | | | | | | | | | | | | | |
| **10% of Train** | 0.123 (H) | 0.003 (H) | 399.837 (H) | 0.947 (H) | 0.215 (H) | 0.006 (V) | 662.682 (V) | 0.926 (H) | 0.306 (H) | 0.006 (V) | 691.935 (V) | 0.900 (H) | 0.391 (H) | 0.007 (V) | 112.012 (R) | 0.858 (H) | 0.473 (H) | 0.002 (V) | 121.978 (R) | 0.674 (H) |
| **20% of Train** | 0.134 (H) | 0.002 (V) | 180.995 (R) | 0.952 (H) | 0.307 (H) | 0.003 (V) | 108.381 (R) | 0.911 (H) | 0.448 (H) | 0.003 (V) | 109.035 (R) | 0.862 (H) | 0.637 (H) | 0.003 (V) | 108.512 (R) | 0.754 (H) | 0.681 (H) | 0.003 (V) | 108.502 (R) | 0.712 (H) |
| **30% of Train** | 0.137 (H) | 0.002 (V) | 167.906 (R) | 0.953 (H) | 0.259 (H) | 0.002 (V) | 159.551 (R) | 0.927 (H) | 0.369 (H) | 0.001 (V) | 175.523 (V) | 0.829 (H) | 0.433 (H) | 0.002 (V) | 108.363 (R) | 0.790 (H) | 0.432 (H) | 0.002 (V) | 108.289 (R) | 0.818 (H) |
| **Groupwise Training** | | | | | | | | | | | | | | | | | | | | |
| **Train A → Test BC** | 0.156 (H) | 0.001 (V) | 159.803 (V) | 0.931 (H) | 0.331 (H) | 0.001 (V) | 126.859 (V) | 0.873 (H) | 0.563 (H) | 0.001 (V) | 108.527 (V) | 0.790 (H) | 0.575 (H) | 0.001 (V) | 130.778 (V) | 0.777 (H) | 0.589 (H) | 0.001 (V) | 126.047 (V) | 0.762 (H) |
| **Train B → Test CA** | 0.172 (H) | 0.001 (V) | 193.936 (V) | 0.952 (H) | 0.263 (H) | 0.001 (V) | 211.710 (V) | 0.905 (H) | 0.495 (H) | 0.001 (V) | 250.811 (V) | 0.895 (H) | 0.520 (H) | 0.002 (V) | 198.748 (V) | 0.811 (H) | 0.743 (H) | 0.002 (V) | 212.328 (V) | 0.766 (H) |
| **Train C → Test AB** | 0.087 (H) | 0.001 (V) | 62.249 (R) | 0.996 (H) | 0.256 (H) | 0.001 (V) | 125.779 (V) | 0.990 (H) | 0.408 (H) | 0.001 (V) | 107.719 (R) | 0.977 (H) | 0.809 (H) | 0.001 (V) | 107.713 (R) | 0.946 (H) | 0.801 (H) | 0.001 (V) | 107.717 (R) | 0.839 (H) |

resource evaluation. The exact sample counts and training schedules for the full-data and low-resource settings are summarized in 9 in the supplementary material.

**Setting III.** The complete corpus is partitioned into geographically distinct Group-A, Group-B, and Group-C groups containing 9,244, 6,992, and 6,992 patches, respectively. Models are trained on one group and evaluated on the remaining geographically disjoint groups, with no shared scene or country between source and target partitions. Group-wise experiments follow a fixed training budget and do not use target-group samples for model selection. The geographic composition of Group-A, Group-B, and Group-C, together with country-level split assignments and low-resource subset membership, is detailed in 10 in the supplementary material.

The benchmark is not explicitly balanced by country or Anderson surface category. Instead, valid spectral support, geographic diversity, and leakage-free scene-disjoint evaluation are prioritized.

### 3.3 Release and Reproducibility

We release the 175-band benchmark representation, the original 242-band data, metadata containing scene, group, and split identifiers, fixed split definitions, degradation generators, and evaluation scripts. The complete release structure is summarized in Table 8 in the supplementary material, while the end-to-end benchmark pipeline is illustrated in Figure 2.

## 4 Experimental Protocols and Benchmark Analysis

We evaluate existing hyperspectral restoration methods under Settings I–III (Section 3.2) using the curated 175-band representation. Rather than identifying a single overall winner, we examine whether method rankings remain stable across degradation severity, training-data availability, metric family, restoration task, and Group-A/Group-B/Group-C geographic transfer. Structured inpainting is evaluated using a fixed $N = 100$ patch protocol. Compact task-level summaries are reported in Tables 2 - 4, while complete per-method results across all metrics are provided in Tables 12-16 in the supplementary material.

### 4.1 Reference Construction and Degradation Protocols

**Gaussian denoising.** Let $X$ denote an unrestored Hyperion Level-1Gst patch, which retains the native sensor characteristics, and let $X_{\mathrm{ref}} = F_{\text{MNF-HySIME}}(X)$ denote its noise-reduced reference. Since native sensor noise is inherent in the Level-1Gst observations, we preserve $X$ as the source observation and apply controlled Gaussian corruption to it rather than to $X_{\mathrm{ref}}$. Specifically, $X$ is normalized by $M = \max(X)$, perturbed with $N \sim \mathcal{N}(0, (\sigma/255)^2)$, clipped to $[0, 1]$, and rescaled to obtain $Y_\sigma = M\,\mathrm{clip}(X/(M + \epsilon) + N, 0, 1)$, where $\sigma \in \{10, 30, 50, 70, 90\}$. Complete denoising results across all five noise levels are reported separately for spatial and efficiency metrics in Table 12 and for spectral metrics in Table 13 in the supplementary material.

**Spatial super-resolution.** The normalized $[0, 1]$ cube is used as the high-resolution reference. For $s \in \{2, 4, 8\}$, low-resolution observations are generated by bicubic anti-aliased downsampling, $L_s = \mathcal{D}_s(X)$; bicubic interpolation is retained where required by an evaluated implementation. Complete inpainting results across the six corruption cases are reported in Table 16 in the supplementary material.

**Structured inpainting.** We construct eight corruption cases spanning six inpainting degradation types motivated by commonly observed satellite-image degradations: stripes, dead lines, block corruption, random missing observations, mixed corruption, and spatial masks with approximately $50\%$, $70\%$, and $90\%$ missing pixels. These cases are compactly represented as $Y = (1 - M) \odot X + M \odot \mathcal{C}(X)$, where $M$ and $\mathcal{C}(\cdot)$ specify the corruption. Spatial masks are replicated across spectral bands. Full MNF–HySIME parameters, corruption generators, and inpainting-sampling details are provided in the supplementary material.

All degradations are controlled stress tests intended for reproducible comparison, rather than complete physical forward models of satellite acquisition artefacts.

### 4.2 Benchmark Setup and Evaluation Criteria

For denoising, we evaluate VolFormer (Yu & Gao (2025)), HSDT (Lai et al. (2023)), SST (Li et al. (2023)), TDSAT (Zhang et al. (2024)), and RND (Zhang et al. (2026)). For spatial super-resolution, we evaluate VolFormer (Yu & Gao (2025)), Fusformer (Hu et al. (2022)), SEST (Yang et al. (2024)), PSRT (Deng et al. (2023)), and ESSAFormer (Zhang et al. (2023)). The supplementary inpainting evaluation includes DIP2D/DIP3D (Sidorov & Yngve Hardeberg (2019)), HLRTF (Luo et al. (2022)), HIR-Diff (Pang et al. (2024)), and R-DLRHyN (Niresi & Chi (2023)).

A complete summary of the evaluated models, including methodological type, abbreviation, venue, and publication year, is provided in Table 11 in the supplementary material.

Table 3: Best super-resolution results across full-data, low-resource, and geographic-transfer protocols at different upscaling factors. Model abbreviations: V-VolFormer, E-ESSAformer, P-PSRT, SE-SEST, and F-Fusformer.

| Protocol | ×2 | | | | | | | | ×4 | | | | | | | | ×8 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ |
| **Full Data** | | | | | | | | | | | | | | | | | | | | | | | | |
| Full Data | 37.098 (F) | 0.889 (F) | 0.041 (F) | 0.829 (F) | 4.217 (SE) | 0.011 (V) | 31.955 (SE) | 0.786 (P) | 35.673 (F) | 0.836 (F) | 0.054 (F) | 0.482 (F) | 3.775 (SE) | 0.012 (V) | 17.022 (SE) | 0.757 (F) | 34.415 (F) | 0.788 (P) | 0.058 (SE) | 0.551 (F) | 5.379 (F) | 0.012 (V) | 11.031 (F) | 0.662 (F) |
| **Low-Resource Fine-Tuning** | | | | | | | | | | | | | | | | | | | | | | | | |
| 10% of Train | 38.166 (SE) | 0.865 (SE) | 0.044 (P) | 0.509 (F) | 3.290 (SE) | 0.013 (V) | 68.672 (SE) | 0.836 (SE) | 36.032 (SE) | 0.804 (SE) | 0.050 (SE) | 0.502 (F) | 3.565 (SE) | 0.014 (V) | 35.012 (SE) | 0.779 (SE) | 34.386 (F) | 0.743 (F) | 0.054 (F) | 0.528 (F) | 4.761 (SE) | 0.013 (V) | 20.532 (SE) | 0.667 (F) |
| 20% of Train | 37.969 (F) | 0.874 (SE) | 0.041 (SE) | 0.498 (F) | 3.930 (F) | 0.011 (V) | 51.534 (F) | 0.831 (F) | 35.196 (F) | 0.813 (SE) | 0.047 (SE) | 0.523 (F) | 5.730 (SE) | 0.010 (V) | 30.500 (F) | 0.765 (F) | 34.636 (F) | 0.756 (SE) | 0.050 (SE) | 0.545 (F) | 5.005 (F) | 0.010 (V) | 13.065 (SE) | 0.686 (F) |
| 30% of Train | 37.599 (F) | 0.879 (SE) | 0.041 (SE) | 0.449 (F) | 4.883 (SE) | 0.009 (V) | 55.273 (SE) | 0.834 (V) | 36.301 (F) | 0.817 (F) | 0.045 (F) | 0.497 (F) | 5.408 (F) | 0.010 (V) | 31.670 (F) | 0.784 (F) | 34.962 (SE) | 0.769 (P) | 0.049 (SE) | 0.518 (F) | 5.225 (F) | 0.010 (V) | 9.433 (SE) | 0.689 (F) |
| **Groupwise Training** | | | | | | | | | | | | | | | | | | | | | | | | |
| Train A → Test BC | 37.305 (F) | 0.861 (F) | 0.046 (P) | 0.513 (F) | 4.381 (SE) | 0.014 (V) | 44.535 (SE) | 0.809 (P) | 35.674 (F) | 0.806 (P) | 0.050 (P) | 0.668 (F) | 4.900 (P) | 0.011 (V) | 18.911 (SE) | 0.756 (F) | 34.275 (F) | 0.747 (P) | 0.056 (P) | 1.414 (F) | 5.601 (F) | 0.015 (V) | 10.910 (SE) | 0.675 (F) |
| Train B → Test CA | 37.169 (F) | 0.862 (F) | 0.042 (F) | 0.541 (F) | 4.969 (SE) | 0.011 (V) | 51.655 (SE) | 0.779 (P) | 35.841 (F) | 0.811 (P) | 0.048 (P) | 0.723 (F) | 4.394 (P) | 0.012 (V) | 19.111 (SE) | 0.748 (P) | 34.067 (F) | 0.755 (V) | 0.053 (F) | 0.697 (F) | 6.467 (F) | 0.012 (V) | 16.808 (SE) | 0.630 (F) |
| Train C → Test AB | 38.267 (F) | 0.842 (V) | 0.037 (F) | 0.479 (F) | 6.916 (SE) | 0.010 (V) | 28.123 (F) | 0.744 (F) | 36.536 (F) | 0.807 (F) | 0.041 (F) | 0.687 (F) | 7.039 (SE) | 0.011 (V) | 22.859 (SE) | 0.710 (F) | 34.938 (F) | 0.748 (F) | 0.048 (F) | 0.618 (F) | 6.530 (V) | 0.012 (V) | 9.068 (F) | 0.638 (V) |

Settings I-II use only the fixed validation partition for checkpoint selection. Setting III uses a fixed training budget without target-group validation or adaptation. Within each task, all methods receive identical inputs, references, splits, and metric implementations.

