

SCo-Cloud: Satellite Constellation Collaboration for Cloud-Aware Onboard-Computed Imaging and Transmission

Jia Liu¹, Qian Li¹, Yongqi Li¹, Cheng Ji², Shangguang Wang¹

¹Beijing University of Posts and Telecommunications

²Beihang University

Abstract

Satellite-acquired optical remote sensing imagery is extensively applied in time-critical applications like traffic surveillance and evaluation of natural disasters. However, *clouds*, as a common atmospheric phenomenon, frequently obscure observation. Current approaches aim to restore visibility in cloud-obscured regions, yet they typically fall short in the presence of dense cloud cover, which are exceedingly prevalent in remote sensing imagery. Alternative approaches rely on the satellite revisit cycle, frequently surpassing ten days, a duration impractical for genuine application scenarios due to target changes and bandwidth limitations. To address these issues, this paper proposes **SCo-Cloud**, a novel satellite constellation collaboration framework for cloud-aware onboard-computed imaging and transmission, which consists of Center-Sat and Edge-Sats. We propose onboard thin cloud removal and re-imaging region location models to locate the impact of clouds. We further design a novel multi-satellite scheduling strategy to eliminate clouds. The models above are integrated within the Center-Sat, with the nearby Edge-Sats collaborating in tandem to execute re-imaging assignments. Furthermore, to facilitate in-depth research, we have meticulously developed a cloud-covered target detection dataset. Comprehensive experiments have conclusively demonstrated that SCo-Cloud effectively surpasses the limitations inherent in current approaches, providing accurate and timely responses within the domain of Earth observation.

Code — <https://github.com/liuzscx/SCo-Cloud>

Introduction

A large number of remote sensing images collected by Earth observation satellites are widely used in scenarios such as traffic monitoring, emergency response, and object detection (Xie et al. 2023; Zhang et al. 2022; Cheng and Lucia 2025; Cheng et al. 2024; Zhou et al. 2024; Liu et al. 2024). For tasks with high real-time requirements, it is crucial to obtain complete surface information of the target area within a short period (Zheng et al. 2023; Luo et al. 2025). However, prior work (Zhang et al. 2024) shows that **observations of satellites are obscured by clouds** (approximately 67%), which severely hinders satellite observation of target regions.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

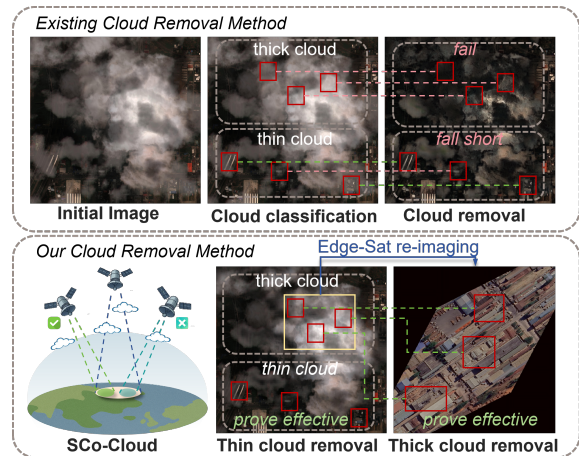


Figure 1: The comparison between our method and existing approaches. SCo-Cloud offers more efficient thin-cloud processing capabilities, while also featuring a constellation scheduling strategy to handle scenarios with heavy cloud cover.

Existing methods struggle to handle the cloud-covered remote sensing images. Some methods focus on single optical images obscured by clouds, using transformer-based models to extract cloud-covered region features in an attempt to restore cloud-free states (Kulkarni and Murala 2023; Guo et al. 2022; Song et al. 2023b,a). Other methods leverage diffusion models, reconstructing the occluded areas through generative model structures (Sui et al. 2024; Wang et al. 2024). However, these methods only work when the clouds are thin and visibility is high. In natural environments, the presence of large amounts of thick clouds renders the above methods ineffective. Other methods leverage multi-temporal imagery, where satellites repeatedly capture the same region at different times (Wang et al. 2023; Ebel et al. 2023, 2022; Long et al. 2023; Sui et al. 2024; Yuan et al. 2024; Wu et al. 2025). All captured images are transmitted to ground stations to reconstruct cloud-free imagery. However, Low Earth Orbit (LEO) satellites typically complete an orbit every 90 minutes, the revisit period for the same region usually extends to several weeks (Hao et al. 2024; Kuang, Xiang, and Guo 2025). During this interval, the target area may un-

dergo significant changes or even disappear. Methods such as (Denby et al. 2023; Du et al. 2025; Tao et al. 2024) reduce the revisit interval for the same region using **satellite constellations** and can capture a cloud-free image every ten days. However, a ten-day gap remains too long for urgent scenarios such as disaster response. Moreover, these methods label cloud-covered regions as low-value and discard them directly. Consequently, no compensation is made for the missing surface information beneath the clouds. Therefore, in this paper, we seek to answer the question: *How to provide timely and effective imaging of cloud-covered target regions with bandwidth-saving transmission?*

