



DeepSpace: Super Resolution Powered Efficient and Reliable Satellite Image Data Acquisition

Chuanhao Sun[†], Yu Zhang[†], Bill Tao[‡], Deepak Vasisht[‡], and Mahesh K. Marina[†]

The University of Edinburgh[†] University of Illinois Urbana-Champaign[‡]

Abstract

Large constellations of low-earth orbit satellites enable frequent high-resolution earth imaging for numerous geospatial applications. They generate large volumes of data in space, hundreds of Terabytes per day, which must be transported to Earth through constrained intermittent connections to ground stations. The large volumes lead to large day-level delay in data download and exorbitant cloud storage costs. We propose DEEPSpace, a new deep learning-based super-resolution approach that compresses satellite imagery by over two orders of magnitude, while preserving image quality using a tailored mixture of experts (MoE) super-resolution framework. DEEPSpace reduces the network bandwidth requirements for space-Earth transfer, and can compress images for cloud storage. DEEPSpace achieves such gains with the limited computational power available on small LEO satellites. We extensively evaluate DEEPSpace against a wide range of state-of-the-art baselines considering multiple satellite image datasets and demonstrate the above mentioned benefits. We further demonstrate the effectiveness of DEEPSpace through several distinct downstream applications (wildfire detection, land use and cropland classification, and fine-grained plastic detection in oceans).

CCS Concepts

- Networks; • Computing methodologies → Machine learning;

Keywords

LEO Satellites, Earth Observation, Generative AI

ACM Reference Format:

Chuanhao Sun[†], Yu Zhang[†], Bill Tao[‡], Deepak Vasisht[‡], and Mahesh K. Marina[†]. 2025. DeepSpace: Super Resolution Powered Efficient and Reliable Satellite Image Data Acquisition. In *ACM SIGCOMM 2025 Conference (SIGCOMM '25), September 8–11, 2025, Coimbra, Portugal*. ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/3718958.3750523>

1 Introduction

In recent years, hundreds of satellites were deployed in Low Earth Orbits (LEO) to capture frequent high-resolution Earth imagery [6, 20]. These satellites generate unprecedented density in spatio-temporal coverage of Earth – capturing multiple images of each location on Earth every day at high resolutions (e.g., 3m per pixel). Such imagery enables wide range of valuable geospatial sensing and computing applications, including disaster detection and relief [15,



This work is licensed under a Creative Commons Attribution 4.0 International License.

SIGCOMM '25, Coimbra, Portugal

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1524-2/25/09

<https://doi.org/10.1145/3718958.3750523>

27, 29, 52], climate change and environment monitoring [11, 47, 69], agriculture [31, 48, 50] and map generation [23, 37, 63].

Images captured by these satellites need to be transferred to Earth and stored in the cloud before a downstream application processes them, as Figure 1 illustrates. In practice, earth imagery transfer and storage faces **three key challenges**:

(1) Networking bottlenecks and delays: First, each satellite generates nearly a Terabyte (TB) of data per day, leading to hundreds of TBs of data per constellation [62]. To make things worse, the low orbits of these satellites lead to intermittent connectivity with ground station receivers on Earth. Each satellite-ground station connection lasts for less than ten minutes and has limited bandwidth (around 200Mbps), and such connections happen few times per day [24]. Due to such intermittent constrained connectivity and large data volumes, satellite imagery can suffer from transfer delays of several hours to a few days [61]. Such delays are impractical for latency-sensitive applications like monitoring natural disasters. Recent work has also shown that the growth in satellite deployments continues to outpace space-earth data capacity, implying that only a small fraction of the data can be downloaded in the future [18, 24].

(2) High cloud storage costs: Cloud storage of earth observation data is an increasingly expensive operation. Using public datasets [21, 37, 63], we estimate that more than 2 Petabytes (PBs) of new observation data to cover the earth gets generated each month. To store this much amount of data for the past few (say, 3) years and support immediate access, it costs millions of dollars each month, based on the storage pricing information of prominent cloud service providers (e.g., AWS [7], Azure [8]). Even cold storage is an expensive proposition. These costs keep adding up as a constellation operates for years, even more so with the increase in constellation and image sizes over time.

(3) Limited onboard compute resources on satellites: Satellites have limited computational resources – the state-of-the-art earth imaging satellites have small GPUs such as NVIDIA Jetson [4]. These resources are further limited by power availability which must prioritize mission-critical satellite operations. So, it is non-trivial to accommodate compute-intensive compression algorithms on the satellite.

The above challenges necessitate the design of end-to-end compression systems that meet the following goals: (a) *high compression ratios (CRs)* to reduce the volume of data that needs to be transferred and stored; (b) *computational efficiency* of onboard operations to run within the compute and power constraints on satellites; (c) *reliable decompression* to meet the image quality requirements of a wide array of downstream applications.

None of the existing approaches for satellite image data acquisition [9, 17, 18, 24, 28, 61] satisfy these goals, as summarized in Table 1 and discussed in §2.2. While lossless methods like “7z” [57]

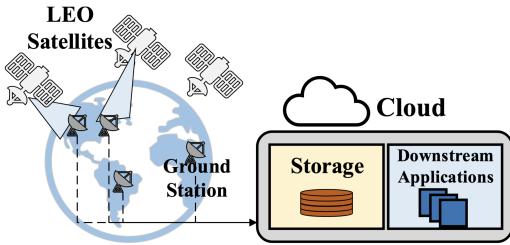


Figure 1: Earth observation with satellite imagery.

are unviable due to their negligible CR (1.03), all other lossy compression methods only offer a CR typically around 10 and up to 32 – much smaller than the 50 or higher desired CR [18]. As such, none of the existing methods meet goal (a). Some of the auto-encoder (AE) based methods (e.g., VQ-VAE 2 [53]) are too computationally heavy to run onboard satellites, and so do not meet goal (b). Most existing methods do not meet goal (c) in that they do not provide any guarantees on decoded/reconstructed image quality.

We propose DEEPSpace, a novel satellite image data acquisition system that meets all three goals mentioned above. DEEPSpace adopts the deep learning-based image super-resolution (SR) approach [76] for the first time in the context of satellite imagery. In DEEPSpace, images captured by a satellite are compressed into low-resolution (LR) versions and sent over the downlink to the cloud, where LR versions are decompressed back to high-resolution (HR) versions that mimic the original images captured by the satellite. DEEPSpace employs image sampling (resizing) onboard satellites for compressing raw captured images to LR versions, as it is not only a lightweight operation that can be run well within onboard compute constraints but is also a flexible way to adapt and control the compression level.

DEEPSpace enables significantly higher CRs than existing satellite image data acquisition systems by leveraging two key insights: (i) images captured by a satellite can be compressed down to the smallest LR form (equivalently, increase the CR to the point) that ensures LR images are distinguishable from each other; (ii) compared to image reconstruction through SR for all regions, per-region reconstruction allows higher CRs without compromising image reconstruction quality. We realize the first insight through compression onboard satellites that selects CR at the image tile level based on its inherent redundancy as well as closeness to reference images, latter based on communication and space efficient hash codes sporadically fetched over the uplink from the cloud. The second insight is reflected in DEEPSpace through a tailored mixture of experts (MoE) [14] style SR framework that employs multiple “expert” SR models, each specializing in different regions and image characteristics. Furthermore, we provide post-hoc measures to estimate the error between recovered HR images and their original counterparts and to guarantee that error stays under a specified level. We implement the DEEPSpace system considering the typical compute hardware suitable for deployment on satellites [4, 18, 61].

We perform a comprehensive evaluation of DEEPSpace using satellite images from multiple different datasets [1, 20, 63] covering diverse regions across the world. We evaluate its performance relative to state-of-the-art systems in terms of compression ratios,

reliability and compute efficiency. We show that DEEPSpace can achieve two orders of magnitudes greater compression (equivalently, 100x less traffic on downlink) while ensuring high reliability. Moreover, the significant compression gain on downlink traffic is achieved with less overhead on the satellite: magnitudes (100s x) faster processing speed and 10x less onboard storage requirement. Finally, we showcase the effectiveness of DEEPSpace through diverse downstream applications: (1) wildfire detection; (2) plastic detection in oceans; and (3) land use and cropland classification. For all these applications, we show that the recovered images with DEEPSpace even after >100x compression has a similar effect to that using ground truth images.

In summary, we make the following key contributions:

- We introduce the approach of deep learning based image super-resolution (SR) combined with image sampling onboard satellites to the satellite based earth observation setting.
- We propose DEEPSpace system that realizes this approach, featuring a tailored MoE based SR framework design for on-demand image decompression in the cloud with control over recovered image quality as well as innovations to adaptively select CR at the image tile level onboard satellites in a lightweight manner.
- We make the DEEPSpace implementation publicly available¹ to benefit the research and innovation on earth observation with LEO satellites.
- Through extensive evaluations considering multiple different satellite image datasets and state-of-the-art baselines, we show that DEEPSpace achieves two orders of magnitude higher CRs without compromising reliability of HR image recovery and using minimal onboard compute resources.
- Through three diverse and representative downstream applications, we demonstrate that reconstructed images with DEEPSpace even after 100x+ compression yield similar application performance as with the ground truth raw uncompressed images.

2 Motivation

2.1 Context and Constraints

Our setting of earth observation using images from low earth orbit (LEO) satellite systems has certain characteristics and constraints, as outlined below.

Satellite orbits: LEO satellites for earth observation operate in orbits around 500-1000 Km from the surface of the Earth. These satellites have an orbital period of around 90 minutes and typically orbit around the Earth’s poles. As the Earth rotates underneath, each satellite scans different locations on Earth during each orbit. Modern megaconstellations for earth observation (e.g., Planet Dove [20]) have hundreds of these satellites deployed to capture multiple images of each location on Earth every day.

Communication constraints: Each satellite captures RGB or multi-spectral images at high resolutions of few square meters per pixel. The images captured from each satellite can amount to more than a Terabyte (TB) of data per day [2], and this in turn translates to hundreds of TBs of data per constellation [62]. These images must be transmitted to Earth for processing and analysis, typically

¹<https://github.com/netsys-edinburgh/DeepSpace/>

in the cloud. However, the data transfer happens through intermittent contacts with ground stations. Due to a satellite’s orbital motion, these contacts last less than ten minutes and are limited to 4-6 contacts per day. Due to such intermittent connectivity and high data volumes, earth imagery data experiences hours to days of delay [62], limiting its use for time-sensitive applications like disaster monitoring as well as necessitating large onboard storage.

Power and computational constraints: Modern constellations consist of relatively small cubesats with size, weight, and power constraints. They generate power through solar panels during periods when they face the sun. This energy is stored for use throughout the orbit and is prioritized for critical satellite operations such as communication, attitude determination and control. The remaining power can be used for compute that effectively limits the compute resource onboard. Past work [18, 61] has therefore explored the use of low-power GPUs such as NVIDIA Jetson AGX Orin in the 15W power draw mode for operation on cubesats. This is consistent with recently launched LEO satellites with edge-computing capabilities [4].

2.2 Limitations of Existing Approaches

By way of motivation, here we discuss the different approaches that have been applied to resolve the satellite image data acquisition problem covering the three goals stated at the outset: (i) high compression ratio (CR), (ii) computational efficiency, and (iii) reliable decompression. Table 1 provides the summary.

From a CR perspective, we first start with an indication of the desired levels of CR. Prior works [18, 24] estimate that only 2% of the images captured by satellites can be downloaded currently due to the satellite-to-ground downlink bottleneck. Overcoming this bottleneck translates to a CR of at least 50. With expected increases in constellation size, downlink will become more of a bottleneck as per the analysis by Denby et al. [18] and so an even higher CR would be needed in the future.

To meet the above CR target, we note that lossless compression methods [57] like “7z” are unviable for two reasons. First, the typical CR achieved with these methods is a mere 1.03 based on our experiments using datasets from Planet Inc’s Dove constellation [20] and DynamicEarthNet (DEN) [63]; even the best case CR with these methods is 1.09 – way less than desired. Second, the image capturing process on satellites introduces some inherent distortion, as elaborated in Appendix A.1. Consequently, all existing works for satellite image compression in the literature [10, 18, 24, 25, 53, 71] (represented in Table 1) employ lossy methods. Even these lossy methods can only support CRs up to 32, as highlighted in Table 1, falling short of the desired CR. Our proposed solution, DEEPSpace, overcomes this barrier, and other limitations of existing methods elaborated below, through a tailored system design taking a new “sampling + deep learning based image super-resolution (SR)” approach.

Starting with the classical methods, Lánczos interpolation [36] is a traditional method that can be applied to resize (downsample) images to reduce their size, then interpolate afterwards to upsample and recover them to their original size. This approach can be effectively error free with Nyquist sampling rate but limits the CR. Compressive sensing (CS) [10] is another classical method that can

Category	Methods	Typical CR ↑	Fast Proc. (✓)	Error Bound (✓)
Conventional Methods	Lanczos Interp. [36]	<5	✓	✓
	CS [10]	5-10	✓	✓
Auto Encoders	DCSN [28]	8 ~ 24	✓	✗
	VQ-VAE 2 [53]	16 ~ 32	✗	✗
Orbital Edge Computing (OEC)	Kodan [18]	<10	✓	✗
	Earth+ [24]	<10	✓	✗
Sampling+Deep SR	RCAN [75]	10 ~ 32	✓	✗
	SR3 [56]	16 ~ 32	✓	✗
	WaveDiff [49]	16 ~ 32	✓	✗
	DEEPSpace (Ours)	> 256	✓	✓

Table 1: Qualitative comparison of different approaches for satellite image data acquisition. CS: Compressive Sensing. SR: Super Resolution.

provide slightly better CR by exploiting the underlying sparsity in the image signal.