The main paper emphasizes PSNR and SAM as representative spatial and spectral criteria. Complete MSSIM, RMSE, SID, ERGAS, CC, runtime, and inpainting results are provided in the supplementary material. Higher PSNR indicates better reconstruction fidelity, whereas lower SAM indicates better preservation of spectral direction.

### 4.3 Benchmark Analysis Across Conditions

The benchmark is most informative when interpreted through the inductive biases of the evaluated models rather than as a collection of isolated winner cells. The interpretations below connect measured behaviour with the published model designs; they should not be interpreted as causal architectural ablations.

**Spatial and spectral fidelity favour different priors.** As shown in Table 2, under Setting I at $\sigma = 10$, VolFormer achieves $50.79$ dB PSNR, whereas HSDT achieves the lowest SAM ($0.12$). The complete spatial and spectral denoising results are reported in Tables 12 and 13, respectively. VolFormer's volumetric interaction models dependencies across the hyperspectral cube and is consistent with strong pixel-level reconstruction. HSDT instead combines spatial–spectral separable processing with guided spectral attention, which is consistent with stronger preservation of inter-band geometry.

A similar separation appears in $\times 2$ super-resolution (Table 3): Fusformer reaches $37.10$ dB PSNR, whereas SEST attains the lowest SAM ($4.22$). The complete super-resolution results are reported in Tables 14 and 16. Fusformer's fusion-oriented representation favours spatial recovery, while SEST explicitly emphasizes spectral structure. Thus, PSNR and SAM can favour different restoration mechanisms, showing why a single image-fidelity metric is insufficient for hyperspectral evaluation.

**Low-resource training changes the ranking.** Table 2 shows that at $\sigma = 10$, the PSNR leader changes from VolFormer under full-data training to HSDT at $10\%$ ($43.63$ versus $37.45$ dB for VolFormer). Similarly, Table 3 shows that the $\times 2$ super-resolution leader changes from Fusformer to SEST at $10\%$ ($38.17$ dB). These shifts suggest different sensitivities to reduced training diversity and architectural priors. Because the $10\%$, $20\%$, and $30\%$ subsets also differ in country composition, Setting II should be interpreted as a low-resource robustness test rather than a pure sample-complexity study; the exact subset composition is given in Table 10 in the supplementary material.

**Failure cases expose operating-range mismatch.** Some methods exhibit degenerate behaviour outside particular benchmark conditions. RND, which is designed around real-noise decomposition, reaches SAM values close to $90^\circ$ and CC near zero under severe controlled Gaussian corruption. TDSAT also produces very low PSNR under some low-resource configurations. These cases are

Table 4: Best inpainting results across eight corruption cases. Model abbreviations: D3–Deep Hyperspectral Prior (3D), H2-HLRTF (2D mask), H3-HLRTF (3D mask), HR-HIR-Diff, and R3-R-DLRHyIN (3D mask).

| Metric | Stripes | Dead lines | Blocks | Random Missing | Mixed | Missing50 | Missing70 | Missing90 |
|---|---|---|---|---|---|---|---|---|
| **MPSNR ↑** | inf (R3) | inf (H2) | 67.685 (H3) | 60.879 (H3) | 59.268 (H3) | inf (R3) | 47.324 (R3) | 43.601 (R3) |
| **MSSIM ↑** | 1.000 (H3/R3) | 1.000 (H2/H3/R3) | 1.000 (H3) | 0.998 (H3/R3) | 0.998 (H3) | 1.000 (R3) | 0.967 (R3) | 0.944 (D3) |
| **RMSE ↓** | 0.000 (H3) | 0.001 (H2/H3/R3) | 0.001 (H3) | 0.002 (H3/R3) | 0.002 (H3) | 0.000 (R3) | 0.009 (D3/R3) | 0.011 (D3) |
| **Time ↓** | 2.251 (HR) | 1.741 (HR) | 1.744 (HR) | 1.728 (HR) | 1.720 (HR) | 1.743 (HR) | 1.743 (HR) | 1.761 (HR) |
| **SAM ↓** | 0.131 (H3) | 0.022 (H3) | 0.202 (HR) | 0.217 (HR) | 0.246 (HR) | 0.038 (R3) | 0.059 (D3) | 0.065 (D3) |
| **SID ↓** | 0.000 (R3) | 0.000 (H3/R3) | 0.003 (R3) | 0.002 (R3) | 0.009 (R3) | 0.000 (R3) | 0.015 (R3) | 0.022 (D3) |
| **ERGAS ↓** | 3.332 (H3) | 1.534 (R3) | 6.940 (H3) | 10.883 (R3) | 18.667 (R3) | 1.269 (R3) | 23.506 (R3) | 30.714 (R3) |
| **CC ↑** | 1.000 (H3/R3) | 1.000 (R3) | 1.000 (R3) | 1.000 (R3) | 0.996 (R3) | 1.000 (R3) | 0.994 (D3) | 0.992 (D3) |

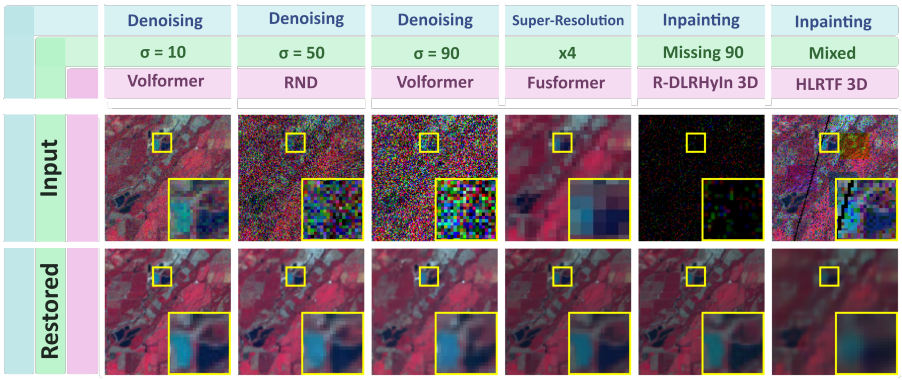



Figure 4: Qualitative comparison of hyperspectral image restoration across multiple tasks. The displayed methods correspond to the top-performing models for each task according to the respective quantitative benchmarks.

reported for transparency but are interpreted as failure modes rather than normal competitive operating points.

**Super-resolution changes the relevant inductive bias.** Increasing the scale from $\times 2$ to $\times 8$ progressively removes spatial information, making reconstruction increasingly dependent on the prior used to infer unobserved high-frequency structure Table. 3, 14, 15 in supplementary material. Fusformer performs strongly on spatial fidelity, SEST favours spectral preservation in several settings, ESSAFormer explicitly models spectral correlations, and PSRT emphasizes multi-scale spatial reconstruction. Their changing order across scales indicates that success at mild degradation does not guarantee robustness when the inverse problem becomes more underdetermined.

VolFormer provides an additional cross-task comparison. Its volumetric interaction is highly effective for denoising PSNR, yet it does not dominate spatial super-resolution. Denoising primarily suppresses perturbation while retaining observed structure, whereas super-resolution must infer spatial content absent from the input. A strong common backbone therefore does not imply a task-independent ranking.

**Geographic transfer changes relative performance.** Setting III fixes the sensor, spectral response, GSD, and product type while changing the geographic source-target distribution, thereby reducing sensor-related confounding. For denoising, VolFormer remains strong on PSNR across Group-A→Group-B+C, Group-B→Group-C+A, and Group-C→Group-A+B, while HSDT remains strong in SAM. Super-resolution rankings are less stable; for example, PSRT is competitive under Setting I but degrades substantially under Group-C→Group-A+B.

These shifts should not be interpreted as evidence of a unique geographic signature of an architecture. Land-cover composition, reflectance, texture, illumination, and acquisition conditions may all differ

between groups. The important empirical result is that rankings measured on the standard split need not persist on geographically held-out scenes.

**Inpainting highlights the role of structural priors.** Inpainting differs from denoising and super-resolution because observations are missing rather than corrupted or spatially downsampled. Accordingly, low-rank and sample-specific priors become particularly relevant. HLRTF exploits hierarchical low-rank tensor structure, R-DLRHyIn combines an untrained CNN with explicit low-rank regularization, DIP2D/DIP3D use sample-specific untrained priors, and HIR-Diff combines low-dimensional structure with a pretrained diffusion prior.

Their relative behaviour changes with mask geometry and severity, supporting evaluation across multiple corruption families. Numerically perfect or near-perfect results, including unbounded PSNR when evaluated MSE becomes zero, are treated as protocol-sensitive cases and are not used as standalone evidence of superiority.

Representative qualitative restoration examples are shown in Figure 4, complementing the quantitative comparisons in Tables 2-4.

**Overall benchmark finding.** No single method ordering survives all tasks, training regimes, metric families, and geographic transfers. Whole-cube interaction, explicit spectral modelling, fusion-based reconstruction, and structural priors become advantageous under different conditions. HSRest therefore acts as a diagnostic benchmark for identifying *which restoration prior remains effective under which condition*, rather than as a single-score leaderboard.

### 4.4 Scope and Limitations

HSRest uses EO-1 Hyperion Level-1Gst imagery only, so the current experiments do not establish cross-sensor generalization. Its degradations provide reproducible paired stress tests rather than complete physical sensor models. Setting II changes both training-set size and country composition. The denoising reference $X_{\mathrm{ref}}$ is an MNF–HySIME reconstruction rather than a directly observed noise-free acquisition, so small inter-method differences should be interpreted cautiously. Finally, HSRest evaluates restoration fidelity rather than downstream classification, detection, spectral unmixing, or material-recognition performance.

## 5 Conclusion

We introduced HSRest, a hyperspectral satellite restoration benchmark comprising $23{,}228$ non-overlapping Hyperion patches from $946$ scenes across $26$ countries, a standardized $175$-band evaluation configuration, controlled restoration protocols, and three complementary settings spanning full-data, low-resource, and Group-A/Group-B/Group-C geographic-transfer evaluation.

Across denoising and spatial super-resolution, the relative ordering of existing restoration methods changes with degradation severity, training-data availability, metric family, and geographic source-target configuration. Spatial reconstruction and spectral fidelity frequently favour different methods, while rankings obtained under full-data training do not necessarily persist under low-resource or geographically held-out evaluation. The supplementary inpainting study further illustrates the dependence of restoration behaviour on corruption structure and the underlying reconstruction prior.

These results show that hyperspectral restoration cannot be characterized reliably using a single split, metric, or operating condition. HSRest provides a controlled and reproducible framework for evaluating which restoration priors remain effective across different data regimes and restoration conditions. We release the benchmark data, metadata, fixed splits, degradation generators, and evaluation code to support reproducible comparison and future hyperspectral restoration research.

### AI use statement

In this work, generative AI tools were used for language editing, including grammar correction, rephrasing, and improving the readability of the manuscript. They were also used for minor LaTeX formatting and presentation assistance. We did not use generative AI tools for research ideation,

hypothesis formulation, methodology or experimental design, implementation of methods, data analysis, interpretation of results, or any other research tasks requiring disclosure under the ICLR 2027 AI Policy. All AI-assisted content was carefully reviewed, revised, and verified by the authors. The authors take full responsibility for the final content, including all claims, analyses, and results.

### Reproducibility statement

We provide an anonymous repository containing the HSRest dataset description and the files required to reproduce part of our experiments. The main paper details the dataset construction, restoration tasks, detailed experimental settings, and evaluation protocols, while the supplementary material provides implementation details. Together with the provided repository, these materials are intended to facilitate reproduction of the reported results.

## References

James Richard Anderson. *A land use and land cover classification system for use with remote sensor data*, volume 964. US Government Printing Office, 1976.

Mahmoud E Awad, Reda Amer, Alberto López-Galindo, Mahmoud M El-Rahmany, Luis F García del Moral, and César Viseras. Hyperspectral remote sensing for mapping and detection of egyptian kaolin quality. *Applied Clay Science*, 160:249–262, 2018.