We propose a satellite constellation architecture, **SCo-Cloud**, to address the challenges of cloud removal for timeliness and downlink bandwidth constraints, as shown in Figure 1. Specifically, the SCo-Cloud constellation comprises a Center-Sat with onboard computing capabilities and multiple collaboratively operating Edge-Sats. The Center-Sat performs onboard cloud detection on its acquired remote sensing imagery and adopts differentiated strategies based on cloud thickness. For **thin cloud** regions, where surface features are faintly visible, the Center-Sat directly conducts onboard thin cloud removal. In contrast, **thick cloud** regions, where the surface is completely obscured, require re-imaging to obtain valid observations. To this end, the Center-Sat determines the re-imaging areas based on the distribution of thick clouds and schedules eligible Edge-Sats to perform the imaging tasks. To ensure scheduling accuracy, we propose a hierarchical constraint-based scheduling algorithm. The Center-Sat first preliminarily filters satellites with feasible observation conditions based on the geometric relationship between the cloud distribution and the orbital positions of the Edge-Sats. It then refines the candidate set by considering imaging angle constraints to finalize task assignment. All critical operations are executed onboard, thereby enhancing the completeness of target detection in designated regions. Given the bandwidth constraints of satellite downlink resources, our constellation architecture introduces a content-aware selective transmission mechanism to avoid redundant data. In summary, the main contributions are as follows:

- **SCo-Cloud Constellation Architecture.** We design a satellite constellation system with on-orbit computing capability, effectively addressing target loss caused by cloud occlusion in remote sensing images.
- **Hierarchical Constraint Scheduling.** We propose a hierarchical constraint scheduling algorithm for autonomous onboard task planning. The algorithm achieves efficient dynamic task assignment and scheduling decisions onboard the Center-Sat.
- **Content-Aware Selective Transmission.** We propose a content-aware selective transmission scheme to address the limitation of satellite downlink resources and improve bandwidth utilization.

Related Work

Cloud Removal in Remote Sensing Images. We have explored various approaches to address the issue of cloud oc-

clusion in satellite imagery. Methods such as (Ebel et al. 2023; Wang et al. 2016; Ebel et al. 2022; Long et al. 2023) adopt multi-temporal strategies to reconstruct occluded information using satellite images captured at different times. However, these approaches rely on the satellite’s revisit cycle, which typically spans several days or even weeks, leading to the risk of missing targets or encountering significant scene changes. Some works that use generative methods (Tao et al. 2022; Xu et al. 2022; Zi et al. 2021), such as diffusion models, have achieved good results to a large extent. However, when natural disasters occur or there are significant changes in the observed area, the performance of generative methods becomes inaccurate. Works (Xiang, Tan, and Yan 2024; Li et al. 2024; Pan et al. 2024; Ghildiyal et al. 2024; Li, Liu, and Li 2023) leverage multimodal fusion by incorporating SAR or infrared imagery to recover information from cloud-covered regions in optical images. However, such methods rely on the availability of large volumes of multimodal data. Under bandwidth constraints, it is difficult to transmit back. Recent methods such as (Du et al. 2025; Denby et al. 2023; Tao et al. 2024) employ onboard processing, utilizing satellite-based cloud detectors, onboard image recognition, or weather forecasts to identify and filter out cloud-covered images before downlink. However, existing approaches have yet to address the problem of recovering targets obscured beneath clouds.

Downlink Transmission. Satellite-to-ground communication resources are inherently limited. Each LEO satellite typically experiences only 4-10 communication windows per day, each lasting approximately 5–15 minutes, and the bandwidth does not exceed several hundred Mbps (Furutani et al. 2025; Gómez and Meoni 2024; Zhang et al. 2024). Our approach selectively transmits only cloud-free, valid images, thereby significantly improving downlink utilization.