In a different category of works, auto-encoders (AEs) are employed to enable better compression ratios (CRs). In fact, there exist custom AE designs for satellite image compression [9, 17, 28] that perform lightweight encoding to obtain compact representation of images captured by satellites in latent space which are then transmitted over the downlink. The reverse decoding process is done on the other side to reconstruct the original images. Using the state-of-the-art AEs for image compression from the computer vision domain [25, 53, 71] can yield higher CRs but these encoders are computationally heavy to run onboard satellites.

Orbital Edge Computing (OEC) [19] has emerged as a new paradigm in recent years that leverages compute onboard satellites to mitigate the satellite to ground downlink communication bottleneck. Kodan [18] is a representative work following the OEC paradigm. For a given geospatial analysis application, Kodan stays within the limits of onboard compute capability and maximizes the value of data communicated through the constrained bottleneck by filtering out low value data and adapting data precision. Serval [61] is broadly similar to Kodan in that it prioritizes transmission of image data for certain latency-sensitive applications (e.g., forest fire detection) through static/dynamic filters, whose computation is distributed between the ground stations and satellites. Umbra [62] is a previous work by the authors of Serval that takes a complementary networking perspective to schedule downlink transmissions across ground stations, considering their differences in ground-cloud segment network characteristics and load.

Earth+ [24] is the latest work under the OEC category in which each satellite employs an image compression scheme that transmits only “changes” with respect to reference images for different locations, as opposed to transmitting raw images themselves. It leverages the uplink bandwidth to equip the satellites with the latest reference images from across the constellation to enable change based compression and transmission.

Note that both Kodan and Earth+ do not provide any error bounds as they are best effort in terms of quality/reliability of images recovered on the cloud side. Like with the above OEC works, AE based methods also do not provide any error bounds on the image reconstruction quality. Also note that, with the exception of Kodan and Earth+, the aforementioned methods require setting a fixed CR across all transferred images and so CR setting needs

to be judiciously chosen to yield acceptable reconstructed image quality.

3 DEEPSpace System

Here we present our solution, DEEPSpace (schematic in Figure 2), for satellite image data acquisition that meets the three goals as laid out at the outset. *Through DEEPSpace, we introduce the deep learning based super-resolution (SR) approach to the satellite image data collection setting for the first time.* This approach inherently involves compression of captured high-resolution (HR) images on satellites to low-resolution (LR) versions onboard through image sampling (resizing), a computationally lightweight operation well suited with limited compute and fast processing onboard. On the cloud side, HR versions reflecting original images are recovered as needed from received LR versions through a SR framework.

3.1 Key Insights

While there exist methods for deep learning based image super-resolution from the computer vision (CV) domain [49, 56, 75, 76], as shown in Table 1 and through our evaluations (§6), their straightforward application to our satellite image data acquisition setting yields CRs that are only marginally better than existing methods designed for this setting. Furthermore, existing SR methods do not provide any guarantees or bounds on the reconstructed image quality. In contrast, DEEPSpace with its tailored design not only enables an order of magnitude higher CR but also provides error bounds on reconstructed image quality, comes with fast processing for onboard compression.

To achieve higher levels of CR than what is possible with existing solutions, our design leverages two key insights:

1) Can use maximal CR that ensures distinguishability among sampled (LR) images. In DEEPSpace, compression is realized via sampling, *i.e.*, by reducing the image resolution to a *recoverable* level. From Sparse Auto Encoder (SAE) theory [16, 45, 51], we know that a faithful reconstruction of encoded data requires the encoding of distinct inputs be distinct, and obviously smaller in size than inputs themselves. In our context, sampling is used to perform encoding. For reliable LR-HR image reconstruction, we need sampled versions of images to be still distinguishable from each other. In other words, with the sampling process $S(\cdot)$, we want to ensure that:

$$\begin{aligned} \text{If : } & \text{Dis}(A, B) > \text{Dis}(A, C), \\ & \rightarrow \text{Dis}_L(S(A), S(B)) > \text{Dis}_L(S(A), S(C)) \end{aligned} \quad (1)$$

where A, B, C are image instances, $\text{Dis}(\cdot)$ is a distance measure quantifying their similarity, and $\text{Dis}_L(\cdot)$ is the same distance measure applied in latent space, *i.e.*, the space of sampled (LR) images. *Limiting $S(\cdot)$ to satisfy condition (1) guarantees that in theory it is possible to distinguish different inputs from their latent representations.* As we use sampling as encoder in DEEPSpace, note that the compressed input still keep the same representation as the corresponding original image, and so $\text{Dis}(\cdot) = \text{Dis}_L(\cdot)$. We use the well-known SSIM [67] for $\text{Dis}(\cdot)$ and $\text{Dis}_L(\cdot)$.

The implication from the above is that we can push the CR to the point that ensures distinguishability between LR images. As elaborated in §3.2, we obtain cues on the maximal CR to use for an image at the satellite end via (a) interpolation based recovery

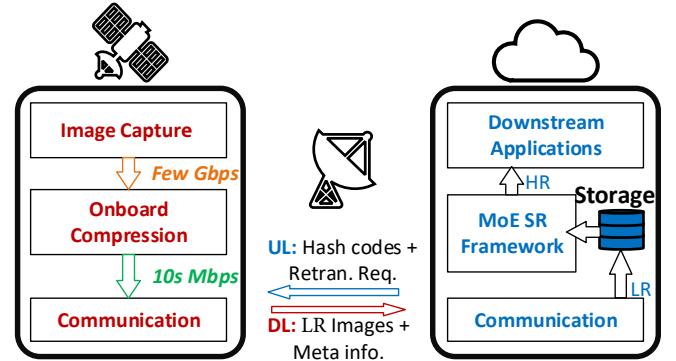


Figure 2: Schematic overview of DEEPSpace. DL: Downlink, UL: Uplink, LR: Low Resolution, HR: High Resolution, MoE: Mixture of Experts, SR: Super Resolution.

of sampled images and (b) similarity with most recent reference images for same region.

2) Using a multitude of image SR models specializing for different contexts allows faithful recovery even at high CRs. The extent of sampling (level of CR) is limited to ensure faithful recovery of images from their sampled versions, when they come from a large input set spanning different regions. As an example, in Figure 3 we compare the similarity between an image of one region A with another image of the same region at a different time (in blue) as well as with images of different regions B and C , with increasing CRs. We use SSIM as the similarity measure for this argument. At low CR, it is very easy to distinguish different regions, as images of the same region A are similar having SSIM around 0.9, while the SSIM to different regions B and C is lower around 0.7. However, as we increase CR, we find the mutual similarity of images from same region A at different time can become lower than the similarity to different regions, suggesting the difficulty of recovering HR images across different regions. If we consider region A , B and C at the same

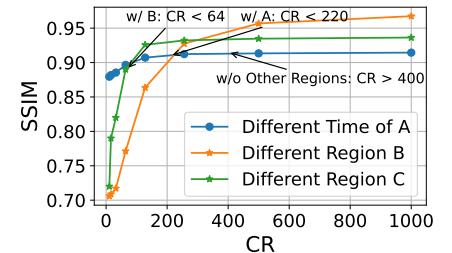


Figure 3: Image recoverability for different regions at different CRs.

time, the maximum CR to identify all regions (64) is lower than the case when only region A and B are considered (220) in Figure 3, and that is in turn lower than the case when only region A at different times is considered (> 400). In other words, generalizing to more regions makes it harder to identify different regions at high CR, and therefore limits the potential CR for achieving faithful reconstruction. The above discussion suggests that we should leverage a pool of “expert” SR models, each specializing in the reliable image recovery (SR) of different regions and image characteristics (e.g., cloud level). This allows us to raise the CR to a higher level than

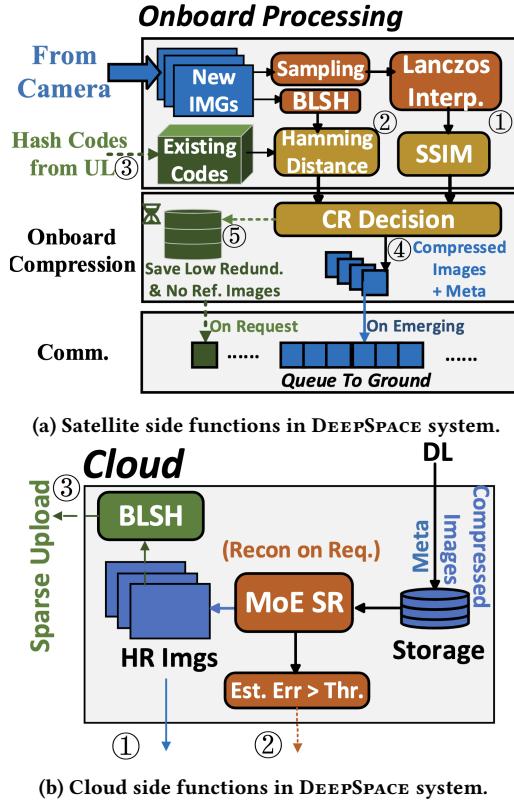


Figure 4: Onboard and cloud based processing.

would otherwise be possible with a single recovery model across regions.

3.2 Design Overview

Figure 2 depicts a high level overview of the DEEPSpace system, broadly composed of two components: (1) image compression onboard a satellite; and (2) on-demand reconstruction of compressed image on the cloud end through SR. Our design builds on the two insights from the previous subsection. Also note that in DEEPSpace, we collect and store *all* images captured by each earth observation satellite in the cloud in a compressed form. Then when a downstream application needs to access all/subset of these images, we decompress them on-demand to provide high-fidelity versions mimicking their original image counterparts.

Image compression onboard satellite. For this, we exploit two complementary opportunities in DEEPSpace, both via sampling (as illustrated in Figure 5): (i) inherent redundancy (ϕ) within an image; (ii) similarity (η) to known (previously seen) images, i.e., if a newly captured image of a region is a close replica of known reference image for the same region. We use (i) to guide the base level of compression. As the next level, we rely on (ii) to further compress the result obtained from applying (i). Across both these levels/dimensions, we pursue a fine-grained compression of images by viewing each of them as a composition of smaller non-overlapping “*tiles*” (with a default size of 256 pixels \times 256 pixels).

More concretely, during the onboard image compression process (depicted in Figure 4a and details in §4.1), DEEPSpace first estimates the inherent redundancy ϕ (① in Figure 4a) and similarity to existing references η (② in Figure 4a). A high value of ϕ indicates high inherent redundancy in the image (e.g., clouds or ice), which symbolizes potential for significant compression. A high value η indicates high similarity with reference images (and hence, higher potential CR). To compute η , DEEPSpace obtains lightweight representation of reference images via binary locality sensitive hashing (BLSH) [38, 66] from the ground over the uplink (③ in Figure 4a). An adaptive compression ratio is applied to each image based on ϕ and η , and then the compressed version is sent out to the ground (④ in Figure 4a). Images with low ϕ and η are temporarily stored onboard for potential retransmission later with minimal or no compression (⑤ in Figure 4a), as they reflect a potential outlier in that they are both complex and not similar to known images (quadrant ③ in Figure 5). For other images, DEEPSpace sends a compressed version of the onboard image to the ground.

On-demand image decompression at cloud For this, DEEPSpace features a tailored deep SR framework (illustrated in Figure 4b and details in §4.2) with a pool of expert SR models à la mixture of experts (MoE) models [14], guided by insight 2) above. Each expert is realized with a separately trained wavelet diffusion model [49], following the reasoning in §4.2. Overall, there are three major functions in the decompression process, namely high resolution data reconstruction (① in Figure 4b), triggering retransmission when the estimated fidelity is lower than expected (② in Figure 4b, details in §4.3), and updation of reference hash codes (③ in Figure 4b). Furthermore, we provide a way to guarantee a specified reconstruction quality (error bound) as well as *post-hoc* measures for estimating quality of recovered images on the cloud side with respect to their original counterparts on the satellite end (details in §4.2 and §4.3).

Example: Figure 6 shows an example to illustrate the end-to-end operation of DEEPSpace, including the combined effect of applying the two levels of onboard compression at the tile granularity.

DEEPSpace first compresses (samples) each of the image tiles separately by a factor of K based on their inherent redundancy and then further increases the CR by a factor α (still at the tile level) if the tile is a close replica of a recent known image tile for the same location. Resulting compressed images are transmitted over the downlink and are recovered with high fidelity through a MoE based SR framework and known information at the cloud.

For comparison, we show Kodan [18], a current state of the art approach. In OEC based methods like Kodan [18, 61], there is onboard processing to identify if an image tile is valuable by ML-based classification, *i.e.*, non-cloudy and required by a certain downstream application. For fair comparison to support all applications, we let Kodan discard only cloudy tiles onboard and replace them with the nearest non-cloudy observation on the ground in Figure 6. This

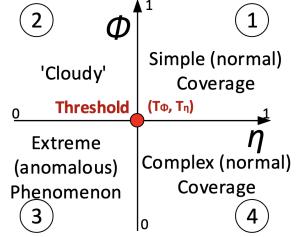


Figure 5: Type of image based on ϕ and η value.

example scenario highlights the orders of magnitude higher CR enabled by DEEPSpace.

4 Detailed Design

Here we elaborate on the design details of key components underlying DEEPSpace. Implementation aspects of DEEPSpace are provided in Appendix A.2.