Marion Baumgardner, Larry Biehl, and David Landgrebe. 220 band aviris hyperspectral image data set: June 12, 1992 indian pine test site 3. *(No Title)*, 2015.

Nassim Ait Ali Braham, Conrad M Albrecht, Julien Mairal, Jocelyn Chanussot, Yi Wang, and Xiao Xiang Zhu. Spectralearth: Training hyperspectral foundation models at scale. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2025.

Chein-I Chang and Qian Du. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Transactions on Information Theory*, 46(5):1927–1932, 2000.

Christian Debes, Andreas Merentitis, Roel Heremans, Jürgen Hahn, Nikolaos Frangiadakis, Tim Van Kasteren, Wenzhi Liao, Rik Bellens, Aleksandra Pižurica, Sidharta Gautama, et al. Hyperspectral and lidar data fusion: Outcome of the 2013 grss data fusion contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):2405–2418, 2014.

Shang-Qi Deng, Liang-Jian Deng, Xiao Wu, Ran Ran, Danfeng Hong, and Gemine Vivone. Psrt: Pyramid shuffle-and-reshuffle transformer for multispectral and hyperspectral image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–15, 2023.

Martin Hermann Paul Fuchs and Begüm Demir. Hyspecnet-11k: A large-scale hyperspectral dataset for benchmarking learning-based hyperspectral image compression methods. In *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1779–1782. IEEE, 2023.

Jin-Fan Hu, Ting-Zhu Huang, Liang-Jian Deng, Hong-Xia Dou, Danfeng Hong, and Gemine Vivone. Fusformer: A transformer-based fusion network for hyperspectral image super-resolution. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2022.

Zeqiang Lai, Chenggang Yan, and Ying Fu. Hybrid spectral denoising transformer with guided attention. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 13019–13029. IEEE, 2023.

Miaoyu Li, Ying Fu, and Yulun Zhang. Spatial-spectral transformer for hyperspectral image denoising. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 1368–1376, 2023.

Yisi Luo, Xi-Le Zhao, Deyu Meng, and Tai-Xiang Jiang. Hlrtf: Hierarchical low-rank tensor factorization for inverse problems in multi-dimensional imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 19303–19312, 2022.

Jakub Nalepa, Bertrand Le Saux, Nicolas Longépé, Lukasz Tulczyjew, Michal Myller, Michal Kawulok, Krzysztof Smykala, and Michal Gumiela. The hyperview challenge: Estimating soil parameters from hyperspectral images. In *2022 IEEE International Conference on Image Processing (ICIP)*, pp. 4268–4272. IEEE, 2022.

Keivan Faghih Niresi and Chong-Yung Chi. Robust hyperspectral inpainting via low-rank regularized untrained convolutional neural network. *IEEE Geoscience and Remote Sensing Letters*, 20:1–5, 2023.

Li Pang, Xiangyu Rui, Long Cui, Hongzhong Wang, Deyu Meng, and Xiangyong Cao. Hir-diff: Unsupervised hyperspectral image restoration via improved diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3005–3014, 2024.

Hyperspectral Remote Sensing Scenes. Available online: http://www. ehu. eus/ccwintco/index. php. *Hyperspectral_Remote_Sensing_Scenes (accessed on 22 April 2020)*, 2020.

Linus Scheibenreif, Michael Mommert, and Damian Borth. Masked vision transformers for hyperspectral image classification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2166–2176, 2023.

Oleksii Sidorov and Jon Yngve Hardeberg. Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pp. 0–0, 2019.

Liwei Sun, Junjie Zhang, Jia Li, Yueming Wang, and Dan Zeng. Sdfc dataset: a large-scale benchmark dataset for hyperspectral image classification. *Optical and Quantum Electronics*, 55(2):173, 2023.

Romain Thoreau, Laurent Risser, Véronique Achard, Béatrice Berthelot, and Xavier Briottet. Toulouse hyperspectral data set: A benchmark data set to assess semi-supervised spectral representation learning and pixel-wise classification techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 212:323–337, 2024.

Di Wang, Meiqi Hu, Yao Jin, Yuchun Miao, Jiaqi Yang, Yichu Xu, Xiaolei Qin, Jiaqi Ma, Lingyu Sun, Chenxing Li, et al. Hypersigma: Hyperspectral intelligence comprehension foundation model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

Yue Wang, Ming Wen, Hailiang Zhang, Jinyu Sun, Qiong Yang, Zhimin Zhang, and Hongmei Lu. Hsimae: A unified masked autoencoder with large-scale pretraining for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17:14064–14079, 2024.

Yuchao Yang, Yulei Wang, Hongzhou Wang, Lifu Zhang, Enyu Zhao, Meiping Song, and Chunyan Yu. Spectral-enhanced sparse transformer network for hyperspectral super-resolution reconstruction. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17:17278–17291, 2024.

Naoto Yokoya and Akira Iwasaki. Airborne hyperspectral data over chikusei. *Space Appl. Lab., Univ. Tokyo, Tokyo, Japan, Tech. Rep. SAL-2016-05-27*, 5(5):5, 2016.

Dabing Yu and Zheng Gao. Volformer: Explore more comprehensive cube interaction for hyperspectral image restoration and beyond. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28091–28101. IEEE, 2025.

Roberta H. Yuhas, Alexander F. H. Goetz, and Joseph W. Boardman. Determination of semi-arid landscape endmembers and seasonal trends using convex geometry spectral unmixing techniques. In *JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop*, volume 1, pp. 147–149, 1992.

Mingjin Zhang, Chi Zhang, Qiming Zhang, Jie Guo, Xinbo Gao, and Jing Zhang. Essaformer: Efficient transformer for hyperspectral image super-resolution. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 23016–23027. IEEE, 2023.

Qiang Zhang, Yushuai Dong, Yaming Zheng, Haoyang Yu, Meiping Song, Lifu Zhang, and Qiangqiang Yuan. Three-dimension spatial–spectral attention transformer for hyperspectral image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–13, 2024.

Yingkai Zhang, Tao Zhang, Jing Nie, and Ying Fu. Real noise decoupling for hyperspectral image denoising. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 12925–12933, 2026.

Yanfei Zhong, Xin Hu, Chang Luo, Xinyu Wang, Ji Zhao, and Liangpei Zhang. Whu-hi: Uav-borne hyperspectral with high spatial resolution (h2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with crf. *Remote Sensing of Environment*, 250:112012, 2020.

# A APPENDIX

This appendix material provides additional documentation for the HSRest benchmark, including EO-1 Hyperion source product details, scene-level acquisition metadata, geographic footprints, representative qualitative results, and the release package structure. These materials support reproducibility by documenting patch-level provenance and the benchmark assets released with HSRest.

## A.1 EO-1 HYPERION SOURCE PRODUCTS

HSRest is constructed from Level-1Gst terrain-corrected radiance products acquired by the Hyperion pushbroom hyperspectral sensor onboard NASA's Earth Observing-1 (EO-1) satellite. EO-1 operated from November 2000 until decommissioning in March 2017. Hyperion records 242 spectral bands from approximately $0.357$ to $2.576$ $\mu$m at $\sim$10 nm spectral resolution and 30 m spatial resolution. Level-1Gst products preserve georeferencing information and are therefore suitable for constructing traceable hyperspectral restoration patches.

Table 5: Representative EO-1 Hyperion source-scene metadata used in HSRest, including region, scene identifier, acquisition date, image size, UTM zone, solar elevation, and sensor look angle.

| **Region** | **Scene ID** | **Date** | **Size** | **UTM** | **Sun Elev.** | **Look Angle** |
|---|---|---|---|---|---|---|
| France | EO1H1950262015271110K4 | 2015-09-28 | $1961 \times 6831$ | 32 | 44.08 | -14.09 |
| California | EO1H0420342015220110KF | 2015-08-08 | $1151 \times 3221$ | 11 | 46.77 | 10.95 |
| Spain | EO1H1990322015177110K2 | 2015-06-26 | $981 \times 3541$ | 30 | 51.311330 | 2.3696 |
| Germany | EO1H1920242015157110PF | 2015-06-06 | $1201 \times 3251$ | 33 | 50.072997 | 7.1723 |
| Hungary | EO1H1860262006023110KF | 2006-01-23 | $1061 \times 3281$ | 34 | 19.304275 | -3.6241 |
| Ireland | EO1H2060232003150110PY | 2003-05-30 | $1041 \times 3151$ | 29 | 56.386781 | -0.3084 |
| Poland | EO1H1920232003292110PZ | 2003-10-19 | $1081 \times 2931$ | 33 | 25.540972 | -0.4527 |

## A.2 SCENE-LEVEL ACQUISITION METADATA

Tables 6 and 7 provide detailed acquisition information for representative EO-1 Hyperion products used in HSRest. These tables document acquisition time windows, image dimensions, UTM zones, solar geometry, sensor viewing angles, and geographic corner coordinates. Here, UL, UR, LL, and LR denote upper-left, upper-right, lower-left, and lower-right image corners, respectively.

Table 6: Scene-level acquisition metadata for representative EO-1 Hyperion products used in HSRest, including acquisition time window, image dimensions, and UTM zone.

| Sample ID | Region | Scene ID | Acquisition Date | Start Time | End Time | Width | Height | UTM Zone |
|---|---|---|---|---|---|---|---|---|
| (1) | France | EO1H1950262015271110K4 | 2015-09-28 | 08:32:57 | 08:37:32 | 1961 | 6831 | 32 |
| (2) | France | EO1H1990292017069110K3 | 2017-03-10 | 08:08:19 | 08:12:39 | 1201 | 3451 | 31 |
| (3) | France | EO1H1950272015320110T3 | 2015-11-16 | 08:34:19 | 08:38:39 | 1161 | 3501 | 32 |
| (4) | France | EO1H1980302016211110KB | 2016-07-29 | 08:19:50 | 08:24:09 | 1101 | 3521 | 31 |
| (5) | California | EO1H0420342015220110KF | 2015-08-08 | 17:10:56 | 17:15:15 | 1151 | 3221 | 11 |
| (6) | California | EO1H0430342017071110KF | 2017-03-12 | 15:55:40 | 15:59:58 | 871 | 3271 | 10 |
| (7) | California | EO1H0440342013247110K9 | 2013-09-04 | 18:19:43 | 18:24:02 | 1211 | 3471 | 10 |
| (8) | Spain | EO1H1990322015177110K2 | 2015-06-26 | 09:14:46 | 09:19:05 | 981 | 3541 | 30 |
| (9) | Spain | EO1H1990322017058110K2 | 2017-02-27 | 08:07:56 | 08:12:15 | 971 | 3531 | 30 |
| (10) | Spain | EO1H2020342016359110KF | 2016-12-24 | 08:33:43 | 08:38:02 | 891 | 3311 | 29 |

Table 7: Solar geometry, sensor look angle, and geographic footprint coordinates for representative EO-1 Hyperion scenes. UL, UR, LL, and LR denote the four image corners.

| Sample ID | Sun Azimuth | Sun Elevation | Look Angle | UL Lat | UL Lon | UR Lat | UR Lon | LL Lat | LL Lon | LR Lat | LR Lon |
|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) | 151.839203 | 44.089302 | -14.0970 | 48.68991 | 7.007979 | 48.70095 | 7.806720 | 46.84739 | 7.077147 | 46.85775 | 7.848174 |
| (2) | 115.342000 | 18.876000 | 7.2101 | 44.22128 | 0.710205 | 44.22942 | 1.160644 | 43.29011 | 0.745458 | 43.29800 | 1.188972 |
| (3) | 141.496000 | 14.663060 | -4.5554 | 48.37939 | 48.385210 | 48.38521 | 7.810068 | 47.43506 | 7.370143 | 47.44069 | 7.831530 |
| (4) | 100.452300 | 37.868690 | -17.0300 | 44.08460 | 0.906588 | 44.09141 | 1.318578 | 43.13439 | 0.939302 | 43.14098 | 1.344862 |
| (5) | 107.578700 | 46.770870 | 10.9510 | 37.78822 | -119.871000 | 37.79711 | -119.479000 | 36.91856 | -119.838000 | 36.92718 | -119.451000 |
| (6) | 109.338300 | 18.837270 | -13.4340 | 37.83374 | -120.592000 | 37.82730 | -120.296000 | 36.95025 | -120.620000 | 36.94402 | -120.327000 |
| (7) | 134.990400 | 51.229170 | 22.3420 | 37.84621 | -122.440000 | 37.84353 | -122.027000 | 36.90791 | -122.447000 | 36.90532 | -122.039000 |
| (8) | 101.868700 | 51.311330 | 2.3696 | 39.97973 | -0.458720 | 39.97167 | -0.114800 | 39.02373 | -0.493320 | 39.01594 | -0.154070 |
| (9) | 115.626600 | 16.281180 | 3.1164 | 39.91800 | -0.478550 | 39.91010 | -0.138440 | 38.96468 | -0.512710 | 38.95705 | -0.177200 |
| (10) | 128.464200 | 8.485979 | -7.3944 | 38.13627 | -6.715660 | 38.12995 | -6.411280 | 37.24195 | -6.742940 | 37.23583 | -6.442180 |

### A.3 RELEASE PACKAGE STRUCTURE

Table 8 summarizes the HSRest release package. The package is organized to enable reproducibility of the results reported in the main paper, including image data, metadata, fixed splits, degradation generators, evaluation scripts, baseline configurations, and dataset documentation.