Methodology

SCo-Cloud is a specialized framework designed to tackle the challenges of cloud-covered remote sensing images that hinder ground target observation. This framework integrates multiple image processing models and scheduling algorithms into a complex pipeline, which is deployed onboard satellites to perform a series of processing tasks through on-orbit computation, as illustrated in Figure 2. The framework consists of three components: **Thin Cloud Filter** restores ground features in thin-cloud regions of the initially captured image, providing optimized information for re-imaging region selection; **Re-imaging Region Locator** identifies suitable regions for re-capture based on the optimized images; **Edge-Sat Selector** selects the optimal satellite to perform the imaging task through precise geo-spatial modeling and scheduling algorithms.

Problem Definition. To formally describe the onboard-computed re-image and transmission, let I be the image captured by the Center-Sat, with M target regions that are covered by clouds and require re-imaging. Each region is represented by a quadrilateral defined by four geographic vertices, and the set of all such target regions is denoted as

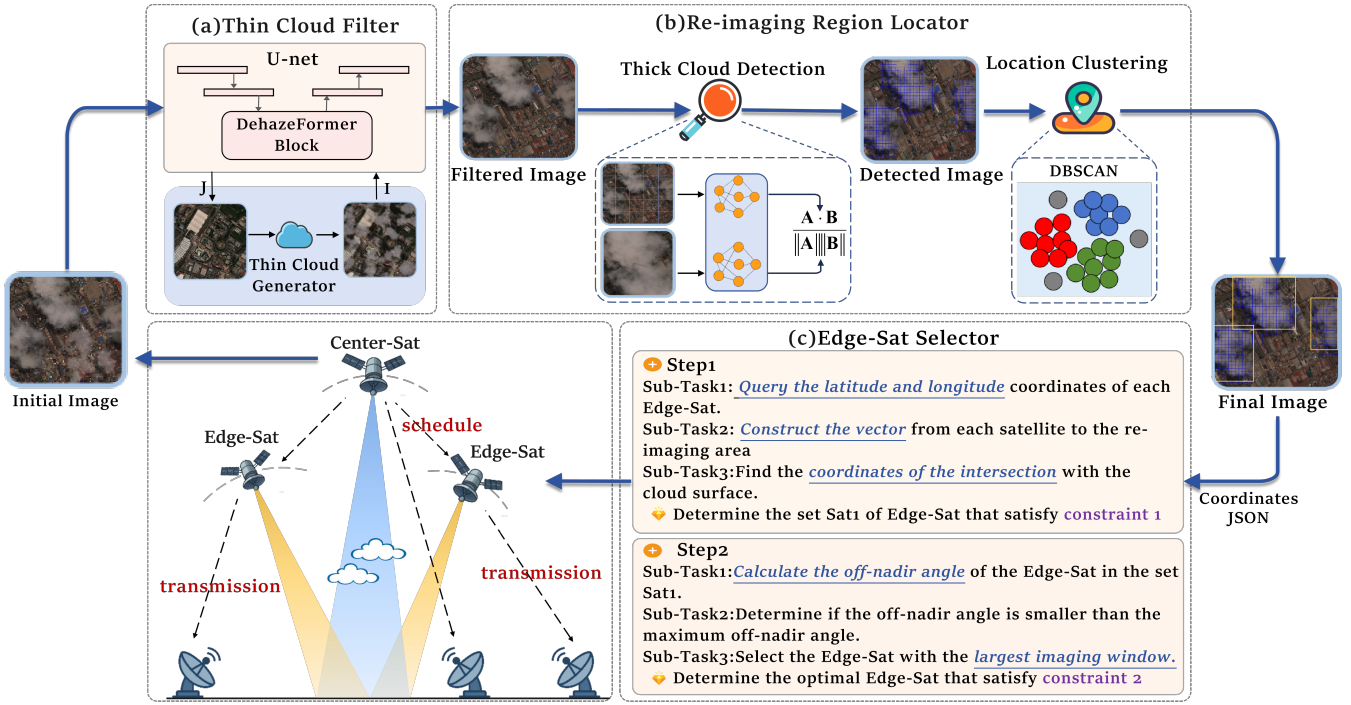


Figure 2: SCo-Cloud system model. The Center-Sat performs three modules of on-orbit processing on the captured images: (a) Thin Cloud Filter, (b) Re-imaging Region localization Module, and (c) Edge-Sat Selector.