4.1 Image Compression Onboard Satellite

Inherent redundancy quantification and CR selection: For redundancy quantification, we first sample each tile of the original (captured) image evenly by K times, and then use Lanczos interpolation [36] to reconstruct the tile, see further discussion and visualization in §A.3. The redundancy is then defined as

$$\phi(I, K) = \text{SSIM}(I, \text{LI}(\mathbf{S}(I, K))) \quad (2)$$

where I is the input image tile, $\text{LI}(\cdot)$ is the Lanczos Interpolation, and $\mathbf{S}(\cdot, K)$ represents evenly sampling by K times. We use SSIM [67] as the similarity measure because it is a commonly used image quality metric and also normalized (between 0 to 1), making it easy to align among different scenarios. Furthermore, it is computationally lightweight and also known to be a robust metric than PSNR for most computer vision applications [46, 67]. Although learned metrics like LPIPS [74] are more accurate for DL-based downstream tasks, they involve model inference and so have high computational overhead for use onboard satellites.

To attain maximal compression while also allowing reliable recovery, we choose K as: $\arg \max_K \phi(I, K) > \tau_\phi$ (by default set to 0.85). Note that K (the CR value obtained in the above manner) could be different for different tiles of an image (see Figure 6, for example).

Near replicate detection: We use binary locality sensitive hashing (BLSH) [38, 66] to help detect if a newly captured image tile of a region nearly replicates the tile of a recent image of the same region, where the recent reference is updated weekly by default. We do this as follows (at grey scale):

$$\eta(I) = \arg \max_{X \in \mathbf{X}_{\text{ref}}} 1 - \frac{\text{HD}(\text{BLSH}(I), X)}{w \times l} \quad (3)$$

where I is the input image tile, X is an instance from reference BLSH code set \mathbf{X}_{ref} . $\text{HD}(\cdot)$ is the Hamming Distance of hash codes. $\text{HD}(\cdot)$ is then normalized by the image size $w \times l$. The satellite receives BLSH of the reference image on ground at the tile level for each region in its orbit from the cloud via the uplink. These reference BLSH codes on the satellite are updated infrequently, once a week in our setup. We choose BLSH for its fast computation and *ultra-small* hash code size, enabling real-time onboard processing and efficient uploading of reference codes from the cloud – around 4 orders of magnitude smaller in size compared to sending even a LR version of reference image over uplink. This approach significantly reduces storage requirements to a negligible size, even when saving codes to cover the entire Earth.

The BLSH based replicate detection serves two purposes in DEEPSpace: (1) Guide compression process by giving tiles with near replicate references higher CR than the CR defined by Equation (2), i.e., increase CR for a tile from K to αK if $\eta(I) > \tau_\eta$ (by

default set to 0.9); (2) Simple yet effective cloud detection mechanism as part of MoE based SR framework used for decompression (see Figure 7 and §4.2). The rationale for the latter is that when an image tile has high $\phi(\cdot)$ and low $\eta(\cdot)$ (quadrant ② in Figure 5), it has higher probability to be cloudy (caused by cloud, smoke, volcano ash, etc.).

Temporary outlier storage for retransmission: If an image is complex and not similar to any reference image, i.e., an outlier, then it does not allow high compression and faithful decompression. Accurate Out-of-Distribution (OOD) detection [70] is complex and hence infeasible to be conducted onboard, considering the compute limitations on satellites. So, we infer potential outliers in a lightweight manner when $\phi(\cdot)$ and $\eta(\cdot)$ are both low (quadrant ③ in Figure 5). In such cases, we transmit a LR version of the image with minimal CR as per Equation (2). We also temporarily store the image onboard to retransmit it as is on request when the transmitted version cannot be recovered faithfully on the cloud side.

4.2 On-Demand Decompression at Cloud

Mixture of Experts (MoE) based SR Framework (Path ① in Figure 4b) Following the reasoning in §3.1, in particular insight 2) to enable high CRs while maintaining reliable image recovery, rather than have one single SR model, we employ a collection of SR models in DEEPSpace to recover HR images from received LR images, each with a focus on specific case (characterized by region type and image characteristics). In essence, each of these models can be viewed as an “expert” and the overall collection resembling a Mixture of Experts (MoE) model [14, 54, 58, 73].

Figure 7 illustrates our MoE based SR framework. Training conventional MoE models [54, 58] is very resource-intensive, as both classifiers and experts are dynamically generated during the training process. We instead take a tailored approach in DEEPSpace. For satellite HR image reconstruction from LR input, as we know the key factors that influence reconstruction quality, we predefine classifier and expert configurations before training. We also limit the selection mechanism to direct each input tile to a single expert, rather than multiple experts in conventional MoE models, thereby further reducing computational overhead.

As shown in Figure 7, we allocate different models (i.e., experts) to different tiles according to the 4-step classification taken, spanning classifiers at different levels. As a result, different tiles (sub-regions) are reconstructed simultaneously by appropriate experts, ensuring both high reconstruction quality and efficient inference. We consider four levels of classifiers as follows:

- **Level 1 (Region Identification):** This stage classifies the input based on the regions it covers. Note that images of water bodies are directed to a standalone category (“Region 0” in Figure 7) using a pretrained binary classifier that takes in the location and LR image. All other images are assigned to their corresponding region types (forests, cities, etc.) based on their location.
- **Level 2:** This level directs the input to specific redundancy level in terms of SSIM after interpolation (defined as in Equation 4), which is computed onboard and included as meta data from satellites. Note that this is different to input size (of transmitted

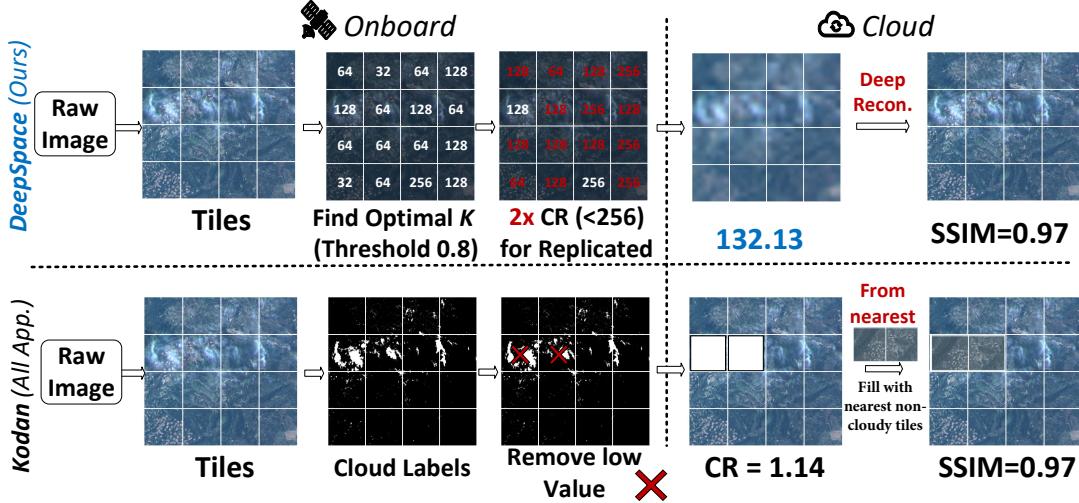


Figure 6: Example illustrating satellite-side onboard image compression and cloud-side image reconstruction with DEEPSpace (top). For comparison, Kodan [18] approach is illustrated (bottom).

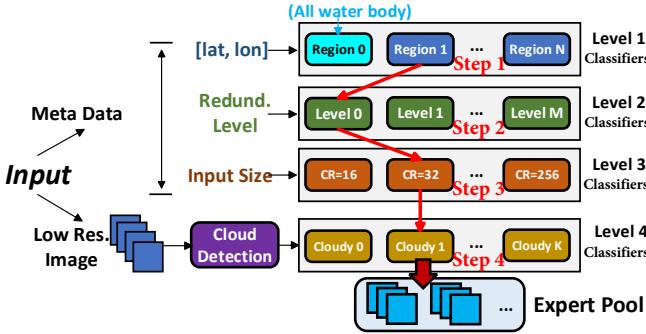


Figure 7: MoE based SR framework in DEEPSpace. Classifiers, akin to “routers” in MoE models, guide selection of an appropriate expert model for each input.

LR image) because we apply higher CR to images with near-replicate reference images. This does not change the inherent redundancy but only size.

- **Level 3:** This level of classifiers steer the input tile depending on its size.
- **Level 4:** This level computes the cloudy level with a pretrained classifier, which takes in LR image and outputs a cloudy level score after softmax. The number of cloudy levels vary by dataset, as Table 17 shows.

After four levels of classification, the input tile is directed to an expert that is best suited to generate the high quality HR version of the tile in question.

Expert Modeling with Wavelet Diffusion. We now consider how to model each expert to perform SR for a given region. Same approach can be applied to other regions. Reliable SR in this context can be formulated as an optimization problem that yields minimum error with respect to the ground truth image. This optimization can be viewed considering both the result of naive SR via interpolation and use of reference images, as shown geometrically with a 2D projection in Figure 8. The red point in Figure 8 is the result of naive SR with interpolation and the red circle shows the constraint

for this approach, defined by Equation (4) using SSIM as the quality measure:

$$T = \text{SSIM}(\text{LI}(S(I)), I), \quad (4)$$

where I is the original input image. Note that T can be measured directly onboard. Figure 8 also illustrates the ground truth’s proximity to its nearest references using BLSH, depicted as blue dashed circles.

The optimization problem can be formulated as:

$$\begin{aligned} \arg \min_{\vec{\Delta}} \quad & \sum_i |D(I' + \vec{\Delta}, I_i) - \delta_i|, \forall I_i \in \mathbb{I}_{\text{ref}} \\ \text{s.t.} \quad & \text{SSIM}(I' + \vec{\Delta}, I') = T \end{aligned} \quad (5)$$

where $I' = \text{LI}(S(I))$ is the result of naive SR, $\delta_i = D(I_{\text{real}}, I_i), \forall I_i \in \mathbb{I}_{\text{ref}}$ are the reference images computed onboard from their corresponding BLSH codes, function $D(\cdot)$ refers to the BLSH difference as follows:

$$D(a, b) = \text{NHD}(\text{BLSH}(a), \text{BLSH}(b)), \quad (6)$$

$\text{NHD}(\cdot)$ is the HD normalized by the image size.

This optimization problem targets minimizing the difference between ‘the total distance of the reconstructed image to all reference images’ and ‘the total distance of original image to all reference images’. It may have multiple solutions as neither SSIM differences nor hash codes can uniquely identify an image. Note that the constraint here means that the reconstructed image has the same SSIM with respect to the image with naive reconstruction by interpolation as that with respect to the ground truth (see equation 4).

Solving this problem optimally is challenging due to the large size and high dimensionality of satellite images, so we instead seek a data-driven heuristic solution. We can represent the change from

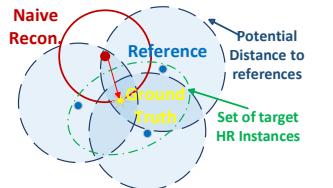


Figure 8: Reliable SR optimization illustrated.

naive reconstruction $\tilde{\Delta}$ as a neural network $\tilde{\Delta} = F(I')$ and optimize it through training, where the loss function would be the objective function. The heuristic solution can be interpreted as a super resolution task, where we find that the Wavelet Diffusion [49] outperforms other methods. Qualitatively too, wavelet diffusion is particularly well-suited for our purpose, as it focuses on reconstructing high-frequency details without compromising low-frequency information, crucial in super-resolution tasks. Moreover, wavelet diffusion is superior to other methods in terms of both output fidelity by leveraging the spectral sparsity of images, enabling the generative model to focus on the most relevant frequency bands in the training data and inference speed. This characteristic aligns perfectly with our MoE based SR framework, where each expert is trained for specific case. In satellite images, spectral sparsity is particularly pronounced, as different types of regions (e.g., forests, buildings) exhibit details at distinct spatial scales, corresponding to specific frequency bands. Given its advantages in both efficiency and reconstruction quality, all experts in DEEPSPACE are based on wavelet diffusion.

Outlier Retransmission (Path ② in Figure 4b). When an image is temporarily stored onboard, we fetch the raw image itself if it cannot be reliably recovered from its minimally compressed LR version. For this assessment, we use two measures – Low Resolution SSIM (LRS) in Equation (7) and Hash Similarity (HSIM) in Equation (8) to ground truth images, as described below.

$$\text{LRS} = \text{SSIM}(I_{\text{in}}, S(I_{\text{recon}})) \quad (7)$$

$$\text{HSIM}(I_{\text{real}}, I_{\text{recon}}) = 1 - \frac{\text{HD}(\text{BLSH}(I_{\text{real}}), \text{BLSH}(I_{\text{recon}}))}{w \times l} \quad (8)$$

Here I_{recon} refers to the reconstructed image with DEEPSPACE, whereas $S(I_{\text{recon}})$ is its LR version downsampled to match the same size as the input for reconstruction, i.e., I_{in} . Note that LRS assesses whether the model preserves the essential low-frequency details of the ground truth during reconstruction. The HSIM, on the other hand, offers a direct bit-level comparison between the reconstructed output and the ground truth, providing a more granular measure of similarity. When used together, LRS and HSIM provide a more robust and comprehensive evaluation of DEEPSPACE reconstruction performance, capturing both broad structural fidelity and fine-grained accuracy. Note that both LRS and HSIM are *post-hoc* measures that can be computed on the cloud side with the aid of LR image input and associated (small amount of) meta data from the satellite. Concretely, the meta data here refers to $\text{BLSH}(I_{\text{real}})$ along with the image size $w \times l$. Retransmission is triggered if LRS or HSIM are lower than a predefined threshold. In other words, such retransmission based on reconstruction fidelity analysis allows the quality of the recovered/received images to be always ensured.