Table 8: HSRest release package structure, including image data, metadata, fixed split files, degradation generators, evaluation scripts, and documentation for reproducible hyperspectral restoration benchmarking.

| **Component** | **Description** |
|---|---|
| `data/` | 242-band patches, 175-band good-band patches, and removed-band subsets. |
| `metadata/` | Patch-level provenance CSV files, metadata schema in JSON format, and Anderson category definitions. |
| `splits/` | Fixed $70/10/20$ train/validation/test files, $10/20/30\%$ low-resource subsets, and Groupwise A/B/C protocols. |
| `degradation/` | Deterministic degradation generators for Gaussian denoising, structured inpainting, and spatial downsampling, with parameter documentation. |
| `eval/` | Evaluation scripts for PSNR, SSIM, SAM, SID, spectral correlation, and baseline configuration files. |
| `dataset_card.md` | Dataset documentation covering source imagery, preprocessing, intended use, limitations, citation, and release structure. |

Table 9: Train–validation–test configurations used in HSRest experiments. For group-wise experiments, one geographic group is used for training and the remaining two groups are used for testing. For percentage-wise experiments, 10%, 20%, or 30% of the training set is used, while the validation and test sets remain fixed.

| Protocol | Train | Validation | Test |
|---|---|---|---|
| **Full** | Train set<br>16,213 | Val set<br>2,369 | Test set<br>4,646 |
| **Group-wise** | A<br>9,244 | – | B+C<br>13,984 |
| | B<br>6,992 | – | A+C<br>16,236 |
| | C<br>6,992 | – | A+B<br>16,236 |
| **Percentage-wise** | 10% of train set<br>1,615 | Val set<br>2,369 | Test set<br>4,646 |
| | 20% of train set<br>3,243 | Val set<br>2,369 | Test set<br>4,646 |
| | 30% of train set<br>4,863 | Val set<br>2,369 | Test set<br>4,646 |

Table 10: Geographic composition and data partitioning of HSRest. Each country belongs exclusively to one geographic group (A–C) and one data split. The low-resource column indicates the percentage-wise training setting in which the corresponding country is included. Percentages denote the share of the total 23,228 patches.

| **Group**<br>Continent-disjoint | **Continent** | **Country** | **Patches** | **Share (%)** | **Split**<br>Country-disjoint | **Low-resource**<br>Country-disjoint |
|---|---|---|---|---|---|---|
| **A** | **Asia** | India | 1,926 | 8.29 | Train | – |
| | | Turkey | 1,053 | 4.53 | Train | 30% |
| | | Mongolia | 1,026 | 4.42 | Train | 10% |
| | | Indonesia | 1,023 | 4.40 | Train | – |
| | | Japan | 972 | 4.18 | Train | 30% |
| | | South Korea | 322 | 1.39 | Test | – |
| | **Africa** | South Africa | 1,351 | 5.82 | Train | – |
| | | Ethiopia | 1,236 | 5.32 | Train | 20% |
| | | Kenya | 335 | 1.44 | Test | – |
| **B** | **North America** | California | 1,939 | 8.35 | Train | – |
| | | Mexico | 1,346 | 5.79 | Train | 30% |
| | | Canada (BC) | 589 | 2.54 | Train | 10% |
| | **Europe** | Spain | 656 | 2.82 | Train | 20% |
| | | Germany | 523 | 2.25 | Train | 30% |
| | | Ukraine | 499 | 2.15 | Train | – |
| | | Norway & Sweden | 358 | 1.54 | Train | 30% |
| | | Hungary | 310 | 1.33 | Train | 30% |
| | | Ireland | 302 | 1.30 | Test | – |
| | | France | 301 | 1.30 | Train | 30% |
| | | Poland | 169 | 0.73 | Val | – |
| **C** | **Oceania** | Australia | 2,455 | 10.57 | Test | – |
| | | New Zealand | 1,257 | 5.41 | Val | – |
| | **South America** | Chile | 1,232 | 5.30 | Test | – |
| | | Brazil | 1,105 | 4.76 | Train | – |
| | | Argentina | 943 | 4.06 | Val | – |

**Overall**
Total: 23,228 patches | Train: 16,213 (69.80%)
Val: 2,369 (10.20%) | Test: 4,646 (20.00%)
**Low-resource subsets:**
10%: 2 countries / 1,615 patches | 20%: 3 countries / 3,243 patches
30%: 7 countries / 4,863 patches

Table 11: Summary of the restoration models evaluated in HSRest, grouped by restoration task. The table reports the model abbreviation, methodological type, publication venue, and publication year.

| Task | Model | Abbr. | Type | Venue | Year |
|---|---|---|---|---|---|
| **Denoising** | VolFormer(Yu & Gao (2025)) | V | Transformer | CVPR | 2025 |
| | RND(Zhang et al. (2026)) | R | Denoising framework | AAAI | 2026 |
| | TDSAT(Zhang et al. (2024)) | T | Transformer | TGRS | 2024 |
| | HSDT(Lai et al. (2023)) | H | Transformer | ICCV | 2023 |
| | SST(Li et al. (2023)) | S | Transformer | AAAI | 2023 |
| **Super-resolution** | VolFormer(Yu & Gao (2025)) | V | Transformer | CVPR | 2025 |
| | ESSAformer(Zhang et al. (2023)) | E | Transformer | ICCV | 2023 |
| | PSRT(Deng et al. (2023)) | P | Transformer | TGRS | 2023 |
| | SEST(Yang et al. (2024)) | SE | Transformer | J-STARS | 2024 |
| | Fusformer(Hu et al. (2022)) | F | Transformer | GRSL | 2022 |
| **Inpainting** | Deep Hyperspectral Prior(Sidorov & Yngve Hardeberg (2019)) | D2/D3 | Deep Image Prior | ICCVW | 2019 |
| | HLRTF(Luo et al. (2022)) | H2/H3 | Low-rank tensor factorization | CVPR | 2022 |
| | HIR-Diff(Pang et al. (2024)) | HR | Diffusion | CVPR | 2024 |
| | R-DLRHyIN(Niresi & Chi (2023)) | R2/R3 | Untrained CNN + low-rank regularization | GRSL | 2023 |

Table 12: Quantitative comparison of hyperspectral image denoising methods under different training regimes and noise levels. MPSNR and MSSIM are higher-is-better (↑), while RMSE and average inference time per sample are lower-is-better (↓). The best and second-best results within each training setting and noise level are shown in deep blue and light blue, respectively.