$\mathcal{R} = \{(p_{i1}, p_{i2}, p_{i3}, p_{i4})\}_{i=1}^M$, where $p = (\phi, \lambda)$ and ϕ, λ denoting the latitude and longitude of vertex p . Let the positions of N Edge-Sats be $\mathcal{S} = \{(\phi_j, \lambda_j)\}_{j=1}^N$, where each element denotes the geographic location of the j -th Edge-Sat. Given the above, a scheduling algorithm is applied to assign Edge-Sats to regions, generating a scheduling table:

$$\mathcal{T} = \{(r_i, s_j) \mid r_i \in \mathcal{R}, s_j \in \mathcal{S}\}, \quad (1)$$

where (r_i, s_j) indicates that the i -th re-imaging region is assigned to the j -th Edge-Sat. During the downlink phase, Center-Sat transmits the cloud-free portion of the captured image I , while Edge-Sats downlink the re-captured images corresponding to their assigned regions. Formally, the total downlinked data D is composed of two parts, $D = I_{\text{clear}} \cup \bigcup_{(r_i, s_j) \in \mathcal{T}} I_{r_i}^{(s_j)}$, where I_{clear} denotes the cloud-free part of image I transmitted by the Center-Sat, and $I_{r_i}^{(s_j)}$ represents the image of region r_i re-imaged by Edge-Sat s_j .

Thin Cloud Filter

The Thin Cloud Filter addresses the inefficiency in Edge-Sat resource utilization by employing a thin cloud removal model. Specifically, Center-Sat is responsible for accurately identifying cloud distribution in each satellite image. However, not all cloud-covered regions fully obscure surface information. In areas affected by thin cloud cover, ground features remain partially visible and can be effectively recovered using computational algorithms. Consequently, the deployment of Edge-Sats for re-imaging in such scenarios is unnecessary. Reacquiring imagery for all cloud-covered re-

gions using Edge-Sats would lead to substantial and avoidable consumption of satellite resources.

Accordingly, the Thin Cloud Filter reconstructs regions affected by thin cloud cover to approximate a cloud-free condition. While existing studies have focused on thin cloud removal in imagery (Kulkarni and Murala 2023; Song et al. 2023a; Guo et al. 2022; Song et al. 2023b), these methods have proven ineffective in the proposed remote sensing scenarios. To improve model generalization under such conditions, we employ a cloud generator (Czerkawski et al. 2023) to synthesize clouds on cloud-free images for model fine-tuning. The loss function is defined as follows:

$$\mathcal{L}_{L_1} = \frac{1}{N} \sum_{i=1}^N \left| \hat{J}(x_i) - J(x_i) \right|, \quad (2)$$

where $J(x_i)$ represents the cloud-free image, $\hat{J}(x_i)$ is the synthesized cloudy image generated by the cloud generator, and N denotes the number of samples. To generate $\hat{J}(x_i)$, the cloud generator \mathcal{G} first constructs a cloud mask and applies it to the input image via the following formulation.

$$\hat{J}(x_i) = \mathcal{G}(J(x_i)) = J(x) \cdot (1 - M(x)) + C(x) \cdot M(x), \quad (3)$$

where $M(x)$ represents the cloud mask image, $C(x)$ denotes the cloud texture map and $J(x)$ refers to the original image. This blending process allows clouds to be realistically added to designated regions of the image. The construction of a thin-cloud removal dataset based on the xView dataset provides high-quality training samples for cloud removal research and effectively simulates real cloud-contaminated images captured by Center-Sat.

Re-imaging Region Locator

The Re-imaging Region Locator localizes and aggregates the identified thick cloud regions, thereby providing a critical foundation for subsequent imaging by Edge-Sats. To address this challenge, we propose a multi-step strategy that effectively mitigates the impact of cloud occlusion. Through the coordinated design of each step, the proposed approach significantly improves the handling of cloud location.

Thick Cloud Detection. We utilize DINOv2 (Oquab et al. 2023), a visual model pretrained on large-scale image data, to perform initial localization of cloud regions. In this strategy, the satellite image I captured by Center-Sat is divided into multiple patches $\{p_i\}_{i=1}^N$, each of which is compared with the feature vector f_{cloud} of a thick cloud image. For each patch p_i , we compute a similarity score s_i as $s_i = \text{sim}(f, f_{cloud})$, where f is the feature vector of patch p_i , and $\text{sim}(\cdot, \cdot)$ denotes a similarity function. Patches with similarity above a predefined threshold τ are retained in $\mathcal{P} = \{p_i | s_i \geq \tau\}$.