4.3 Configurable Reconstruction Quality

Here we argue how we ensure the reconstruction quality is better than a desired and configurable error bound. SR output from an expert will have a reconstruction error bounded by Equation (9), which can be easily obtained via the geometry in Figure 8. The reconstruction should be strictly on the red circle to maintain the

SSIM after sampling and also preserve the similarity to the references. For an image with arbitrary dimensions, we have:

$$\text{Err}_{\max} = \min \left(T', \arg \min_{\mathbf{b}, \mathbf{c} \in \mathbb{I}_{\text{ref}}} T' \cdot \cos^{-1} \left(\frac{\vec{\mathbf{b}} \cdot \vec{\mathbf{c}}}{\|\vec{\mathbf{b}}\| \|\vec{\mathbf{c}}\|} \right) \right) \quad (9)$$

where vector $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are the vector representation of images, and \mathbf{a} is the result of naive SR with interpolation, T' is the Euclidean distance to the ground truth after interpolation. Vectors \mathbf{b} and \mathbf{c} are from reference set \mathbb{I}_{ref} . The proper solution of the optimization problem must significantly outperform the naive interpolation when $\exists \mathbf{b}, \mathbf{c} \in \mathbb{I}_{\text{ref}} \rightarrow \cos^{-1} \left(\frac{\vec{\mathbf{b}} \cdot \vec{\mathbf{c}}}{\|\vec{\mathbf{b}}\| \|\vec{\mathbf{c}}\|} \right) \ll 1$, which can be easily verified by checking two conditions: (i) if \mathbf{b} and \mathbf{c} are similar to the ground truth; (ii) if both \mathbf{b} and \mathbf{c} are significantly different to \mathbf{a} .

Calculating the exact error bound, however, is computationally expensive, as it involves pixel-wise comparisons of all potential vectors \mathbf{b} and \mathbf{c} across all reference images. A more efficient approach is to verify that the error bound is small by jointly using conditions (i) and (ii). Onboard operations ensure high similarity to the reference before compression; if not, lower compression ratio is used, and retransmission is prepared, thereby satisfying condition (i) by design. Simple images that remain stable through sampling and interpolation are excluded when assessing (i) as the reconstruction is indeed trivial. Condition (ii) is straightforward, as compression significantly alters the image that was initially similar to the reference before compression. Therefore, we can reasonably expect the reconstruction quality to surpass interpolation, using the SSIM after interpolation as a stringent error bound. Note that the SSIM after interpolation can be easily computed onboard, taking the fidelity configuration (in SSIM) as an input to decide CR.

5 Evaluation Methodology

5.1 Baselines

We compare against four classes of state-of-the-art baselines shown in Table 1:

- **Conventional compression** schemes such as Lánczos interpolation [36] and compressive sensing (CS) [10]. For CS, we consider three representative methods based on the comparative study in [30]. Specifically, the selected methods are ADMM [13], gOMP [65], and CoSaMP [43].
- **Autoencoders** such as DSCN [28] and VQ-VAE-2 [53] where the encoder runs on the satellite and communicates the latent space representation while the decoder on the cloud side reconstructs the image.
- **Orbital edge computing (OEC)** schemes that leverage compute onboard satellites: Kodan [18] and Earth+ [24]. Note that we compare DEEPSPACE with the OEC schemes that employ image filtering primarily to gauge their bandwidth/storage savings with respect to each other. But in fact both DEEPSPACE and OEC can be applied simultaneously to achieve even higher efficiency.
- **Deep super-resolution** techniques from the CV domain applied to satellite image data acquisition. We use the state-of-the-art wavelet diffusion (WaveDiff) [49] as the representative method from this category.

Dataset	Length	Granularity	Coverage	Region	Type	File Size
Planet-CAL [20]	20 days	Hourly	$2.5 \times 10^3 \text{ km}^2$	California	RGB	65 GB
Planet-HK [20]	20 days	Hourly	$2.5 \times 10^3 \text{ km}^2$	Hong Kong	RGB	65 GB
FarmVibes [11]	6 months	Monthly	$1 \times 10^2 \text{ km}^2$	Amazon (Cropland)	Multi-Spectral	1GB
DEN-3 [63]	2 years	Daily	$1.5 \times 10^3 \text{ km}^2$	Global	RGB	480 GB
DEN-12 [63]	2 years	Monthly	$3.6 \times 10^3 \text{ km}^2$	Global	Multi-Spectral	75 GB

Table 2: Datasets Statistics. Multi-Spectral: 12 channels.

5.2 Datasets

We compare the above baselines on a mix of custom and public datasets described in Table 2. We collect RGB images from Planet Inc’s Dove constellation [20] for California and Hong Kong for our custom dataset. In addition, we use publicly available data from Sentinel satellites [6]. These datasets include both RGB images and multi-spectral images. The RGB images in DynamicEarthNet (DEN) [63] was pre-processed by the authors to remove clouds before the dataset release. For the OEC methods that compress an image by removing clouds, we refer to the cloud coverage at same region from other climate datasets [42, 59] and apply an ideal cloud removal operated on tile level. DEN-3 refers to the RGB images and DEN-12 refers to the 12 channel Sentinel-2 multi-spectral images in dataset [63].

5.3 Performance Metrics

We consider three sets of metrics for our evaluation:

- (i) **Compression ratio:** Compression ratio (CR) is defined as $\text{CR} = \frac{\text{Original Size}}{\text{Compressed Size}}$. High compression ratio reduces bandwidth consumption and storage costs.
- (ii) **Reliable reconstruction:** We capture the similarity between reconstructed and ground truth images using two standard image quality/fidelity metrics: structural similarity (SSIM) and peak signal-to-noise ratio (PSNR)². We report both average and worst-case numbers for these metrics.

Besides SSIM and PSNR, we also consider learned fidelity metrics in the context of several deep learning based downstream applications in §7. We believe such metrics on domain specific and pixel-level real world applications are more representative than general learned metrics like LPIPS [74].

- (iii) **Compute and storage overheads:** We measure the onboard computing speed, decompression speed, and storage requirements for different methods.

6 Evaluation Results

6.1 DEEPSpace Brings Orders of Magnitude Compression Gains

First, we analyze DEEPSpace’s ability to achieve high compression gains. For this evaluation, here we first focus on comparisons with OEC based schemes – Kodan and Earth+; other baselines operate

²We set the maximum PSNR to be 50 dB in our evaluations. This is because PSNR is unbounded and some corner cases produce extreme high PSNR values that can hide more challenging cases when taking the average. For example, simple images like thick cloud, basically purely white, can be easily recovered with perfect fidelity (> 80 dB). We empirically find 50 dB to be a good maximum value from our datasets. Setting a maximum PSNR value is common in similar contexts. For instance, video encoding [22] and quality assessment [44] generally use a maximum PSNR between 50 dB to 60 dB.

with fixed compression ratios and are evaluated later in §6.2. We also visualize the reconstruction results on city-scale in Appendix A.4.1.

We report the reconstruction fidelity in Figure 9a and compression ratio in Figure 9b for the different approaches (more details in Appendix A.4.2, Table 18). DEEPSpace achieves high-fidelity reconstructions while achieving a two orders of magnitude higher compression ratio.

While Earth+ and Kodan reject tiles that are not relevant to applications (e.g., cloudy tiles), DEEPSpace performs per-tile compression with an appropriate compression ratio. This allows DEEPSpace to achieve a higher compression ratio while maintaining reconstruction fidelity. Earth+ replaces high-similarity images with the reference copy which allows it to achieve higher compression ratio than Kodan, but it also leads to worse fidelity.

DEEPSpace also requires lower uplink usage as shown in Figure 10. DEEPSpace’s uplink usage (about 10 Kbps) translates to a small fraction of the 256 Kbps uplink capacity of deployed earth observation satellites [21], while the change based Earth+, DCSN and VQ-VAE-2, requires 230 Kbps, 380 Kbps, and 396 Kbps respectively, fully occupying or exceeding the uplink capacity. The full setup of Earth+ (*i.e.*, detect all differences) requires > 2 Mbps uplink throughput, which cannot be met by many deployed satellites [21]. Note that Kodan is not included in Figure 10 because no model update is specified in the original paper [18], although such updates are essential for long term deployment.

Results for onboard storage requirement are shown in Figure 11, where the “reference” refers to the reference hash code, *i.e.*, with BLSH, in DEEPSpace and images in Earth+, “model” refers to the trained models for onboard inference, and “captured” refers to the image storage onboard for further processing. DEEPSpace uses less onboard storage than the baselines, and these storage gains are achieved through the use of BLSH. Instead of sending low resolution images as references in Earth+, we utilize hashing to represent reference images, resulting in significantly smaller sizes and higher computational efficiency.

6.2 DEEPSpace Provides High Reliability for a Given Compression Ratio

We compare DEEPSpace with the other data-driven alternatives, which require setting a fixed CR. Table 3, 4, 5 and 6 show the fidelity performance of different methods for a given CR with different datasets.

For CS methods, CR that can be enabled while maintaining fidelity is depending on the data sparsity. For image data, sparsity is calculated as the ratio of near-zero frequency components to the image size. Intuitively, greater the sparsity higher the CR. With the satellite image data we consider, we find that sparsity is less than 87.5% (as opposed to the ideal 99%). This in turn limits the effective CR with CS methods to 8. Therefore, we first compare DEEPSpace with different CS methods – gOMP [65], CoSaMP [43] and ADMM [13] – for CRs up to 8. These results are shown in Table 7. As expected, these CS methods perform quite well with DEEPSpace offering only marginal gain. Among the CS methods, gOMP yields the best fidelity performance. We then include gOMP as well as the general ADMM in the evaluations focusing on the higher CR regime we target. These results shown in Table 3, 4, 5

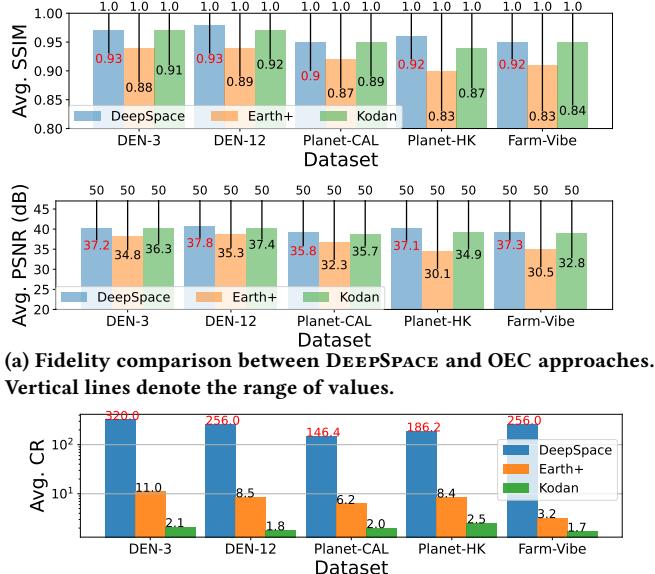


Figure 9: DEEPSpace achieves two orders of magnitudes higher CR with similar/better fidelity.

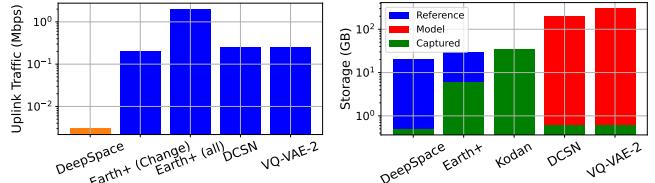


Figure 10: Uplink usage of different methods. Figure 11: Onboard storage usage of different methods.

Dataset	Planet-CAL					
	CR=16		CR=64		CR=256	
	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑
DEEPSpace	0.97	40.2	0.93	38.3	0.86	33.9
WaveDiff	0.93	36.6	0.89	36.2	0.79	32.9
SR3	0.92	35.1	0.90	35.4	0.77	30.8
VQ-VAE-2	0.92	35.4	0.85	34.8	0.78	31.2
DCSN	0.90	35.2	0.83	33.2	0.75	30.3
CS (gOMP)	0.89	35.6	0.81	32.0	0.71	28.9
CS (ADMM)	0.91	35.9	0.81	31.5	0.68	28.8
Interpolation	0.89	35.2	0.80	31.7	0.69	27.3

Table 3: Fidelity for a given CR with Planet-CAL dataset.

and 6 demonstrate that CS methods sacrifice fidelity with increase in CR, whereas DEEPSpace maintains fidelity at higher CRs.

Surprisingly, the autoencoder methods do not exhibit better fidelity under high compression ratios compared to super-resolution models. Our findings indicate that these methods fail to ensure fidelity, even with low CRs. The root cause lies in the encoder's inability to guarantee a reliable latent space representation, particularly when using a variational encoder (e.g., DCSN). Similarly, the decoder works as a black box and provides no guarantee or performance estimation.