| Method | Venue–Year | Type | $\sigma = 10$ | | | | $\sigma = 30$ | | | | $\sigma = 50$ | | | | $\sigma = 70$ | | | | $\sigma = 90$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ |
| **Full Data** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | **50.794** | **0.987** | **0.005** | 1.863 | **42.895** | **0.940** | 0.011 | 1.827 | 40.807 | **0.922** | 0.016 | 1.944 | **40.826** | **0.918** | **0.016** | 2.341 | **40.108** | **0.897** | **0.016** | 2.794 |
| RND | AAAI 2026 | Supervised | 44.055 | 0.920 | 0.009 | 0.783 | 42.846 | 0.922 | **0.010** | 0.786 | **41.197** | 0.903 | **0.013** | 0.766 | 27.154 | 0.215 | 0.140 | 0.792 | 27.154 | 0.215 | 0.140 | 0.801 |
| TDSAT | TGRS 2024 | Supervised | 25.772 | 0.770 | 0.006 | 0.523 | 21.688 | 0.671 | **0.010** | 0.513 | 17.152 | 0.536 | 0.017 | 0.517 | 17.353 | 0.506 | 0.017 | 0.900 | 17.255 | 0.500 | 0.017 | 0.525 |
| SST | AAAI 2023 | Supervised | 34.602 | 0.685 | 0.027 | **0.094** | 28.852 | 0.505 | 0.071 | **0.101** | 25.625 | 0.425 | 0.114 | **0.179** | 23.390 | 0.369 | 0.156 | **0.152** | 21.997 | 0.343 | 0.198 | **0.128** |
| HSDT | ICCV 2023 | Supervised | 39.518 | 0.820 | 0.015 | 4.670 | 29.219 | 0.530 | 0.066 | 2.950 | 26.712 | 0.469 | 0.086 | 0.214 | 22.391 | 0.371 | 0.153 | 0.214 | 19.517 | 0.292 | 0.192 | 0.214 |
| **Low-Resource Fine-Tuning** | | | | | | | | | | | | | | | | | | | | | | |
| **10% of Train** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 37.455 | 0.853 | 0.018 | 2.020 | 35.802 | **0.782** | 0.023 | 1.574 | **35.389** | **0.758** | **0.023** | 1.721 | **34.861** | **0.751** | **0.026** | 1.893 | **40.494** | **0.897** | **0.016** | 3.247 |
| RND | AAAI 2026 | Supervised | 38.712 | 0.845 | 0.014 | 0.581 | **35.963** | 0.773 | **0.021** | 1.646 | 32.510 | 0.695 | 0.031 | 1.174 | 27.522 | 0.217 | 0.137 | 0.596 | 27.429 | 0.217 | 0.137 | 1.852 |
| TDSAT | TGRS 2024 | Supervised | 3.396 | 0.199 | 0.071 | 0.465 | 9.891 | 0.345 | 0.033 | 0.226 | 7.794 | 0.294 | 0.041 | 0.225 | 3.577 | 0.228 | 0.071 | 0.585 | 4.905 | 0.206 | 0.059 | 1.887 |
| SST | AAAI 2023 | Supervised | 29.790 | 0.514 | 0.039 | **0.138** | 22.786 | 0.212 | 0.089 | **0.134** | 19.269 | 0.103 | 0.135 | **0.070** | 17.024 | 0.069 | 0.176 | 0.745 | 15.419 | 0.052 | 0.212 | **0.143** |
| HSDT | ICCV 2023 | Supervised | **43.628** | **0.909** | **0.009** | 0.214 | 35.854 | 0.758 | 0.024 | 0.214 | 30.624 | 0.600 | 0.045 | 0.339 | 28.013 | 0.475 | 0.064 | **0.214** | 23.445 | 0.316 | 0.101 | 0.213 |
| **20% of Train** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 41.662 | **0.923** | 0.012 | 2.134 | **41.105** | **0.898** | **0.014** | 2.146 | **40.573** | **0.886** | **0.016** | 2.278 | **39.989** | **0.877** | **0.018** | 2.421 | **39.605** | **0.869** | **0.019** | 2.557 |
| RND | AAAI 2026 | Supervised | 40.401 | 0.845 | 0.013 | 0.583 | 27.546 | 0.217 | 0.137 | 1.846 | 27.543 | 0.217 | 0.137 | 0.576 | 27.548 | 0.217 | 0.137 | 1.533 | 27.548 | 0.217 | 0.137 | 1.526 |
| TDSAT | TGRS 2024 | Supervised | 15.286 | 0.532 | 0.017 | 0.480 | 12.331 | 0.426 | 0.025 | 0.225 | 3.259 | 0.139 | 0.074 | 0.349 | 8.775 | 0.274 | 0.037 | 0.864 | 7.433 | 0.219 | 0.044 | 0.692 |
| SST | AAAI 2023 | Supervised | 30.805 | 0.555 | 0.035 | 0.222 | 23.544 | 0.237 | 0.083 | **0.130** | 20.017 | 0.136 | 0.124 | **0.070** | 17.713 | 0.087 | 0.162 | 0.749 | 16.102 | 0.059 | 0.196 | **0.140** |
| HSDT | ICCV 2023 | Supervised | **41.779** | 0.888 | **0.011** | **0.214** | 31.362 | 0.540 | 0.034 | 0.339 | 27.195 | 0.350 | 0.057 | 0.339 | 22.742 | 0.149 | 0.100 | **0.214** | 22.089 | 0.196 | 0.108 | 0.213 |
| **30% of Train** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 40.330 | **0.928** | 0.017 | 1.638 | **41.473** | **0.920** | **0.013** | 2.684 | **41.296** | **0.917** | **0.014** | 2.913 | **40.494** | **0.897** | **0.016** | 3.247 | **40.484** | **0.894** | **0.015** | 3.156 |
| RND | AAAI 2026 | Supervised | **43.310** | 0.898 | **0.009** | 1.585 | 39.004 | 0.824 | 0.015 | 1.526 | 37.314 | 0.766 | 0.018 | 0.594 | 27.549 | 0.217 | 0.137 | 1.659 | 27.550 | 0.217 | 0.137 | 1.180 |
| TDSAT | TGRS 2024 | Supervised | 18.619 | 0.612 | 0.012 | 0.257 | 15.232 | 0.523 | 0.018 | 0.226 | 12.269 | 0.404 | 0.026 | 0.501 | 8.972 | 0.333 | 0.036 | 0.880 | 11.693 | 0.360 | 0.027 | 0.563 |
| SST | AAAI 2023 | Supervised | 16.447 | 0.048 | 0.187 | **0.141** | 23.943 | 0.234 | 0.079 | **0.139** | 20.291 | 0.138 | 0.120 | **0.071** | 18.143 | 0.086 | 0.154 | 0.774 | 16.447 | 0.048 | 0.187 | **0.141** |
| HSDT | ICCV 2023 | Supervised | 42.129 | 0.868 | 0.011 | 0.214 | 33.770 | 0.659 | 0.029 | 0.339 | 26.205 | 0.403 | 0.070 | 0.214 | 23.602 | 0.300 | 0.125 | **0.213** | 23.231 | 0.353 | 0.111 | 0.214 |
| **Groupwise Training** | | | | | | | | | | | | | | | | | | | | | | |
| **Train A → Test BC** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | **47.387** | **0.962** | **0.007** | 1.935 | **45.160** | **0.948** | **0.009** | 1.977 | **43.723** | **0.931** | **0.012** | 1.985 | **43.261** | **0.923** | **0.013** | 1.986 | **43.015** | **0.914** | **0.014** | 1.908 |
| RND | AAAI 2026 | Supervised | 44.458 | 0.926 | 0.008 | 0.707 | 41.003 | 0.876 | 0.013 | 1.171 | 38.869 | 0.836 | 0.018 | 0.588 | 29.401 | 0.252 | 0.128 | 0.586 | 29.401 | 0.252 | 0.128 | 0.585 |
| TDSAT | TGRS 2024 | Supervised | 19.106 | 0.662 | 0.010 | 0.516 | 17.786 | 0.574 | 0.013 | 0.515 | 13.526 | 0.512 | 0.019 | 0.513 | 14.506 | 0.464 | 0.018 | 0.863 | 12.384 | 0.388 | 0.023 | 0.331 |
| SST | AAAI 2023 | Supervised | 32.389 | 0.618 | 0.031 | **0.096** | 24.804 | 0.297 | 0.074 | **0.099** | 21.296 | 0.167 | 0.111 | **0.114** | 19.040 | 0.113 | 0.144 | **0.129** | 17.404 | 0.083 | 0.174 | **0.121** |
| HSDT | ICCV 2023 | Supervised | 39.889 | 0.849 | 0.014 | 0.710 | 31.769 | 0.553 | 0.031 | 0.725 | 25.030 | 0.204 | 0.068 | 0.676 | 24.116 | 0.201 | 0.075 | 0.744 | 24.238 | 0.239 | 0.082 | 0.772 |
| **Train B → Test CA** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | **46.298** | **0.956** | **0.008** | 1.947 | **44.642** | **0.935** | **0.010** | 1.817 | **43.502** | **0.921** | **0.012** | 1.972 | **42.816** | **0.909** | **0.014** | 1.968 | **42.618** | **0.904** | **0.015** | 1.914 |
| RND | AAAI 2026 | Supervised | 44.495 | 0.934 | **0.008** | 0.809 | 36.867 | 0.846 | 0.017 | 0.588 | 35.909 | 0.834 | 0.020 | 0.589 | 26.225 | 0.195 | 0.144 | 0.586 | 26.225 | 0.195 | 0.144 | 0.586 |
| TDSAT | TGRS 2024 | Supervised | 22.795 | 0.736 | 0.009 | 0.519 | 15.624 | 0.603 | 0.020 | 0.513 | 17.767 | 0.575 | 0.016 | 0.516 | 16.805 | 0.534 | 0.019 | 0.907 | 14.676 | 0.465 | 0.024 | 0.618 |
| SST | AAAI 2023 | Supervised | 32.086 | 0.624 | 0.029 | **0.096** | 24.442 | 0.266 | 0.070 | **0.102** | 20.878 | 0.151 | 0.106 | **0.189** | 18.605 | 0.096 | 0.138 | **0.154** | 16.997 | 0.067 | 0.166 | **0.123** |
| HSDT | ICCV 2023 | Supervised | 40.745 | 0.832 | 0.011 | 0.605 | 35.418 | 0.660 | 0.024 | 0.794 | 26.647 | 0.327 | 0.071 | 0.694 | 24.876 | 0.273 | 0.106 | 0.610 | 20.923 | 0.102 | 0.121 | 0.788 |
| **Train C → Test AB** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | **48.719** | **0.971** | **0.006** | 1.879 | **45.531** | **0.948** | **0.010** | 1.906 | **44.919** | **0.939** | **0.011** | 1.913 | **43.878** | **0.920** | **0.013** | 1.889 | **43.869** | **0.914** | **0.013** | 1.931 |
| RND | AAAI 2026 | Supervised | 41.058 | 0.888 | 0.012 | 1.347 | 39.871 | 0.888 | 0.013 | 0.589 | 27.696 | 0.225 | 0.136 | 0.588 | 27.696 | 0.225 | 0.136 | 0.582 | 27.696 | 0.225 | 0.136 | 0.592 |
| TDSAT | TGRS 2024 | Supervised | 14.105 | 0.505 | 0.021 | 0.515 | 16.728 | 0.534 | 0.016 | 0.514 | 13.713 | 0.513 | 0.022 | 0.741 | 15.691 | 0.500 | 0.019 | 1.550 | 15.447 | 0.453 | 0.021 | 1.011 |
| SST | AAAI 2023 | Supervised | 29.616 | 0.453 | 0.037 | **0.099** | 24.038 | 0.225 | 0.072 | **0.095** | 20.561 | 0.104 | 0.108 | **0.188** | 18.490 | 0.075 | 0.138 | **0.204** | 16.909 | 0.057 | 0.166 | **0.130** |
| HSDT | ICCV 2023 | Supervised | 41.340 | 0.855 | 0.010 | 0.671 | 33.644 | 0.536 | 0.024 | 0.775 | 29.686 | 0.361 | 0.038 | 0.596 | 23.111 | 0.073 | 0.081 | 0.669 | 22.561 | 0.042 | 0.081 | 0.768 |

*Time denotes average inference time per sample. Metric order within each noise level: MPSNR, MSSIM, RMSE, and average time per sample. Best and second-best values are determined independently within each training setting across the five compared models.*

Table 13: Quantitative comparison of hyperspectral image denoising methods under different training regimes and noise levels. Cross-correlation is higher-is-better (↑), while SAM, SID, and ERGAS are lower-is-better (↓). The best and second-best results within each training setting and noise level are shown in deep blue and light blue, respectively.

| Method | Venue–Year | Type | $\sigma = 10$ | | | | $\sigma = 30$ | | | | $\sigma = 50$ | | | | $\sigma = 70$ | | | | $\sigma = 90$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ |
| **Full Data** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 1.841 | **0.002** | **34.963** | 0.878 | 3.979 | **0.002** | **43.433** | 0.768 | 4.499 | **0.004** | **47.104** | 0.716 | 4.613 | **0.005** | **37.837** | 0.679 | 4.987 | **0.006** | **53.656** | 0.653 |
| RND | AAAI 2026 | Supervised | 3.676 | 0.320 | 80.619 | 0.725 | 3.623 | 0.538 | 187.329 | 0.730 | 4.501 | 0.539 | 126.640 | 0.646 | 90.000 | 1.027 | 107.130 | 0.000 | 90.000 | 1.031 | 107.129 | 0.000 |
| TDSAT | TGRS 2024 | Supervised | 2.611 | 0.038 | 142.157 | 0.787 | 3.901 | 0.072 | 220.685 | 0.692 | 5.946 | 0.114 | 548.704 | 0.564 | 6.297 | 0.148 | 403.722 | 0.528 | 5.894 | 0.194 | 404.458 | 0.520 |
| SST | AAAI 2023 | Supervised | 12.661 | 1.124 | 1669.151 | 0.934 | 28.069 | 2.532 | 4979.628 | 0.774 | 37.319 | 4.662 | 8289.972 | 0.659 | 44.458 | 7.241 | 11592.745 | 0.575 | 49.747 | 7.823 | 14817.531 | 0.504 |
| HSDT | ICCV 2023 | Supervised | **0.121** | 0.005 | 929.897 | **0.983** | **0.296** | 0.031 | 3245.777 | **0.852** | **0.319** | 0.055 | 4460.236 | **0.884** | **0.436** | 0.142 | 8340.847 | **0.804** | **0.482** | 0.126 | 10012.375 | **0.795** |
| **Low-Resource Fine-Tuning** | | | | | | | | | | | | | | | | | | | | | | |
| **10% of Train** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 8.212 | 0.004 | 551.017 | 0.668 | 10.606 | **0.006** | **662.682** | 0.545 | 10.993 | **0.006** | **691.935** | 0.519 | 11.961 | **0.007** | 742.419 | 0.721 | 7.297 | **0.002** | 298.906 | 0.671 |
| RND | AAAI 2026 | Supervised | 8.440 | 0.324 | 535.681 | 0.622 | 11.001 | 0.444 | 740.767 | 0.511 | 12.869 | 0.349 | 1573.049 | 0.439 | 90.000 | 1.087 | **112.012** | 0.000 | 89.997 | 1.099 | **121.978** | 0.000 |
| TDSAT | TGRS 2024 | Supervised | 33.976 | 2.931 | 3094.004 | 0.284 | 17.191 | 1.071 | 1463.275 | 0.408 | 18.438 | 0.974 | 1852.399 | 0.364 | 21.416 | 1.100 | 2961.471 | 0.325 | 20.784 | 0.945 | 2797.399 | 0.266 |
| SST | AAAI 2023 | Supervised | 19.915 | 1.756 | 1509.632 | 0.851 | 33.914 | 3.944 | 3237.236 | 0.701 | 43.690 | 6.226 | 5032.855 | 0.586 | 49.897 | 7.968 | 6728.672 | 0.495 | 53.587 | 9.094 | 8383.303 | 0.432 |
| HSDT | ICCV 2023 | Supervised | **0.123** | **0.003** | **399.837** | **0.947** | **0.215** | 0.014 | 1063.872 | **0.926** | **0.306** | 0.034 | 2087.132 | **0.900** | **0.391** | 0.062 | 2717.259 | **0.858** | **0.473** | 0.093 | 4824.428 | **0.674** |
| **20% of Train** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 6.254 | **0.002** | 254.983 | 0.734 | 7.330 | **0.003** | 365.658 | 0.691 | 7.731 | **0.003** | 361.767 | 0.633 | 8.352 | **0.003** | 310.637 | 0.698 | 8.912 | **0.003** | 269.921 | 0.618 |
| RND | AAAI 2026 | Supervised | 9.142 | 0.633 | **180.995** | 0.649 | 90.000 | 1.085 | **108.381** | 0.000 | 90.000 | 1.084 | **109.035** | 0.000 | 90.000 | 1.083 | **108.512** | 0.000 | 90.000 | 1.083 | **108.502** | 0.000 |
| TDSAT | TGRS 2024 | Supervised | 10.394 | 0.466 | 893.265 | 0.579 | 14.679 | 1.189 | 792.960 | 0.481 | 33.531 | 3.095 | 1795.158 | 0.208 | 15.417 | 0.498 | 1773.995 | 0.296 | 19.455 | 1.334 | 1386.708 | 0.246 |
| SST | AAAI 2023 | Supervised | 18.947 | 1.847 | 1242.520 | 0.868 | 33.076 | 4.233 | 2912.500 | 0.721 | 41.360 | 5.811 | 4556.559 | 0.602 | 47.032 | 7.068 | 6209.693 | 0.516 | 51.538 | 8.277 | 7757.561 | 0.462 |
| HSDT | ICCV 2023 | Supervised | **0.134** | 0.003 | 412.165 | **0.952** | **0.307** | 0.023 | 1264.731 | **0.911** | **0.448** | 0.047 | 2715.138 | **0.862** | **0.637** | 0.102 | 3439.623 | **0.754** | **0.681** | 0.108 | 3897.994 | **0.712** |
| **30% of Train** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 7.120 | **0.002** | 268.779 | 0.754 | 7.209 | **0.002** | 256.290 | 0.712 | 7.611 | **0.001** | **175.523** | 0.707 | 7.297 | **0.002** | 298.906 | 0.716 | 8.237 | **0.002** | 227.342 | 0.657 |
| RND | AAAI 2026 | Supervised | 7.165 | 0.385 | **167.906** | 0.725 | 9.469 | 0.762 | **159.551** | 0.611 | 9.820 | 0.598 | 394.772 | 0.525 | 90.000 | 1.083 | **108.363** | 0.000 | 90.000 | 1.083 | **108.289** | 0.000 |
| TDSAT | TGRS 2024 | Supervised | 8.845 | 0.509 | 558.853 | 0.646 | 10.115 | 0.409 | 744.703 | 0.567 | 12.212 | 0.442 | 981.500 | 0.452 | 13.653 | 0.361 | 1725.775 | 0.372 | 13.689 | 0.740 | 954.137 | 0.408 |
| SST | AAAI 2023 | Supervised | 52.551 | 8.336 | 7495.180 | 0.495 | 32.132 | 4.185 | 2785.341 | 0.745 | 38.997 | 4.844 | 4483.289 | 0.615 | 46.818 | 6.923 | 5931.476 | 0.545 | 52.551 | 8.336 | 7495.180 | 0.495 |
| HSDT | ICCV 2023 | Supervised | **0.137** | 0.004 | 485.800 | **0.953** | **0.259** | 0.016 | 1330.441 | **0.927** | **0.369** | 0.048 | 2897.971 | **0.829** | **0.433** | 0.116 | 4974.587 | **0.790** | **0.432** | 0.101 | 5038.641 | **0.818** |
| **Groupwise Training** | | | | | | | | | | | | | | | | | | | | | | |
| **Train A → Test BC** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 5.138 | **0.001** | **159.803** | 0.768 | 6.042 | **0.001** | **126.859** | 0.725 | 6.828 | **0.001** | **108.527** | 0.681 | 7.162 | **0.001** | 130.778 | 0.665 | 7.042 | **0.001** | 126.047 | 0.632 |
| RND | AAAI 2026 | Supervised | 5.702 | 0.261 | 188.934 | 0.709 | 7.843 | 0.529 | 176.985 | 0.610 | 9.771 | 0.602 | 247.054 | 0.483 | 90.000 | 1.269 | **108.253** | 0.000 | 90.000 | 1.269 | **108.252** | 0.000 |
| TDSAT | TGRS 2024 | Supervised | 5.937 | 0.125 | 469.726 | 0.699 | 8.350 | 0.524 | 446.045 | 0.617 | 9.279 | 0.373 | 782.379 | 0.564 | 10.243 | 0.551 | 646.093 | 0.502 | 10.876 | 0.416 | 812.735 | 0.421 |
| SST | AAAI 2023 | Supervised | 17.332 | 1.632 | 1147.991 | 0.903 | 32.080 | 3.711 | 2880.916 | 0.747 | 42.110 | 5.807 | 4518.455 | 0.632 | 48.687 | 7.208 | 6135.332 | 0.536 | 51.022 | 7.477 | 7699.247 | 0.478 |
| HSDT | ICCV 2023 | Supervised | **0.156** | 0.004 | 595.104 | **0.931** | **0.331** | 0.018 | 1235.143 | **0.873** | **0.563** | 0.062 | 3057.492 | **0.790** | **0.575** | 0.075 | 3332.492 | **0.777** | **0.589** | 0.078 | 3877.643 | **0.762** |
| **Train B → Test CA** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 8.548 | **0.001** | **193.936** | 0.784 | 7.382 | **0.001** | **211.710** | 0.738 | 7.943 | **0.001** | **250.811** | 0.649 | 9.544 | **0.002** | 198.748 | 0.659 | 9.690 | **0.002** | 212.328 | 0.637 |
| RND | AAAI 2026 | Supervised | 4.206 | 0.114 | 194.072 | 0.796 | 7.109 | 0.310 | 442.599 | 0.637 | 7.018 | 0.264 | 375.569 | 0.617 | 90.000 | 1.109 | **107.832** | 0.000 | 90.000 | 1.109 | **107.839** | -0.000 |
| TDSAT | TGRS 2024 | Supervised | 4.989 | 0.191 | 273.961 | 0.767 | 7.138 | 0.176 | 640.257 | 0.653 | 7.349 | 0.244 | 474.335 | 0.616 | 8.440 | 0.323 | 594.835 | 0.577 | 9.309 | 0.355 | 738.815 | 0.509 |
| SST | AAAI 2023 | Supervised | 14.221 | 1.218 | 792.515 | 0.921 | 27.962 | 3.091 | 1973.465 | 0.789 | 36.350 | 4.339 | 3125.378 | 0.674 | 42.532 | 5.530 | 4254.731 | 0.585 | 46.635 | 6.398 | 5339.177 | 0.525 |
| HSDT | ICCV 2023 | Supervised | **0.172** | 0.004 | 479.458 | **0.952** | **0.263** | 0.016 | 1258.719 | **0.905** | **0.495** | 0.071 | 3329.814 | **0.895** | **0.520** | 0.118 | 4891.358 | **0.811** | **0.743** | 0.145 | 5533.077 | **0.766** |
| **Train C → Test AB** | | | | | | | | | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 10.372 | **0.001** | 95.468 | 0.834 | 11.370 | **0.001** | **125.779** | 0.786 | 12.411 | **0.001** | 110.805 | 0.661 | 11.886 | **0.001** | 118.811 | 0.623 | 12.332 | **0.001** | 118.309 | 0.669 |
| RND | AAAI 2026 | Supervised | 4.229 | 0.209 | **62.249** | 0.758 | 4.720 | 0.091 | 347.761 | 0.678 | 90.000 | 1.311 | **107.719** | 0.000 | 90.000 | 1.311 | **107.713** | -0.000 | 90.000 | 1.311 | **107.717** | 0.000 |
| TDSAT | TGRS 2024 | Supervised | 9.344 | 0.612 | 562.690 | 0.644 | 6.972 | 0.265 | 533.853 | 0.581 | 8.221 | 0.145 | 759.163 | 0.581 | 7.378 | 0.138 | 556.060 | 0.549 | 8.089 | 0.202 | 518.525 | 0.490 |
| SST | AAAI 2023 | Supervised | 18.187 | 1.335 | 1391.796 | 0.926 | 30.081 | 2.965 | 2284.022 | 0.795 | 42.525 | 5.785 | 3379.415 | 0.667 | 48.337 | 7.326 | 4373.746 | 0.567 | 52.281 | 8.397 | 5413.717 | 0.480 |
| HSDT | ICCV 2023 | Supervised | **0.087** | 0.002 | 295.089 | **0.996** | **0.256** | 0.018 | 899.153 | **0.990** | **0.408** | 0.034 | 1299.432 | **0.977** | **0.809** | 0.078 | 3304.953 | **0.946** | **0.801** | 0.090 | 2191.323 | **0.839** |

*CC denotes cross-correlation. Metric order within each noise level: SAM, SID, ERGAS, and cross-correlation. Best and second-best values are determined independently within each training setting across the five compared models.*

Table 14: Quantitative comparison of hyperspectral image super-resolution methods under different training regimes and upscaling factors. MPSNR and MSSIM are higher-is-better (↑), while RMSE and average inference time per sample are lower-is-better (↓). The best and second-best results within each training setting and upscaling factor are shown in deep blue and light blue, respectively.