Dataset	Planet-HK					
	CR=16		CR=64		CR=256	
	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑
DEEPSpace	0.98	40.8	0.95	39.7	0.91	37.3
WaveDiff	0.93	37.8	0.90	36.7	0.80	33.5
SR3	0.93	36.6	0.88	35.3	0.81	34.4
VQ-VAE-2	0.92	36.8	0.88	34.6	0.79	32.5
DCSN	0.90	35.5	0.84	33.7	0.77	31.4
CS (gOMP)	0.92	33.1	0.84	34.0	0.70	31.2
CS (ADMM)	0.89	35.9	0.82	34.0	0.70	31.5
Interpolation	0.90	35.8	0.81	32.9	0.72	31.9

Table 4: Fidelity for a given CR with Planet-HK dataset.

Dataset	FarmVibes					
	CR=16		CR=64		CR=256	
	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑
DEEPSpace	0.99	41.5	0.94	38.4	0.91	37.2
WaveDiff	0.94	38.9	0.89	33.9	0.83	31.5
SR3	0.88	36.3	0.85	34.5	0.80	31.8
VQ-VAE-2	0.90	36.6	0.90	35.6	0.84	34.2
DCSN	0.90	35.2	0.83	33.2	0.75	30.3
CS (gOMP)	0.90	35.8	0.83	33.2	0.73	32.9
CS (ADMM)	0.92	36.5	0.81	30.5	0.66	25.5
Interpolation	0.92	36.8	0.82	31.8	0.69	29.8

Table 5: Fidelity for a given CR with FarmVibes dataset.

Dataset	DEN-3					
	CR=16		CR=64		CR=256	
	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑
DEEPSpace	0.98	41.5	0.98	40.7	0.98	40.3
WaveDiff	0.98	40.5	0.93	38.4	0.88	36.7
SR3	0.975	40.3	0.91	37.2	0.86	34.9
VQ-VAE-2	0.97	39.9	0.92	36.8	0.84	34.6
DCSN	0.95	40.1	0.88	36.1	0.81	32.5
CS (gOMP)	0.94	38.9	0.84	34.8	0.79	31.8
CS (ADMM)	0.98	42.2	0.84	34.5	0.70	28.0
Interpolation	0.94	38.8	0.83	33.5	0.74	29.8

Table 6: Fidelity for a given CR with DEN-3 dataset.

Reliability analysis: DEEPSpace achieves a fidelity that is significantly better than the error tolerance configuration of $\phi(\cdot) \geq 0.85$. Moreover, DEEPSpace estimates reconstruction quality to avoid errors caused by outliers. From the results shown in Table 9, we observe that DEEPSpace consistently shows a minimum observed SSIM fidelity that is higher than the 0.85 threshold. We also evaluated the impact of tuning this threshold; results are included in Table 8. While higher threshold contributes to a significant higher fidelity, it also limits the compression ratio. Empirically, we find that setting the threshold to 0.85 yields best balance between compression gain and reliability.

In Figure 12, we show the near-linear correlation between LRS and actual SSIM performance of reconstruction – a low LRS value always points to poor reconstruction quality because the input is an outlier to the model. This implies that DEEPSpace can use LRS to estimate the fidelity loss without transmitting the raw file. For a small fraction of the images (0.1% to 0.3%), marked by \times , the LRS values are below the threshold (i.e., unlikely to yield a high-fidelity reconstruction). These images need to be retransmitted at

Dataset	Planet-CAL								Planet-HK								FarmVibes								DEN-3							
	CR=2		CR=4		CR=8		CR=2		CR=4		CR=8		CR=2		CR=4		CR=8		CR=2		CR=4		CR=8									
Method	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑				
DEEPSpace	0.99	45.1	0.98	42.4	0.97	41.1	0.99	45.8	0.98	43.7	0.98	41.8	0.99	44.5	0.99	43.2	0.99	42.8	0.995	46.2	0.99	43.3	0.98	41.8								
ADMM	0.97	42.0	0.95	40.1	0.92	35.5	0.99	44.3	0.97	40.8	0.94	38.6	0.99	43.3	0.99	43.2	0.96	39.5	0.99	44.0	0.98	42.8	0.96	39.3								
gOMP	0.98	43.9	0.97	41.6	0.94	38.1	0.99	44.5	0.98	43.2	0.95	39.9	0.99	44.3	0.99	43.5	0.98	41.5	0.995	45.5	0.99	44.2	0.97	40.9								
CuSaMP	0.99	44.3	0.96	40.5	0.93	36.8	0.99	44.5	0.98	42.8	0.95	39.1	0.99	45.0	0.99	43.3	0.97	40.6	0.99	44.9	0.99	43.4	0.97	39.5								

Table 7: Fidelity of reconstruction for given CR with different compressive sensing based methods. DEEPSpace shows higher fidelity across different datasets and CR.

Dataset	Exp. SSIM $\phi(\cdot) \geq 0.9$					Exp. SSIM $\phi(\cdot) \geq 0.85$ (Default)					Exp. SSIM $\phi(\cdot) \geq 0.8$				
	SSIM (>Min.) ↑	PSNR ↑	CR ↑	LRS ↑	HSIM ↑	SSIM (>Min.) ↑	PSNR ↑	CR ↑	LRS ↑	HSIM ↑	SSIM (>Min.) ↑	PSNR ↑	CR ↑	LRS ↑	HSIM ↑
Planet-CAL	0.96 (0.92)	39.5	96.3	0.993	0.99	0.95 (0.90)	39.0	146.4	0.99	0.99	0.91 (0.87)	37.6	195.2	0.98	0.98
Planet-HK	0.96 (0.93)	40.6	124.2	0.995	0.99	0.96 (0.92)	39.8	186.2	0.993	0.99	0.93 (0.89)	38.6	244	0.99	0.98
Farm Vibe	0.97 (0.95)	41.1	154.2	0.99	0.98	0.95 (0.92)	38.9	256	0.98	0.97	0.94 (0.92)	39.2	256	0.97	0.97
DEN-3	0.98 (0.93)	39.8	156	0.995	0.992	0.98 (0.93)	39.8	310	0.992	0.99	0.94 (0.90)	37.9	366	0.992	0.99
Average	0.96 (0.918)	39.4	132.7	0.993	0.988	0.96 (0.918)	39.4	224.65	0.99	0.985	0.93 (0.89)	38.3	265.3	0.983	0.98

Table 8: Fidelity of reconstruction and CRs for DEEPSpace with different fidelity thresholds.

Metric	SSIM (Min) ↑	PSNR ↑	CR ↑	LRS ↑	HSIM ↑
Planet-CAL	0.95 (0.90)	39.0	146.4	0.99	0.99
Planet-HK	0.96 (0.92)	39.8	186.2	0.993	0.99
Farm Vibes	0.95 (0.92)	38.9	256	0.98	0.97
DEN-3	0.98 (0.93)	39.8	310	0.992	0.99
Average	0.96 (0.918)	39.4	224.65	0.99	0.985

Table 9: Reliability analysis with DEEPSpace using $\phi(\cdot) \geq 0.85$.

full resolution. HSIM, on the other hand, is a direct BLSH based similarity measurement. We find that it also shows similar correlation behavior with actual SSIM performance.

6.3 Computational and Storage Efficiency

The onboard computation of DEEPSpace involves matrix computations and can be significantly accelerated by 2 to 3 orders of magnitude with a cheap graphics card. As mentioned before, such graphics cards are increasingly available on satellites. With NVIDIA Jetson GPU, as Figure 13 shows, DEEPSpace has about 2 orders of magnitudes faster processing speed than existing OEC methods. This is because DEEPSpace does *not* deploy ML models onboard, unlike past work in orbital edge computing. The faster processing speed on satellites also translates to equivalent gains in energy consumption, considering the same NVIDIA Jetson onboard satellite compute platform.

Storage saving and cost evaluation: The high efficiency of DEEPSpace also makes the long-term fine granularity storage of global observation data efficient and more affordable. Storage costs are incurred on a pay-as-you-go basis, even for infrequently accessed data. Providing immediate access to tens of petabytes of satellite imagery costs millions of dollars per month [7, 8]. With >200x compression, DEEPSpace makes cloud storage affordable besides reducing data transmitted over the downlink path from satellite to cloud. The data reconstruction is done on-demand.

Based on our tests and previous works [49], it takes <0.1s for DEEPSpace to reconstruct a 1024×1024 RGB image with a single NVIDIA A series GPU, corresponding to a satellite raw file with around 8.5 MB. Since DEEPSpace uses MoE structure, different tiles can be reconstructed in parallel, where each GPU contributes $640 \sim 800$ Mbps throughput. This means the end user only needs 4 common GPUs to achieve the maximum write speed of a SATA SSD during decompression. This fast reconstruction speed is comparable

Method	SSIM (>Min.)↑	PSNR↑	CR↑	LRS↑	HSIM↑
DEEPSpace	0.95 (0.9)	39.0	146.4	0.99	0.99
No Adaptive	0.91 (0.85)	36.9	82.6	0.98	0.97
No Re-transmission	0.95 (0.82)	37.2	147.5	0.99	0.99
w/o MoE	0.90 (0.85)	36.3	92.5	0.93	0.92
MoE w/o Level3	0.92 (0.87)	37.4	112.4	0.95	0.93
MoE w/o Level4	0.93 (0.88)	36.8	128	0.96	0.95
WaveDiff (base)	0.81 (0.69)	33.2	144	0.88	0.85
w/o Deep Reconstruction*	0.83 (0.75)	33.5	71.6	1	0.83

Table 10: Ablation study of DEEPSpace. *: Using naive interpolation for SR.

with most existing lightweight methods, and the comparison is showed in Figure 14.

6.4 Ablation Study

We carry ablation study of DEEPSpace considering the Planet-CAL dataset, which includes various coverage types. The error bound is set to 0.85 when CR = 64. The result is showed in Table 10. Overall, the adaptive nature of the compression in DEEPSpace contributes to the CR significantly, where a constant CR for all images can only achieve 82.6 CR with worse fidelity. Retransmission occurs infrequently, affecting less than 1 in every 100 tiles. As Table 10 shows, the worst case only has a 0.82 SSIM. Enabling retransmission improves the worst-case SSIM performance to 0.9, with only a slight reduction in CR. The MoE plays a critical role in improving reconstruction quality, boosting average SSIM fidelity by 0.05. Note that the case without Level 1 classifier overlaps with WaveDiff, where we use one model that generalizes over all regions. Similarly, the ablation test of level 2 classifier is included in “No adaptive”, where the redundancy level is not computed. It also reduces reconstruction uncertainty, minimizing unnecessary retransmissions. We also demonstrate the performance of wavelet diffusion at comparable CR. Without the mechanisms in DEEPSpace, a single SR model based on wavelet diffusion struggles to accurately reconstruct the image at high CR. The core principle of the deep reconstruction method is demonstrated by substituting wavelet diffusion with interpolation, resulting in a notable decline across all performance metrics (LRS=1 due to interpolation).

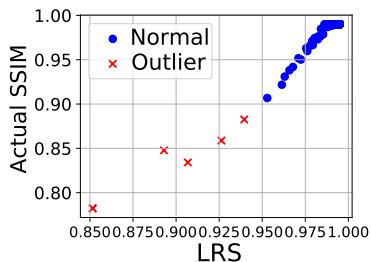


Figure 12: Actual SSIM has near-linear correlation with LRS. \times : re-transmission triggered.

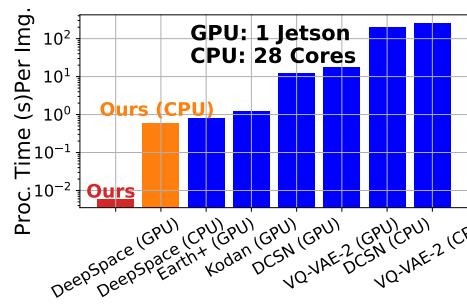


Figure 13: Onboard processing time per image of different methods

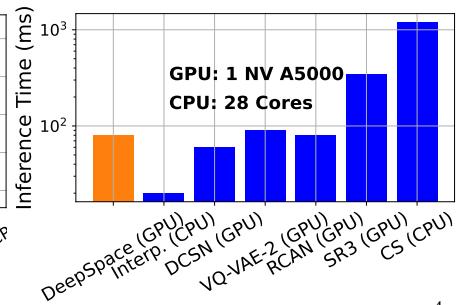


Figure 14: Decompression Speed of a $10^4 \times 10^4$ pixels region, 3 channels.

Method	Med. Resp. Time (h)	Rel. Size Accuracy	Class. Accuracy
Interpolation	0.45	0.55	0.63
CS (gOMP)	8.55	0.65	0.65
Kodan	39.8	0.74	0.71
Earth+	24.4	0.72	0.69
DEEPSPACE	0.48	0.993	0.885
Ground Truth	71.71	1	0.89

Table 11: Wild fire detection result with the image recovery by different methods.

7 Case Studies

In this section, we evaluate three use cases of satellite imagery and how they are affected by the use of reconstructed imagery instead of ground truth images.

7.1 Wildfire Detection

We compare the accuracy of wildfire detection on reconstructed imagery and ground truth images. We use the classical U-net [55] structure for the classification task, which is a common method used for this task [52]. We train the model with Planet-CAL dataset between July and August 2021, where the wildfire location is well labeled. Note that the cloud detection used in Kodan [18] and Earth+ [24] can confuse between clouds and smoke, and have significant false detection rates (i.e., can reject images with smoke). We also report end-to-end delivery time using the simulation framework proposed in Serval [61].

The quantitative result on response time and accuracy is shown in Table 11, where we use Median Response Time (Med. Resp. Time) to measure the response speed, Relative Size (Rel. Size) accuracy to show the quality of fire coverage measurement, and classification (Class.) accuracy to show whether we can identify areas containing wildfire. DEEPSPACE achieves best performance on both response time and accuracy (size measure in column 3 and label accuracy in column 4). The improvement in response time is due to the high CR achieved by DEEPSPACE, thereby not limited by the constrained satellite-cloud network path. This result demonstrates that DEEPSPACE reduces the disaster warning latency from few days to <30 minutes, enabling rapid and reliable detection.

Method	F1 Score \uparrow	Recall \uparrow	Precision \uparrow	SSIM \uparrow	PSNR \uparrow	CR \uparrow
Interpolation	0.42	0.69	0.30	0.85	35.2	225
CS (gOMP)	0.86	0.82	0.91	0.91	36.6	8
Kodan	0.98	0.96	0.99	0.93	36.9	2.4
Earth+	0.84	0.81	0.90	0.92	36.3	10.1
DEEPSPACE $\eta(\cdot) \geq 0.85$	0.92	0.95	0.91	0.97	41.2	232.5
DEEPSPACE $\eta(\cdot) \geq 0.93$	0.98	0.97	0.98	0.98	41.6	98.6
Ground Truth	1	1	1	1	N/A	1

Table 12: Small plastic detection with reconstructed imagery from different methods.

7.2 Plastic Detection in Oceans

Next, we evaluate DEEPSPACE on an extremely fine-grained task – plastic detection in oceans. This requires sub-pixel level analysis and very high reconstruction fidelity. We adopt the data and detection methods introduced in [11, 12], aiming to detect the objects with a shape that is close or even smaller than pixel size of multiple-channel satellite images. We evaluate the detection performance with pixel level binary labels, where each label can be plastic or no-plastic. To conduct the experiment, we simply replace the original image with reconstructed images with different methods, leaving the rest of the pipeline exactly the same as [11, 12].

We test two configurations of DEEPSPACE on this task with different values of $\eta(\cdot)$. A higher $\eta(\cdot)$ increases the sensitivity of DEEPSPACE and is better suited for this task. Results in Table 12 show that DEEPSPACE can achieve a reliable detection at the cost of lower compression ratio, which is still 50x higher than the state-of-the-art method with high accuracy. The other methods failed in this task as they cannot preserve or recover pixel-level information.

7.3 Land Use and Cropland Classification

Imagery based remote sensing has been extensively utilized for land-use classification in agriculture and other sectors [20, 23, 29, 31, 50]. Unlike a use case like wildfire detection, land use measurement prioritizes accuracy and cost over response speed.

We first verify if the reconstructed images can be recognized by other ML models by applying the reconstruction with DEEPSPACE to the image segmentation tasks in [33, 63]. The output is illustrated in Figure 15. Segmentation tests the performance on more general detection and measurement tasks. To quantify the performance, we count the number of object with Segment Anything (SA) [33], named as "#SA". From the result in Table 13, we see DEEPSPACE, Kodan, and Earth+ all show very high fidelity on cropland and water

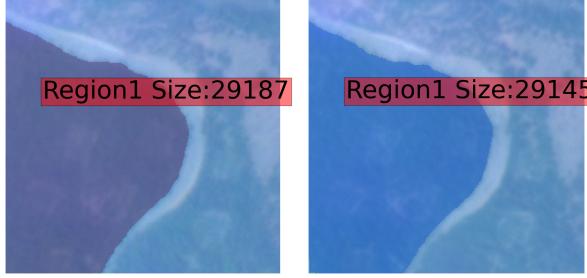


Figure 15: (L) Orig. result. (R) Recon. by DEEPSpace 256x.

Method	Rel. #SA	Rel. Size	Classification	CR↑
	Error↓	Accuracy ↑	Accuracy ↑	
Interpolation	0.4	0.45	0.31	256
CS (gOMP)	0.10	0.89	0.80	16
Kodan	0.01	0.997	0.87	2.1
Earth+	0.02	0.99	0.86	9.8
DEEPSpace	0.01	0.995	0.86	256
Ground Truth	0	1	0.87	1

Table 13: Segmentation and classification with the reconstruction of DEEPSpace.

body detection tasks³. But DEEPSpace is two orders of magnitude more efficient than these other OEC methods, while providing an experience akin to using real data. For tasks requiring detailed classification and precise size measurement with pixel-level accuracy, both DEEPSpace and conventional OECs deliver equally reliable results. This is because such tasks are less time-sensitive, and in some cases, even older image copies can produce accurate outcomes.

We now zoom-in to the particular task of cropland classification, a classical application for satellite image data, to further showcase the ability of DEEPSpace to seamlessly and efficiently support such a real world application. Our labels of cropland comes from the official dataset of California [3]. We use the satellite images from the Planet dataset captured from the same year as this official dataset. We label these satellite images by first segmenting with [33] and then labeling them based on [3]. We consider the pretrained GFM model [40] for the classification task and treat the performance with original images with this model as the “ground truth”. With respect to this ground truth, we evaluate the classification performance of reconstructed images with different methods including DEEPSpace.

Results in Table 14 show that DEEPSpace not only achieves a high CR of 146.4 but also a cropland classification performance similar to ground truth. While Kodan and Earth+ can achieve high classification performance, they yield significantly low CR (around 2-6). The other CS and ML based methods sacrifice the classification performance to achieve as high a CR as DEEPSpace, thus making those methods unsuitable for this task.

8 Discussion

Here we discuss limitations of DEEPSpace and some potential future research directions. The approach taken in DEEPSpace is inspired by

³To showcase the fidelity on typical remote sensing tasks, we train a U-Net classification model with the cropland type label from [5].

Method	Precision↑	Recall ↑	F1 Score ↑	CR ↑
Interpolation	0.40	0.45	0.43	144
CS (gOMP)	0.45	0.51	0.48	144
VQ-VAE-2	0.70	0.75	0.73	144
Kodan	0.91	0.95	0.93	2.0
Earth+	0.88	0.92	0.90	6.2
WaveDiff	0.79	0.74	0.76	144
DEEPSpace	0.90	0.95	0.92	146.4
Ground Truth	0.91	0.95	0.93	1

Table 14: Cropland classification with pretrained GFM model for different reconstructions.

prior efforts in the context of multimedia communications [68, 72], where the video frames are sent at a low bit-rate or low resolution, and then recovered on the receiver side with custom DL methods – frame prediction [68] or super resolution [72]. More recently in NetGSR [60], the idea of super resolution was used in the temporal domain in the network monitoring context to enable efficient telemetry data collection. DEEPSpace is similar in spirit to these earlier works but its design is tailored to account for the unique characteristics and constraints of satellite based earth observation setting through novel techniques like MoE-based recovery system. As with these prior systems, there is a cost associated with training DEEPSpace in terms of GPU time, especially to cover all regions on earth. However, this cost is outweighed by the value it creates across many application areas. A potential approach to tackle the training cost is to build a pre-trained global scale foundation model along the lines of [32], and then fine tune it for different types of satellites and images.

Besides the single-image SR focus of DEEPSpace, there exist other satellite image processing works. In particular, multiple image SR [35, 41] leverage LR images from different sources to obtain HR images. This provokes us to extend DEEPSpace to cover different data sources (*i.e.*, type of satellite) to achieve even higher compression gains with good generalization while maintaining reconstruction quality.

Recent works also leverage multi-modal language models for satellite image processing [32]. An aspect for future work along these lines is to reconstruct images with customized features as prompts. Another direction for future work is to exploit higher onboard compute resources [26] and uplink bandwidth [64] where available to unlock even higher compression gains.

9 Conclusions

Efficient high-quality satellite image data acquisition is challenging due to large data volumes, downlink bottlenecks and high cloud storage costs. Earth observation satellites are also limited in their compute capabilities. We introduce DEEPSpace, a deep learning-based system that compresses satellite images by hundreds of times in real-time onboard, while ensuring faithful reconstruction with fidelity guarantees. DEEPSpace performs lightweight computation onboard the satellite and shifts the compute burden to the cloud, where images are decompressed on demand.

10 Ethics Statement

This work does not require ethics approval for reporting human subjects or potentially sensitive data.

Acknowledgments

We thank our shepherd and the anonymous reviewers for their helpful comments and suggestions that greatly improved this paper.