| Method | Venue–Year | Type | ×2 | | | | ×4 | | | | ×8 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ | MPSNR ↑ | MSSIM ↑ | RMSE ↓ | Time ↓ |
| **Full Data** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 34.351 | 0.877 | 0.054 | 1.337 | 32.634 | 0.813 | 0.059 | 1.643 | 32.122 | 0.762 | 0.068 | 1.493 |
| ESSAFORMER | ICCV 2023 | Supervised | 34.035 | 0.872 | 0.051 | 3.225 | 31.009 | 0.760 | 0.080 | 3.686 | 31.719 | 0.747 | 0.070 | 3.789 |
| PSRT | IEEE TGRS 2023 | Supervised | 34.760 | 0.855 | 0.065 | 3.512 | 33.329 | 0.816 | 0.059 | 3.563 | 33.140 | **0.788** | 0.062 | 3.481 |
| SEST | IEEE JSTARS 2024 | Supervised | 35.907 | 0.867 | 0.054 | 1.473 | 34.031 | 0.785 | 0.070 | 1.338 | 33.510 | 0.775 | **0.058** | 1.422 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | **37.098** | **0.889** | **0.041** | **0.829** | **35.673** | **0.836** | **0.054** | **0.482** | **34.415** | 0.785 | 0.059 | **0.551** |
| **Low-Resource Fine-Tuning** | | | | | | | | | | | | | | |
| **10% of Train** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 34.254 | 0.847 | 0.049 | 1.732 | 32.238 | 0.771 | 0.057 | 0.992 | 31.576 | 0.723 | 0.057 | 1.525 |
| ESSAFORMER | ICCV 2023 | Supervised | 30.363 | 0.742 | 0.055 | 3.431 | 29.361 | 0.707 | 0.062 | 3.342 | 28.209 | 0.607 | 0.066 | 3.455 |
| PSRT | IEEE TGRS 2023 | Supervised | 34.813 | 0.849 | **0.044** | 3.377 | 32.905 | 0.769 | 0.051 | 3.430 | 32.781 | 0.739 | 0.055 | 3.585 |
| SEST | IEEE JSTARS 2024 | Supervised | **38.166** | **0.865** | **0.044** | 7.643 | **36.032** | **0.804** | **0.050** | 7.665 | 33.806 | 0.733 | 0.056 | 8.143 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | 37.753 | 0.858 | 0.046 | **0.509** | 35.269 | 0.794 | 0.051 | **0.502** | **34.386** | **0.743** | **0.054** | **0.528** |
| **20% of Train** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 32.880 | 0.832 | 0.051 | 1.019 | 31.912 | 0.785 | 0.053 | 1.304 | 32.177 | 0.740 | 0.054 | 1.265 |
| ESSAFORMER | ICCV 2023 | Supervised | 31.956 | 0.786 | 0.051 | 3.037 | 29.995 | 0.715 | 0.062 | 2.631 | 28.662 | 0.663 | 0.072 | 3.279 |
| PSRT | IEEE TGRS 2023 | Supervised | 34.560 | 0.832 | 0.045 | 3.868 | 28.175 | 0.688 | 0.074 | 3.319 | 30.003 | 0.716 | 0.058 | 3.937 |
| SEST | IEEE JSTARS 2024 | Supervised | 37.127 | **0.874** | **0.041** | 0.638 | 34.922 | **0.813** | **0.047** | 0.922 | 34.446 | **0.756** | **0.050** | 1.390 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | **37.969** | 0.872 | 0.042 | **0.498** | **35.196** | 0.794 | 0.048 | **0.523** | **34.636** | 0.753 | 0.051 | **0.545** |
| **30% of Train** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 31.884 | 0.873 | 0.056 | 1.561 | 33.401 | 0.813 | 0.049 | 1.568 | 33.204 | 0.760 | 0.051 | 1.238 |
| ESSAFORMER | ICCV 2023 | Supervised | 32.828 | 0.840 | 0.048 | 3.903 | 31.764 | 0.762 | 0.050 | 3.603 | 31.343 | 0.726 | 0.057 | 3.857 |
| PSRT | IEEE TGRS 2023 | Supervised | 35.617 | 0.861 | 0.043 | 4.548 | 30.664 | 0.744 | 0.066 | 3.325 | 34.065 | **0.769** | 0.050 | 3.556 |
| SEST | IEEE JSTARS 2024 | Supervised | 37.375 | **0.879** | **0.041** | 1.465 | 33.115 | 0.805 | 0.050 | 1.416 | **34.962** | 0.765 | **0.049** | 1.364 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | **37.599** | 0.866 | 0.042 | **0.449** | **36.301** | **0.817** | **0.045** | **0.497** | 34.294 | 0.746 | 0.050 | **0.518** |
| **Groupwise Training** | | | | | | | | | | | | | | |
| **Train A → Test BC** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 35.583 | 0.843 | 0.052 | 1.366 | 34.180 | 0.805 | 0.052 | 1.020 | 32.990 | 0.733 | 0.061 | 2.145 |
| ESSAFORMER | ICCV 2023 | Supervised | 33.700 | 0.823 | 0.052 | 3.533 | 33.136 | 0.780 | 0.055 | 3.172 | 31.892 | 0.721 | 0.060 | 3.353 |
| PSRT | IEEE TGRS 2023 | Supervised | 36.398 | 0.854 | **0.046** | 3.356 | 35.483 | **0.806** | **0.050** | 5.123 | 33.158 | **0.747** | **0.056** | 3.321 |
| SEST | IEEE JSTARS 2024 | Supervised | 36.610 | 0.843 | 0.053 | 4.012 | 35.093 | 0.790 | 0.056 | 4.011 | 33.427 | 0.728 | 0.061 | 3.847 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | **37.305** | **0.861** | 0.047 | **0.513** | **35.674** | 0.799 | 0.054 | **0.668** | **34.275** | 0.739 | 0.060 | **1.414** |
| **Train B → Test CA** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 34.931 | 0.853 | 0.050 | 1.152 | 33.972 | 0.806 | 0.056 | 0.880 | 32.995 | **0.755** | 0.058 | 2.666 |
| ESSAFORMER | ICCV 2023 | Supervised | 33.711 | 0.837 | 0.049 | 2.691 | 32.676 | 0.770 | 0.056 | 1.733 | 31.906 | 0.735 | 0.056 | 1.747 |
| PSRT | IEEE TGRS 2023 | Supervised | 35.606 | 0.858 | 0.044 | 3.544 | 35.055 | **0.811** | **0.048** | 3.508 | 31.215 | 0.720 | 0.055 | 3.557 |
| SEST | IEEE JSTARS 2024 | Supervised | 37.087 | 0.854 | 0.048 | 1.395 | 35.554 | 0.800 | 0.051 | 1.336 | 33.475 | 0.747 | 0.057 | 1.331 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | **37.169** | **0.862** | **0.042** | **0.541** | **35.841** | 0.808 | 0.050 | **0.723** | **34.067** | 0.754 | **0.053** | **0.697** |
| **Train C → Test AB** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 34.522 | **0.842** | 0.045 | 1.552 | 32.164 | 0.772 | 0.052 | 2.142 | 33.098 | 0.746 | 0.056 | 2.381 |
| ESSAFORMER | ICCV 2023 | Supervised | 32.034 | 0.779 | 0.049 | 1.512 | 30.978 | 0.709 | 0.058 | 1.556 | 29.886 | 0.626 | 0.062 | 1.520 |
| PSRT | IEEE TGRS 2023 | Supervised | 21.736 | 0.466 | 0.131 | 3.657 | 21.356 | 0.522 | 0.156 | 3.389 | 23.578 | 0.570 | 0.126 | 3.415 |
| SEST | IEEE JSTARS 2024 | Supervised | 33.600 | 0.825 | 0.050 | 2.430 | 34.518 | 0.787 | 0.050 | 2.356 | 33.821 | 0.736 | 0.051 | 1.985 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | **38.267** | 0.824 | **0.037** | **0.479** | **36.536** | **0.807** | **0.041** | **0.687** | **34.938** | **0.748** | **0.048** | **0.618** |

*Time denotes average inference time per sample. Metric order within each upscaling factor: MPSNR, MSSIM, RMSE, and average time per sample. Best and second-best values are determined independently within each training setting across the five compared models.*

Table 15: Super-resolution performance across full-data training, low-resource fine-tuning, and group-wise training settings. Each entry reports SAM/SID/ERGAS/Cross-Correlation for the corresponding upsampling factor. Lower values are better for SAM, SID, and ERGAS, while higher values are better for Cross-Correlation. Best and second-best values within each setting and upsampling factor are highlighted.

| Method | Venue–Year | Type | x2 | | | | x4 | | | | x8 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ | SAM ↓ | SID ↓ | ERGAS ↓ | CC ↑ |
| **Full Data** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 4.528 | **0.011** | 62.329 | 0.703 | 5.238 | **0.012** | 108.420 | 0.725 | 5.607 | **0.012** | 61.141 | 0.633 |
| ESSAFORMER | ICCV 2023 | Supervised | 5.369 | 0.131 | 145.363 | 0.738 | 5.607 | 0.128 | 58.878 | 0.684 | 6.870 | 0.101 | 32.693 | 0.615 |
| PSRT | IEEE TGRS 2023 | Supervised | 4.616 | 0.151 | 100.378 | **0.786** | 5.292 | 0.165 | 95.319 | 0.682 | 5.648 | 0.158 | 23.387 | 0.596 |
| SEST | IEEE JSTARS 2024 | Supervised | **4.217** | 0.076 | **31.955** | 0.775 | **3.775** | 0.093 | **17.022** | 0.745 | 5.868 | 0.126 | 34.654 | 0.616 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | 5.063 | 0.039 | 94.757 | 0.765 | 5.168 | 0.032 | 24.687 | **0.757** | **5.379** | 0.031 | **11.031** | **0.662** |
| **Low-Resource Fine-Tuning** | | | | | | | | | | | | | | |
| **10% of Train** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 5.411 | **0.013** | 115.034 | 0.785 | 6.148 | **0.014** | 57.822 | 0.691 | 6.832 | **0.013** | 23.250 | 0.618 |
| ESSAFORMER | ICCV 2023 | Supervised | 8.330 | 0.157 | 333.925 | 0.586 | 9.640 | 0.168 | 177.720 | 0.516 | 10.799 | 0.210 | 92.438 | 0.414 |
| PSRT | IEEE TGRS 2023 | Supervised | 6.442 | 0.149 | 129.027 | 0.747 | 5.482 | 0.153 | 61.154 | 0.658 | 4.965 | 0.169 | 27.093 | 0.608 |
| SEST | IEEE JSTARS 2024 | Supervised | **3.290** | 0.183 | **68.672** | **0.836** | **3.565** | 0.075 | **35.012** | **0.779** | **4.761** | 0.164 | **20.532** | 0.663 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | 3.869 | 0.143 | 90.470 | 0.828 | 5.765 | 0.246 | 55.028 | 0.753 | 6.298 | 0.212 | 23.367 | **0.667** |
| **20% of Train** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 7.104 | **0.011** | 114.904 | 0.719 | 7.227 | **0.010** | 79.239 | 0.701 | 6.933 | **0.010** | 23.181 | 0.611 |
| ESSAFORMER | ICCV 2023 | Supervised | 7.646 | 0.138 | 195.007 | 0.630 | 9.854 | 0.186 | 107.914 | 0.536 | 10.705 | 0.195 | 63.391 | 0.439 |
| PSRT | IEEE TGRS 2023 | Supervised | 7.376 | 0.354 | 89.032 | 0.742 | 14.280 | 0.267 | 397.425 | 0.592 | 8.923 | 0.140 | 47.329 | 0.491 |
| SEST | IEEE JSTARS 2024 | Supervised | 5.650 | 0.190 | 56.580 | 0.801 | **5.730** | 0.128 | 31.655 | 0.756 | 5.207 | 0.125 | **13.065** | 0.673 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | **3.930** | 0.069 | **51.534** | **0.831** | 5.819 | 0.124 | **30.500** | **0.765** | **5.005** | 0.094 | 16.109 | **0.686** |
| **30% of Train** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 5.116 | **0.009** | 113.403 | **0.834** | 5.864 | **0.010** | 41.540 | 0.708 | 6.864 | **0.010** | 25.210 | 0.682 |
| ESSAFORMER | ICCV 2023 | Supervised | 6.244 | 0.108 | 120.294 | 0.729 | 8.241 | 0.148 | 87.631 | 0.618 | 7.169 | 0.107 | 35.868 | 0.567 |
| PSRT | IEEE TGRS 2023 | Supervised | 4.923 | 0.086 | 80.766 | 0.774 | 8.206 | 0.306 | 69.712 | 0.610 | 5.514 | 0.092 | 20.010 | 0.637 |
| SEST | IEEE JSTARS 2024 | Supervised | **4.883** | 0.072 | **55.273** | 0.820 | 7.437 | 0.094 | 47.351 | 0.717 | 6.152 | 0.127 | **9.433** | 0.675 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | 5.562 | 0.121 | 94.571 | 0.825 | **5.408** | 0.086 | **31.670** | **0.784** | **5.225** | 0.103 | 13.313 | **0.689** |
| **Groupwise Training** | | | | | | | | | | | | | | |
| **Train A → Test BC** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 5.000 | **0.014** | 84.602 | 0.779 | 5.331 | **0.011** | 53.821 | 0.700 | 5.825 | **0.015** | 20.494 | 0.662 |
| ESSAFORMER | ICCV 2023 | Supervised | 6.675 | 0.154 | 153.121 | 0.687 | 7.535 | 0.146 | 105.277 | 0.664 | 7.895 | 0.141 | 45.219 | 0.587 |
| PSRT | IEEE TGRS 2023 | Supervised | 5.217 | 0.204 | 66.629 | **0.809** | **4.900** | 0.170 | 26.004 | 0.753 | 6.222 | 0.356 | 20.968 | 0.636 |
| SEST | IEEE JSTARS 2024 | Supervised | **4.381** | 0.093 | **44.535** | 0.792 | 6.477 | 0.534 | **18.911** | 0.749 | 7.377 | 0.244 | **10.910** | 0.650 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | 5.665 | 0.102 | 82.255 | 0.797 | 5.362 | 0.093 | 34.116 | **0.756** | **5.601** | 0.100 | 17.896 | **0.675** |
| **Train B → Test CA** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 5.762 | **0.011** | 95.864 | 0.740 | 6.445 | **0.012** | 55.243 | 0.695 | 7.598 | **0.012** | 29.821 | 0.614 |
| ESSAFORMER | ICCV 2023 | Supervised | 6.857 | 0.124 | 154.810 | 0.663 | 8.341 | 0.180 | 74.837 | 0.608 | 7.778 | 0.160 | 39.377 | 0.518 |
| PSRT | IEEE TGRS 2023 | Supervised | 5.679 | 0.104 | 93.271 | **0.779** | **4.394** | 0.076 | 36.908 | **0.748** | 7.979 | 0.174 | 64.680 | 0.533 |
| SEST | IEEE JSTARS 2024 | Supervised | **4.969** | 0.083 | **51.655** | 0.758 | 5.524 | 0.110 | **19.111** | 0.706 | 7.663 | 0.124 | **16.808** | 0.624 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | 6.690 | 0.143 | 195.801 | 0.741 | 5.724 | 0.110 | 36.686 | 0.723 | **6.467** | 0.126 | 25.222 | **0.630** |
| **Train C → Test AB** | | | | | | | | | | | | | | |
| VOLFORMER | CVPR 2025 | Supervised | 8.080 | **0.010** | 72.653 | 0.694 | 9.591 | **0.011** | 64.668 | 0.634 | **6.530** | **0.012** | 24.136 | **0.638** |
| ESSAFORMER | ICCV 2023 | Supervised | 10.469 | 0.300 | 201.748 | 0.589 | 13.145 | 0.916 | 85.103 | 0.533 | 12.166 | 0.404 | 54.365 | 0.468 |
| PSRT | IEEE TGRS 2023 | Supervised | 17.632 | 0.708 | 816.676 | 0.494 | 18.255 | 0.754 | 612.001 | 0.464 | 15.728 | 0.613 | 158.904 | 0.338 |
| SEST | IEEE JSTARS 2024 | Supervised | **6.916** | 0.172 | 59.034 | 0.696 | **7.039** | 0.194 | **22.859** | 0.640 | 7.349 | 0.173 | 12.687 | 0.597 |
| FUSFORMER | IEEE GRSL 2022 | Supervised | 8.394 | 2.360 | **28.123** | **0.744** | 8.312 | 0.239 | 26.010 | **0.710** | 7.419 | 0.172 | **9.068** | 0.616 |

*Best and second-best values are marked by **bold** and underlined, respectively. ↓ indicates lower is better and ↑ indicates higher is better. All values are reproduced directly from the supplied CSV.*

Table 16: Inpainting performance across eight corruption cases. The results are divided into four metric groups for compact presentation: MPSNR/MSSIM, RMSE/Time, SAM/SID, and ERGAS/Cross-Correlation. Higher values are better for MPSNR, MSSIM, and Cross-Correlation, while lower values are better for RMSE, Time, SAM, SID, and ERGAS. Best and second-best values within each case are highlighted.

| Method | Venue–Year | Type | Stripes | | Dead lines | | Blocks | | Random Missing | | Mixed | | Missing50 | | Missing70 | | Missing90 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **MPSNR / MSSIM** | | | | | | | | | | | | | | | | | | |
| | | | MPSNR↑ | MSSIM↑ | MPSNR↑ | MSSIM↑ | MPSNR↑ | MSSIM↑ | MPSNR↑ | MSSIM↑ | MPSNR↑ | MSSIM↑ | MPSNR↑ | MSSIM↑ | MPSNR↑ | MSSIM↑ | MPSNR↑ | MSSIM↑ |
| HIRDIFF | CVPR 2024 | Diff | 43.008 | 0.885 | 41.258 | 0.862 | 44.282 | 0.932 | 41.581 | 0.920 | 40.980 | 0.904 | 33.406 | 0.626 | 32.016 | 0.550 | 30.671 | 0.417 |
| DIP2D | ICCV 2019 | Unsupervised | 49.136 | 0.980 | 46.563 | 0.970 | 47.783 | 0.976 | 42.781 | 0.893 | 39.812 | 0.867 | 32.629 | 0.515 | 31.433 | 0.456 | 30.479 | 0.378 |
| DIP3D | ICCV 2019 | Unsupervised | 18.093 | 0.216 | 31.982 | 0.719 | 12.788 | 0.099 | 12.149 | 0.100 | 27.679 | 0.619 | 44.646 | 0.960 | 45.000 | 0.952 | 43.371 | **0.944** |
| HLRTF(2D mask) | IEEE TGRS 2023 | Tensor | 42.358 | 0.989 | **inf** | **1.000** | 40.249 | 0.985 | 32.822 | 0.735 | 29.506 | 0.719 | 43.210 | 0.960 | 40.356 | 0.922 | 37.496 | 0.876 |
| HLRTF(3D mask) | IEEE TGRS 2023 | Tensor | 75.923 | **1.000** | 69.000 | **1.000** | **67.685** | **1.000** | **60.879** | **0.998** | **59.268** | **0.998** | 48.934 | 0.964 | 45.958 | 0.922 | 43.101 | 0.875 |
| R-DLRHyIn(2D MASK) | IEEE GRSL 2023 | Tensor | 8.541 | 0.248 | 25.729 | 0.267 | 8.569 | 0.248 | 8.611 | 0.249 | 26.374 | 0.283 | 26.258 | 0.257 | 26.230 | 0.254 | 26.193 | 0.251 |
| R-DLRHyIn(3D MASK) | IEEE GRSL 2023 | Tensor | **inf** | **1.000** | 71.684 | **1.000** | 58.573 | 0.998 | 59.621 | **0.998** | 50.286 | 0.984 | **inf** | **1.000** | **47.324** | **0.967** | **43.601** | 0.927 |
| **RMSE / Time** | | | | | | | | | | | | | | | | | | |
| | | | RMSE↓ | Time↓ | RMSE↓ | Time↓ | RMSE↓ | Time↓ | RMSE↓ | Time↓ | RMSE↓ | Time↓ | RMSE↓ | Time↓ | RMSE↓ | Time↓ | RMSE↓ | Time↓ |
| HIRDIFF | CVPR 2024 | Diff | 0.020 | **2.251** | 0.023 | **1.741** | 0.015 | **1.744** | 0.022 | **1.728** | 0.025 | **1.720** | 0.073 | **1.743** | 0.092 | **1.743** | 0.110 | **1.761** |
| DIP2D | ICCV 2019 | Unsupervised | 0.008 | 207.795 | 0.012 | 207.292 | 0.016 | 206.869 | 0.025 | 207.500 | 0.030 | 206.601 | 0.089 | 206.416 | 0.106 | 206.031 | 0.120 | 207.526 |
| DIP3D | ICCV 2019 | Unsupervised | 0.178 | 5645.355 | 0.056 | 5665.280 | 0.270 | 4921.576 | 0.294 | 5660.604 | 0.078 | 5648.536 | 0.009 | 5635.960 | **0.009** | 2715.918 | **0.011** | 2716.695 |
| HLRTF(2D mask) | IEEE TGRS 2023 | Tensor | 0.008 | 7.950 | **0.001** | 7.951 | 0.015 | 7.921 | 0.036 | 7.934 | 0.040 | 7.901 | 0.009 | 7.887 | 0.013 | 7.903 | 0.019 | 7.915 |
| HLRTF(3D mask) | IEEE TGRS 2023 | Tensor | **0.000** | 24.066 | **0.001** | 23.579 | **0.001** | 23.512 | **0.002** | 23.805 | **0.002** | 23.381 | 0.009 | 22.323 | 0.013 | 24.037 | 0.019 | 22.313 |
| R-DLRHyIn(2D MASK) | IEEE GRSL 2023 | Tensor | 0.411 | 1295.766 | 0.142 | 532.181 | 0.410 | 569.013 | 0.409 | 647.899 | 0.142 | 474.304 | 0.143 | 922.269 | 0.143 | 945.913 | 0.143 | 929.671 |
| R-DLRHyIn(3D MASK) | IEEE GRSL 2023 | Tensor | 0.001 | 594.696 | **0.001** | 386.375 | 0.002 | 603.539 | **0.002** | 382.369 | 0.006 | 316.225 | **0.000** | 340.508 | **0.009** | 569.044 | 0.014 | 826.850 |
| **SAM / SID** | | | | | | | | | | | | | | | | | | |
| | | | SAM↓ | SID↓ | SAM↓ | SID↓ | SAM↓ | SID↓ | SAM↓ | SID↓ | SAM↓ | SID↓ | SAM↓ | SID↓ | SAM↓ | SID↓ | SAM↓ | SID↓ |
| HIRDIFF | CVPR 2024 | Diff | 0.267 | 0.600 | 0.240 | 0.682 | **0.202** | 0.485 | **0.217** | 0.478 | **0.246** | 0.482 | 0.371 | 1.018 | 0.451 | 1.561 | 0.847 | 7.475 |
| DIP2D | ICCV 2019 | Unsupervised | 0.184 | 0.639 | 0.141 | 0.439 | 0.204 | 0.522 | 0.243 | 0.352 | 0.303 | 0.891 | 0.612 | 6.881 | 0.814 | 9.404 | 1.089 | 14.586 |
| DIP3D | ICCV 2019 | Unsupervised | 1.203 | 16.203 | 0.365 | 1.888 | 1.178 | 13.596 | 1.093 | 10.878 | 0.498 | 2.573 | 0.059 | 0.023 | **0.059** | 0.019 | **0.065** | **0.022** |
| HLRTF(2D mask) | IEEE TGRS 2023 | Tensor | 5.300 | 0.168 | 0.032 | 0.016 | 5.444 | 0.209 | 15.699 | 0.133 | 19.939 | 0.324 | 1.571 | 1.104 | 2.536 | 1.571 | 3.809 | 2.081 |
| HLRTF(3D mask) | IEEE TGRS 2023 | Tensor | **0.131** | 0.007 | **0.022** | **0.000** | 0.772 | 0.063 | 3.390 | 0.528 | 3.633 | 0.585 | 1.513 | 0.008 | 2.514 | 0.016 | 3.812 | 0.024 |
| R-DLRHyIn(2D MASK) | IEEE GRSL 2023 | Tensor | 44.488 | 1.435 | 44.633 | 1.441 | 44.407 | 1.429 | 44.596 | 1.439 | 44.874 | 1.446 | 45.612 | 1.477 | 45.598 | 1.473 | 45.402 | 1.456 |
| R-DLRHyIn(3D MASK) | IEEE GRSL 2023 | Tensor | 0.234 | **0.000** | 0.024 | **0.000** | 0.958 | **0.003** | 0.901 | **0.002** | 1.513 | **0.009** | **0.038** | **0.000** | 2.423 | **0.015** | 3.701 | 0.028 |
| **ERGAS / CC** | | | | | | | | | | | | | | | | | | |
| | | | ERGAS↓ | CC↑ | ERGAS↓ | CC↑ | ERGAS↓ | CC↑ | ERGAS↓ | CC↑ | ERGAS↓ | CC↑ | ERGAS↓ | CC↑ | ERGAS↓ | CC↑ | ERGAS↓ | CC↑ |
| HIRDIFF | CVPR 2024 | Diff | 449.461 | 0.860 | 82.104 | 0.882 | 65.555 | 0.917 | 69.586 | 0.902 | 265.365 | 0.874 | 104.445 | 0.750 | 114.390 | 0.669 | 130.339 | 0.364 |
| DIP2D | ICCV 2019 | Unsupervised | 470.951 | 0.890 | 49.463 | 0.931 | 51.695 | 0.920 | 53.677 | 0.910 | 302.689 | 0.852 | 107.128 | 0.538 | 108.727 | 0.393 | 120.036 | 0.211 |
| DIP3D | ICCV 2019 | Unsupervised | 3096.066 | 0.005 | 900.086 | 0.678 | 9114.503 | -0.057 | 6982.240 | -0.028 | 2854.857 | 0.489 | 51.020 | 0.994 | 39.834 | **0.994** | 54.235 | **0.992** |
| HLRTF(2D mask) | IEEE TGRS 2023 | Tensor | 518.308 | 0.907 | 4.247 | 0.999 | 14.659 | 0.984 | 34.299 | 0.922 | 483.009 | 0.823 | 35.017 | 0.959 | 43.743 | 0.933 | 54.938 | 0.893 |
| HLRTF(3D mask) | IEEE TGRS 2023 | Tensor | **3.332** | **1.000** | 4.498 | 0.999 | **6.940** | 0.998 | 19.751 | 0.987 | 21.527 | 0.984 | 35.341 | 0.959 | 43.827 | 0.933 | 54.784 | 0.893 |
| R-DLRHyIn(2D MASK) | IEEE GRSL 2023 | Tensor | 14134.249 | 0.056 | 201.158 | 0.015 | 13705.052 | 0.104 | 13958.689 | 0.029 | 128.322 | 0.041 | 125.077 | -0.002 | 125.577 | -0.003 | 126.238 | 0.004 |
| R-DLRHyIn(3D MASK) | IEEE GRSL 2023 | Tensor | 3.734 | **1.000** | **1.534** | **1.000** | 11.682 | **1.000** | **10.883** | **1.000** | **18.667** | **0.996** | **1.269** | **1.000** | **23.506** | 0.992 | **30.714** | 0.982 |

*Best and second-best values are marked by **bold** and underlined, respectively. ↑ indicates higher is better and ↓ indicates lower is better. All values are reproduced directly from the supplied CSV files.*