References

- [1] [n. d.]. Multi-Modal GeoSpatial ML Models for Agriculture and Sustainability. <https://github.com/microsoft/farmvibes-ai>. ([n. d.]).
- [2] [n. d.]. NASA Turns to the Cloud for Help With Next-Generation Earth Missions. <https://sealevel.nasa.gov/news/226/nasa-turns-to-the-cloud-for-help-with-next-generation-earth-missions/>. ([n. d.]).
- [3] [n. d.]. Statewide Crop Mapping Map Service. <https://catalog.data.gov/dataset/statewide-crop-mapping-5fcda/resource/5811cc03-2687-4ad6-855f-0b85c573bc6c>. ([n. d.]).
- [4] 2024. Planet Labs PBC Announces Real-Time Insights Technology Using NVIDIA Jetson Platform. <https://tinyurl.com/y65rafab>. (June 2024).
- [5] California Natural Resources Agency. 2021. 2021 Statewide Crop Mapping Geo-database. <https://data.cnra.ca.gov/dataset/statewide-crop-mapping>. (2021).
- [6] The European Space Agency. 2024. Sentinel-2. <https://sentinels.copernicus.eu/web/sentinel/home>. (2024).
- [7] AWS. 2024. Pricing calculator. <https://calculator.aws/>. (2024).
- [8] Microsoft Azure. 2024. Pricing calculator. <https://azure.microsoft.com/en-us/pricing/calculator/?cdn=disable>. (2024).
- [9] Pascal Bacchus, Renaud Fraisse, Aline Roumy, and Christine Guillemot. 2022. Quasi lossless satellite image compression. In *IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 1532–1535.
- [10] Richard G Baraniuk. 2007. Compressive sensing [lecture notes]. *IEEE signal processing magazine* 24, 4 (2007), 118–121.
- [11] Lauren Biermann, Daniel Clewley, Victor Martinez-Vicente, and Konstantinos Topouzelis. 2020. Finding plastic patches in coastal waters using optical satellite data. *Scientific reports* 10, 1 (2020), 5364.
- [12] Lauren Biermann, Daniel Clewley, Victor Martinez-Vicente, and Konstantinos Topouzelis. 2020. A plastic detection pipeline for coastal waters. <https://github.com/tsdataclinic/PMILDataPipeline?tab=readme-ov-file>. (2020).
- [13] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning* 3, 1 (2011), 1–122.
- [14] Weilin Cai, Juyong Jiang, Fan Wang, Jing Tang, Sunghun Kim, and Jiayi Huang. 2024. A Survey on Mixture of Experts. (2024). arXiv:cs.LG/2407.06204 <https://arxiv.org/abs/2407.06204>
- [15] Kejie Chen, Jean-Philippe Avouac, Saif Aati, Chris Milliner, Fu Zheng, and Chuang Shi. 2020. Cascading and pulse-like ruptures during the 2019 Ridgecrest earthquakes in the Eastern California Shear Zone. *Nature communications* 11, 1 (2020), 22.
- [16] Min Chen, Xiaobo Shi, Yin Zhang, Di Wu, and Mohsen Guizani. 2017. Deep feature learning for medical image analysis with convolutional autoencoder neural network. *IEEE Transactions on Big Data* 7, 4 (2017), 750–758.
- [17] Vinicius Alves de Oliveira, Marie Chabert, Thomas Oberlin, Charly Poulliat, Mickael Bruno, Christophe Latry, Mikael Carlan, Simon Henrot, Frederic Falzon, and Roberto Camarero. 2022. Satellite image compression and denoising with neural networks. *IEEE Geoscience and Remote Sensing Letters* 19 (2022), 1–5.
- [18] Bradley Denby, Krishna Chintalapudi, Ranveer Chandra, Brandon Lucia, and Shadi Noghabi. 2023. Kodam: Addressing the computational bottleneck in space. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 3*. 392–403.
- [19] Bradley Denby and Brandon Lucia. 2020. Orbital Edge Computing: Nanosatellite Constellations as a New Class of Computer System. In *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS '20)*. Association for Computing Machinery, New York, NY, USA, 939–954. <https://doi.org/10.1145/3373376.3378473>
- [20] Kiruthika Devaraj, Ryan Kingsbury, Matt Ligon, Joseph Breu, Vivek Vittaldev, Bryan Klofas, Patrick Yeon, and Kyle Colton. 2017. Dove high speed downlink system. (2017).
- [21] Kiruthika Devaraj, Matt Ligon, Eric Blossom, Joseph Breu, Bryan Klofas, Kyle Colton, and Ryan Kingsbury. 2019. Planet high speed radio: Crossing Gbps from a 3U cubesat. (2019).
- [22] SVT-AV1 Developers. 2025. SVT-AV1. <https://gitlab.com/AOMediaCodec/SVT-AV1>. (2025).
- [23] Chen Du, Yitong Wang, Zhicheng Yang, Hang Zhou, Mei Han, and Jui-Hsin Lai. 2023. Parcs: A deployment-oriented ai system for robust parcel-level cropland segmentation of satellite images. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 15775–15781.
- [24] Kunhai Du, Yihua Cheng, Peder Olsen, Shadi Noghabi, Ranveer Chandra, and Junchen Jiang. 2025. Earth+: On-board satellite imagery compression leveraging historical earth observations. *Proceedings of the 30th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 1* (2025).
- [25] Zhihao Duan, Ming Lu, Jack Ma, Yuning Huang, Zhan Ma, and Fengqing Zhu. 2023. Qarv: Quantization-aware resnet vae for lossy image compression. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2023).
- [26] Nafiseh Ghasemi, Jon Alvarez Justo, Marco Celesti, Laurent Despoisse, and Jens Nieke. 2025. Onboard Processing of Hyperspectral Imagery: Deep Learning Advancements, Methodologies, Challenges, and Emerging Trends. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2025).
- [27] Ritwik Gupta, Bryce Goodman, Nirav Patel, Ricky Hosfelt, Sandra Sajeev, Eric Heim, Jigal Doshi, Keane Lucas, Howie Choset, and Matthew Gaston. 2019. Creating xBD: A dataset for assessing building damage from satellite imagery. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 10–17.
- [28] Chih-Chung Hsu, Chia-Hsiang Lin, Chi-Hung Kao, and Yen-Cheng Lin. 2020. DCSCN: Deep compressed sensing network for efficient hyperspectral data transmission of miniaturized satellite. *IEEE Transactions on Geoscience and Remote Sensing* 59, 9 (2020), 7773–7789.
- [29] Pallavi Jain, Bianca Schoen-Phelan, and Robert Ross. 2020. Automatic flood detection in Sentinel-2 images using deep convolutional neural networks. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*. 617–623.
- [30] Jon Alvarez Justo, Daniela Lupu, Milica Orlandić, Ion Necula, and Tor Arne Johansen. 2022. A comparative study of compressive sensing algorithms for hyperspectral imaging reconstruction. In *2022 IEEE 14th Image, Video, and Multi-dimensional Signal Processing Workshop (IVMSP)*. IEEE, 1–5.
- [31] Nikolai Kalischek, Nico Lang, Cécile Renier, Rodrigo Caye Daudt, Thomas Adadoah, William Thompson, Wilma Blaser-Hart, Rachael Garrett, Konrad Schindler, and Jan Dirk Wegner. 2023. Cocoa plantations are associated with deforestation in Côte d'Ivoire and Ghana. *Nature Food* 4, 5 (2023), 384–393.
- [32] Samar Khanna, Patrick Liu, Linqi Zhou, Chenlin Meng, Robin Rombach, Marshall Burke, David Lobell, and Stefano Ermon. 2023. Diffusionsat: A generative foundation model for satellite imagery. *arXiv preprint arXiv:2312.03606* (2023).
- [33] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. 2023. Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 4015–4026.
- [34] Devaraj Kiruthika, Ligon Matt, Blossom Eric, Breu Joseph, Klofas Bryan, Colton Kyle, and Kingsbury Ryan. 2019. Planet High Speed Radio: Crossing Gbps from a 3U CubeSat. *33rd Annual AIAA/USU Conference on Small Satellites* 10, 1 (2019), 5364.
- [35] Paweł Kowaleczko, Tomasz Tarasiewicz, Maciej Ziaja, Daniel Kostrzewa, Jakub Nalepa, Przemysław Rokita, and Michał Kawulok. 2023. A real-world benchmark for Sentinel-2 multi-image super-resolution. *Scientific Data* 10, 1 (2023), 644.
- [36] Cornelius Lanczos. 1996. *Linear differential operators*. SIAM.
- [37] Nico Lang, Walter Jetz, Konrad Schindler, and Jan Dirk Wegner. 2023. A high-resolution canopy height model of the Earth. *Nature Ecology & Evolution* 7, 11 (2023), 1778–1789.
- [38] Huawei Liu, Wenhua Zhou, Zongda Wu, Shichao Zhang, Gang Li, and Xuelong Li. 2023. Refining codes for locality sensitive hashing. *IEEE Transactions on Knowledge and Data Engineering* (2023).
- [39] BN Madhukar and R Narendra. 2013. Lanczos resampling for the digital processing of remotely sensed images. In *Proceedings of International Conference on VLSI, Communication, Advanced Devices, Signals & Systems and Networking (VCASAN-2013)*. Springer, 403–411.
- [40] Matías Mendieta, Boran Han, Xingjian Shi, Yi Zhu, and Chen Chen. 2023. Towards geospatial foundation models via continual pretraining. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 16806–16816.
- [41] Andrea Bordone Molini, Diego Valsesia, Giulia Fracastoro, and Enrico Magli. 2019. Deepsum: Deep neural network for super-resolution of unregistered multitemporal images. *IEEE Transactions on Geoscience and Remote Sensing* 58, 5 (2019), 3644–3656.
- [42] NASA. 2024. Cloud Fraction. https://earthobservatory.nasa.gov/global-maps/MODAL2_M_CLD_FR. (2024).
- [43] Deanna Needell and Joel A Tropp. 2009. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and computational harmonic analysis* 26, 3 (2009), 301–321.
- [44] Netflix. 2025. VMAF - Video Multi-Method Assessment Fusion. <https://gitlab.com/AOMediaCodec/SVT-AV1>. (2025).
- [45] Andrew Ng et al. 2011. Sparse autoencoder. *CS294A Lecture notes* 72, 2011 (2011), 1–19.
- [46] Jim Nilsson and Tomas Akenine-Möller. 2020. Understanding ssim. *arXiv preprint arXiv:2006.13846* (2020).
- [47] Joel R Norris, Robert J Allen, Amato T Evan, Mark D Zelinka, Christopher W O'Dell, and Stephen A Klein. 2016. Evidence for climate change in the satellite cloud record. *Nature* 536, 7614 (2016), 72–75.
- [48] Victor Mackenhauer Olsen, Rasmus Fensholt, Pontus Olofsson, Rogerio Bonifacio, Van Butsic, Daniel Druce, Deepak Ray, and Alexander V Prishchepov. 2021. The impact of conflict-driven cropland abandonment on food insecurity in South

- Sudan revealed using satellite remote sensing. *Nature Food* 2, 12 (2021), 990–996.
- [49] Hao Phung, Quan Dao, and Anh Tran. 2023. Wavelet diffusion models are fast and scalable image generators. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10199–10208.
- [50] Peter Potapov, Svetlana Turubanova, Matthew C Hansen, Alexandra Tyukavina, Viviana Zelles, Ahmad Khan, Xiao-Peng Song, Amy Pickens, Quan Shen, and Jocelyn Cortez. 2022. Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nature Food* 3, 1 (2022), 19–28.
- [51] Yunchen Pu, Zhe Gan, Ricardo Henao, Xin Yuan, Chunyuan Li, Andrew Stevens, and Lawrence Carin. 2016. Variational autoencoder for deep learning of images, labels and captions. *Advances in neural information processing systems* 29 (2016).
- [52] Dmitry Rashkovetsky, Florian Mauracher, Martin Langer, and Michael Schmitt. 2021. Wildfire detection from multisensor satellite imagery using deep semantic segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), 7001–7016.
- [53] Ali Razavi, Aaron Van den Oord, and Oriol Vinyals. 2019. Generating diverse high-fidelity images with vqvae-2. *Advances in neural information processing systems* 32 (2019).
- [54] Carlos Riquelme, Joan Puigcerver, Basil Mustafa, Maxim Neumann, Rodolphe Jenatton, André Susano Pinto, Daniel Keysers, and Neil Houlsby. 2021. Scaling vision with sparse mixture of experts. *Advances in Neural Information Processing Systems* 34 (2021), 8583–8595.
- [55] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III* 18. Springer, 234–241.
- [56] Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J Fleet, and Mohammad Norouzi. 2022. Image super-resolution via iterative refinement. *IEEE transactions on pattern analysis and machine intelligence* 45, 4 (2022), 4713–4726.
- [57] Khalid Sayood. 2002. *Lossless compression handbook*. Elsevier.
- [58] Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. 2017. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. *arXiv preprint arXiv:1701.06538* (2017).
- [59] Climate Data Store. 2024. Cloud properties global gridded monthly and daily data from 1979 to present derived from satellite observations. <https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-cloud-properties?tab=overview>. (2024).
- [60] Chuanhao Sun, Kai Xu, Gianni Antichi, and Mahesh K. Marina. 2024. NetGSR: Towards Efficient and Reliable Network Monitoring with Generative Super Resolution. *Proc. ACM Netw.* 2, CoNEXT4, Article 29 (Nov. 2024), 27 pages. <https://doi.org/10.1145/3696400>
- [61] Bill Tao, Om Chabra, Ishani Janveja, Indranil Gupta, and Deepak Vasisht. 2024. Known Knowns and Unknowns: Near-realtime Earth Observation Via Query Bifurcation in Serval. In *21st USENIX Symposium on Networked Systems Design and Implementation (NSDI 24)*. 809–824.
- [62] Bill Tao, Maleeha Masood, Indranil Gupta, and Deepak Vasisht. 2023. Transmitting fast and slow: Scheduling satellite traffic through space and time. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*. 1–15.
- [63] Aysim Toker, Lukas Kondmann, Mark Weber, Marvin Eisenberger, Andrés Camero, Jingliang Hu, Ariadna Pregel Hoderlein, Çağlar Şenaras, Timothy Davis, Daniel Cremers, et al. 2022. Dynamicearthnet: Daily multi-spectral satellite dataset for semantic change segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 21158–21167.
- [64] Huiwen Wang, Liang Liao, Jing Xiao, Weisi Lin, and Mi Wang. 2023. Uplink-assist downlink remote-sensing image compression via historical referencing. *IEEE Transactions on Geoscience and Remote Sensing* 61 (2023), 1–15.
- [65] Jian Wang, Seokbeop Kwon, and Byonghyo Shim. 2012. Generalized orthogonal matching pursuit. *IEEE Transactions on signal processing* 60, 12 (2012), 6202–6216.
- [66] Jingdong Wang, Ting Zhang, Nicu Sebe, Heng Tao Shen, et al. 2017. A survey on learning to hash. *IEEE transactions on pattern analysis and machine intelligence* 40, 4 (2017), 769–790.
- [67] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* 13, 4 (2004), 600–612.
- [68] Jiangkai Wu, Yu Guan, Qi Mao, Yong Cui, Zongming Guo, and Xinggong Zhang. 2023. ZGaming: Zero-latency 3D cloud gaming by image prediction. In *Proceedings of the ACM SIGCOMM 2023 Conference*. 710–723.
- [69] Jun Yang, Peng Gong, Rong Fu, Minghua Zhang, Jingming Chen, Shunlin Liang, Bing Xu, Jiancheng Shi, and Robert Dickinson. 2013. The role of satellite remote sensing in climate change studies. *Nature climate change* 3, 10 (2013), 875–883.
- [70] Jingkang Yang, Kaiyang Zhou, Yixuan Li, and Ziwei Liu. 2024. Generalized out-of-distribution detection: A survey. *International Journal of Computer Vision* (2024), 1–28.
- [71] Yibo Yang, Robert Bamler, and Stephan Mandt. 2020. Improving inference for neural image compression. *Advances in Neural Information Processing Systems* 33 (2020), 573–584.

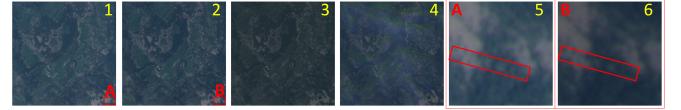


Figure 16: Common reason of inconsistent observation in short period (Planet dataset). 1: high brightness, 2: medium brightness, 3: low brightness, 4: RGB channel mismatch, 5: Zoom-in A region, with artifact, 6: Zoom-in B region, without artifact

- [72] Hyunho Yeo, Hwijoon Lim, Jaehong Kim, Youngmok Jung, Juncheol Ye, and Dongsu Han. 2022. Neuroscaler: Neural video enhancement at scale. In *Proceedings of the ACM SIGCOMM 2022 Conference*. 795–811.
- [73] Seniya Esen Yüksel, Joseph N Wilson, and Paul D Gader. 2012. Twenty years of mixture of experts. *IEEE transactions on neural networks and learning systems* 23, 8 (2012), 1177–1193.
- [74] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. 2018. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 586–595.
- [75] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. 2018. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)*. 286–301.
- [76] Fangyuan Zhu. 2022. A Review of Deep Learning Based Image Super-resolution Techniques. (2022). [arXiv:cs.CV/2201.10521](https://arxiv.org/abs/2201.10521) <https://arxiv.org/abs/2201.10521>

A Appendix

Appendices are supporting material that has not been peer-reviewed.

A.1 Inherent Distortion and Cause of Changes

As Figure 19 shows, even without cloud coverage and known event on the ground, the images of the same location in a short period is not exact the same. Before compression, it is critical to analysis the root causes of such difference, as not all variation (e.g., *cause by noise*) need to be perfectly recovered.

We analyze both open-source and private earth observation datasets, and summarize the following common causes of difference in continuous observations, except the weather or special events. The common causes of change on observation in short period is illustrated in Figure 16, where the pictures are taken by different satellite at very close timestamp (within 2 days). We can assume there is no significant change on the ground in such a short period, and they are suppose to have very high similarity — but this is not the case. First, there is unpredictable artifact when satellite produce the final images, for example, between 1 and 2 in Figure 16, besides marginal brightness difference, some of the details are not aligned. Zoom in to region A and B (correspond to 5 and 6 in Figure 16), in region A, there is artifact on the image, leading to mismatching effect on nearby pixels (A and B is exactly the same region), and the SSIM between A and B is 0.9. Second, brightness varies a lot on the same region, as 3 in Figure 16 shows, the SSIM between 2 and 3 is just 0.83. Third, the satellite takes different channels independently and then merge them together. Sometime the satellite is not very stable when shooting different channels, causing mismatching between channels. The 4 in Figure 16 has “rainbow” like pattern on the image, which is caused by mismatched RGB channels.

Image Type	Resolution	# Channels	Code Size	# Code (# Land only)	Total Size (GB)
Planet RGB (3m)	3m	3	16 Bytes	8.52×10^7 (2.56×10^8)	13.6 (4.1)
Planet RGB (4m)	4m	3	16 Bytes	4.94×10^8 (1.48×10^8)	7.9 (2.4)
Sentinel-2	10m	12	64 Bytes	0.75×10^8 (0.22×10^8)	4.8 (1.4)

Table 15: Number of Codes and Storage Requirement.

Dataset	Training Set	Input Granularity	*GSD	Training Ratio	Supported CR	#Ref.
Planet-CAL [20]	First 3 days	5 ~ 6 hours	3.5 m	15%	[16,32,64,128,256]	6
Planet-HK [20]	First 3 days	5 ~ 6 hours	3.5 m	15%	[16,32,64,128,256]	6
Farm Vibes [1]	First 3 Months	Monthly	10m	50%	[16,32,64,128,256,512]	3
DEN-3 [63]	First Year	Weekly	3.5 m	7%	[16,32,64,128,256,512]	10
DEN-12 [63]	First Year	Monthly	10m	50%	[16,32,64,128,256,512]	10

Table 16: Training set and compression configuration. *GSD: Ground sampling distance.

Among all the causes in Figure 16, besides the brightness, the other artifacts are hard to avoid or fix as they appear in an unpredictable manner. This phenomenon means even the raw observation has a certain level random distortion. In theory, no method can address the inherent uncertainty of the data. Therefore, based on available observation, we come up with a target fidelity at $\text{SSIM} > 0.9$, which is around the similarity we can expect when using satellite observe same region in a short period.

Takeaway: Earth observation data have inherent random distortion. The average level distortion can be taken as a fidelity tolerance when carry out reconstruction as the data uncertainty is irreducible.

A.2 Implementation Aspects

A.2.1 Onboard Compression Setup. By default, onboard computations, including interpolation and metrics such as SSIM and Hamming distance, are GPU-accelerated. To align with the latest onboard environment [18, 61], we utilize the Jetson AGX Orin GPU in a 15W power configuration as our reference compute platform on satellite.

As in Figure 4a, each image is divided into non-overlap 256×256 tiles, and each tile is first sampled to 64×64 , 45×45 , 32×32 , 23×23 and 16×16 (*i.e.*, $\text{CR}=16, \dots, 256$), then reconstruct to 256×256 using naive interpolation. The basic CR per image is selected as the one meets fidelity requirement with highest CR, following the $\phi(\cdot)$ defined in Equation 2, and the rules in §4.1. As for the computation of $\eta(\cdot)$, by default we use a 16 bytes to represent a 256×256 tile with 3 channels in BLSH – 1.23×10^4 times smaller than the original size, enabling storage of multiple variants of reference codes for images across the globe. The codes, *i.e.*, hashed reference images, are updated weekly for the same period of same region, the typical uplink consumption for code updating is $< 10\text{ kbps}$ following the code size information in Table 15, which is just 4% of the typical uplink bandwidth of planet LEO satellite [34]. The configuration for different datasets is shown in Table 16. The threshold of $\eta(\cdot)$ (as defined in §4.1, Equation 3) is set to 0.9, *i.e.*, if reference yields $\eta(\cdot) > 0.9$, we double the CR ($\alpha = 2$ in §4.1).

A.2.2 Decompression Setup. We display the configuration of the MoE system in Table 17. The coverage area for each expert is constrained by the dataset. For instance, in the DynamicEarthNet (DEN) dataset, the default resolution is 1024×1024 pixels per region, which we adopt as the expert’s default coverage since they do not need to generalize across different regions. The threshold of $\phi(\cdot)$ defines the

Dataset	#Region	Region Size (Pixel)	$\phi(\cdot)$	Threshold	Cloud Levels
Planet-CAL [20]	16	4096×4096	$\phi(\cdot) > 0.85$	[0,1]	
Planet-HK [20]	16	4096×4096	$\phi(\cdot) > 0.85$	[0,0.5,1]	
Farm Vibes [1]	1	1024×1024	$\phi(\cdot) > 0.85$	[0,1]	
DEN-3 [63]	96	1024×1024	$\phi(\cdot) > 0.85$	N/A	
DEN-12 [63]	36	1024×1024	$\phi(\cdot) > 0.85$	[0,1]	

Table 17: Configuration of MoE system (tile size 256×256).

inherent redundancy of image following Equation 4. For instance, the tile that has $\text{SSIM} > 0.85$ to the original image after applying $64\times$ sampling and interpolation but not after $128\times$, will be classified to the same group during both training and testing phase. We train and evaluate images at varying levels of cloud thickness. For two levels, we divide by 50% visibility. As an exception, in the Planet-HK dataset, there are three levels: Level 0 (visibility $> 67\%$), Level 0.5 ($33\% \leq \text{visibility} \leq 67\%$), and Level 1 ($\text{visibility} < 33\%$) – Hong Kong experiences more cloud levels due to its proximity to the ocean and numerous islands, resulting in more complex cloud coverage.

The trained checkpoint of trained expert is only 168MB, less than the size of a single satellite image. Thresholds for both LRS and HSIM to trigger retransmission are set to 0.95. Temporary onboard storage is triggered if $\phi(\cdot) < 0.85$ and $CR \leq 64$. Those configurations are empirical value based on the datasets we have access to. The retransmission rarely happens and negligible in terms of communication overhead.

A.2.3 Continual Training Data Collection. During training data collection, DEEPSPACE aims to save space by ensuring the training dataset collection is highly sparse and tolerant of high latency. By default, DEEPSPACE only needs one “raw” image weekly in total of a target region for model training. Since there are multiple satellites covering each region, each satellite only needs to send $1/N_S$ image weekly for training where N_S is the number of times satellites pass by that region each week. The average extra downlink traffic is $\sim \frac{1}{N_S} \text{ kbps}$ for a 500km^2 region, negligible to general data transmission. Those training data will be transmitted with the spare bandwidth as there will be significant spare bandwidth when using DEEPSPACE (256x compression, see §6).

Considering the fact that some datasets do not provide satellite identification [63], we do not limit the number of satellite engaged in this work. By default for both training and testing phases, we will leverage all images that potentially come from different satellites. This is also the fairest setup for methods that require information from multiple satellites. For instance, Earth+ [24] leverages the image from the other satellites to detect near replicated images at a different satellite.

A.3 Sources of Redundancy

The satellite images can be compressed according to the redundancy in two aspects:

(1) Single Image Redundancy: some part of image has very simple pattern, and can be recovered even after massively sampled. For instance the cloud in Figure 17 and forest in Figure 18. In Figure 17 (b) and Figure 18 (b), we first sample the image by 64 times smaller (*i.e.*, 8 times evenly along X and Y axis), then recover the image to original resolution by Lanczos interpolation [36, 39]. Compared with the original image, on the spectral domain, only the significant components are reserved after interpolation, which is distributed around

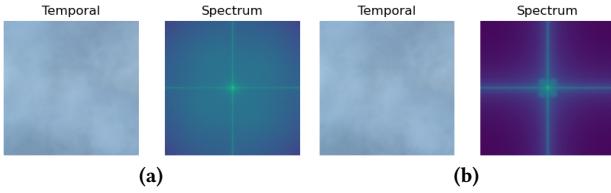


Figure 17: (a) Original Thick Cloudy Image and Spectrum. (b) Sampled by 64 times, recovered by Lanczos interpolation. After 64 times sampling, the interpolated image has 0.99 SSIM to the original image.

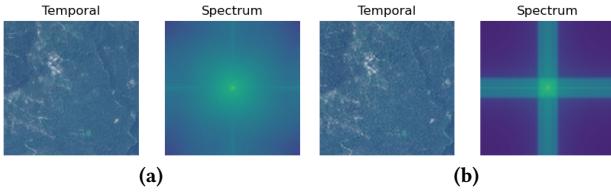


Figure 18: (a) Original Forest Image and Spectrum in California, July 2021. (b) Sampled by 64 times, recovered by Lanczos interpolation. After 64 times sampling, the interpolated image has 0.98 SSIM to the original image.

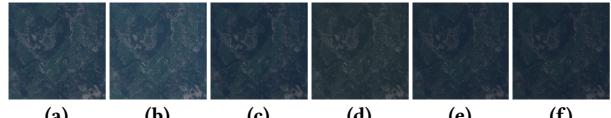


Figure 19: (a) to (f) Image without cloud coverage of the same place from 10th July to 15th July.

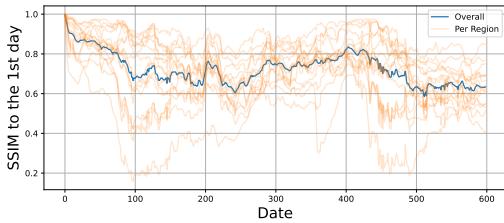


Figure 20: SSIM to the first date in the DynamicEarthNet dataset of one large region (10km^2) and each sub-regions (0.625km^2)

the low frequency part. With those low frequency components, we recover an image with > 0.98 SSIM (Structural SIMilarity) [67] to the original image, indicating that the high frequency part only has negligible power. Similarly, we can apply same analysis to the coverage with huge water body, desert, and etc.

(2) Near Replicate Images: the geographical structure of the earth evolve much slower than the frequency of observation, leading near replicate images (or part of image) in continuous observations. In Figure 19, where we analyze the similarity between the continuous observations in a 5-day period. Even with different brightness, their mutual SSIM is between 0.83 to 0.98, and average mutual SSIM is 0.91, when there is no significant cloud coverage. Those phenomena indicate that even for the images with complex pattern (low redundancy), if we can only transmit the “difference” to few known reference images, the data amount to be sent can be reduced significantly.



Figure 21: Ground truth image.

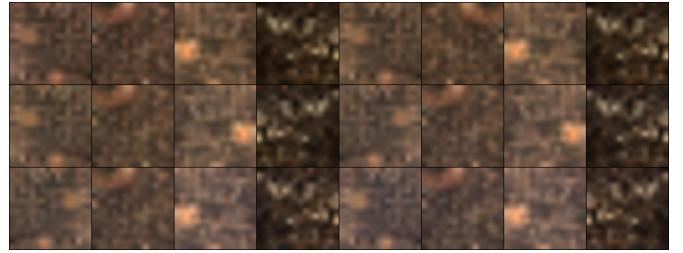


Figure 22: Reconstruction by interpolation after 256x compression.



Figure 23: Reconstruction with DEEPSpace after 256x compression.

Leveraging those redundancy for compression is non-trivial. From Figure 20, we see there is seasonal variation for the overall similarity in a large region. However, there can be > 0.15 overall SSIM change in less than one week time. Figure 20 also highlights the importance of operating on small region with an adaptive manner, as the neighbouring regions might have distinct dynamics in temporal, some region changes more significant and fast than others.

Takeaway: Earth observation image data have high redundancy. However, leveraging this for compression is non-trivial due to the spatial diversity and temporal dynamic.

A.4 Detailed Results

A.4.1 Visualizing Reconstruction at City Scale. We include the visualization on city scale in Figure 21 for the original image, Figure 22 for low resolution image recovered by interpolation, and Figure 23 for the reconstruction with DEEPSpace. The image of the same place might be taken from different time or angles, where DEEPSpace consistently reconstructs image with high fidelity. The results of interpolation in Figure 22 illustrates that the reconstruction is non-trivial, the details cannot be recovered simply through interpolation.

Method	Dataset	SSIM (>Min.)↑	PSNR (>Min.)↑	CR↑
DEEPSpace (SSIM>0.85)	DEN-3	0.97 (0.93)	40.3 (37.2)	320
	DEN-12	0.98 (0.93)	40.8 (37.8)	256
	Planet-CAL	0.95 (0.9)	39.1 (35.8)	146.4
	Planet-HK	0.96 (0.92)	40.3 (37.1)	186.2
	Farm-Vibe	0.95 (0.92)	39.3 (37.3)	256
Earth+ (Ideal ¹)	DEN-3	0.94 (0.88)	38.1 (34.8)	11
	DEN-12	0.94 (0.89)	38.6 (35.3)	8.5
	Planet-CAL	0.92 (0.87)	36.7 (32.3)	6.2
	Planet-HK	0.90 (0.83)	34.6 (30.1)	8.4
	Farm-Vibe	0.91 (0.83)	34.9 (30.5)	3.2
Kodan (Ideal, All APP.)	DEN-3	0.97 (0.91)	40.2 (36.3)	2.1
	DEN-12	0.97 (0.92)	40.3 (37.4)	1.8
	Planet-CAL	0.95 (0.89)	38.8 (35.7)	2
	Planet-HK	0.94 (0.87)	39.2 (34.9)	2.5
	Farm-Vibe	0.95 (0.84)	39.0 (32.8)	1.7

Table 18: Comparison between DEEPSpace and orbital edge computing (OEC) schemes. ¹:in Earth+ and Kodan, we assume ideal cloud detection, fine granularity of tiling, leverage all satellites in dataset.

Note that the images on the same column covers approximately the same region from different satellite at different time, and therefore they are rotated with difference in details.

A.4.2 Comparison with Orbital Edge Computing (OEC) Schemes. We present the comparison between DEEPSpace and OEC image filtering in terms of overall fidelity and bandwidth/storage savings in Table 18.