Location Clustering. The objective of the Location Clustering is to provide the geographic coordinates (latitude and longitude) for each re-imaging region. Considering both the imaging constraints of satellites and the need for coordinate-based scheduling of Edge-Sats, the re-imaging areas are defined in a rectangular shape. The Thick Cloud Detector tends to scatter cloud regions across multiple patches; treating each patch as an independent re-imaging unit is inefficient. To address this, we adopt algorithm groups patches based on density. Specifically, a patch p_i in the set of patches \mathcal{P} , where $p_i \in \mathcal{P}$, is assigned to a cluster if it has at least MinPts neighbors within radius ε .

$$C_k = \{p_i \in \mathcal{P} \mid |\{p_j \in \mathcal{P} \mid d(p_i, p_j) \leq \varepsilon\}| \geq \text{MinPts}\}, \quad (4)$$

where $d(p_i, p_j)$ is the distance between patch centers. By applying a clustering algorithm to spatially adjacent cloud patches into coherent regions. Subsequently, the vertex coordinates of each aggregated region are computed and stored in a JSON file.

Edge-Sat Selector

Edge-Sat Selector, the final module before Center-Sat generates the scheduling table, is responsible for selecting the optimal Edge-Sat to perform the re-imaging task. Considering limitations such as cloud interference and imaging distortion, we propose a two-step *Hierarchical Constrained Scheduling Algorithm* within this module. The specific constraints are defined as follows.

Constraint 1. No Cloud Occlusion in the Field of View.

The primary prerequisite for Edge-Sat to successfully perform an imaging task is the absence of cloud occlusion over the target area within its field of view. To enable accurate geometric computation, it is necessary to convert the geographic coordinate (Latitude, Longitude, Altitude) into the Earth-Centered Earth-Fixed (ECEF) coordinate system.

$$\begin{aligned} \mathbf{l} &= (x, y, z) \\ &= (r \cdot \cos(\phi) \cdot \cos(\lambda), r \cdot \cos(\phi) \cdot \sin(\lambda), r \cdot \sin(\phi)), \end{aligned} \quad (5)$$

where \mathbf{l} represents the vector in the ECEF coordinate system derived from latitude and longitude, ϕ and λ represent latitude and longitude, respectively. $r = R_{\text{earth}} + \text{alt}$, where R_{earth} represents the Earth's radius and alt denotes the altitude relative to the Earth's surface. In the ECEF coordinate system, the position vector of an Edge-Sat is denoted as \mathbf{l}_s , and the position vectors of each vertex of the re-imaging region are denoted as \mathbf{l}_t , then $\mathbf{l}_{dir} = \mathbf{l}_t - \mathbf{l}_s$. By constructing the vector representation $\mathbf{p}(t)$ of a random point on \mathbf{l}_{dir} , we determine whether the line passes through the cloud layer.

$$\mathbf{p}(t) = \mathbf{l}_s + t \cdot (\mathbf{l}_t - \mathbf{l}_s) = \mathbf{l}_s + t \cdot \mathbf{l}_{dir}, \quad t \in [0, 1]. \quad (6)$$

When \mathbf{l}_{dir} intersects the cloud sphere, we have

$$\|\mathbf{p}(t)\|^2 = R_{cloud}^2, \quad (7)$$

where R_{cloud}^2 represents the cloud layer radius.

Proposition 1 (Cloud Sphere Intersection) *The imaging ray from the Edge-Sat intersects the spherical cloud layer when $\sin^2 \theta \leq \frac{R_{cloud}^2}{\|\mathbf{l}_s\|^2}$, where $\mathbf{l}_s \in \mathbb{R}^3$ denotes the position vector of the Edge-Sat, \mathbf{l}_{dir} is the unit direction vector from the satellite to the target region, θ is the angle between \mathbf{l}_s and \mathbf{l}_{dir} , R_{cloud} is the radius of the spherical cloud layer. $\frac{R_{cloud}^2}{\|\mathbf{l}_s\|^2}$ is a geometric threshold. This inequality indicates that the imaging ray can only intersect the cloud sphere when the angle θ between \mathbf{l}_{dir} and \mathbf{l}_s is sufficiently small. Proof of Proposition 1 can be found in the Appendix.*

In proposition 1, $\sin^2 \theta \leq \frac{R_{cloud}^2}{\|\mathbf{l}_s\|^2}$ means that the Edge-Sat and the target vertex are on opposite sides of the cloud. In this case, the physical intersection point between the imaging ray and the cloud surface is computed by solving for $\mathbf{l}_p = \mathbf{l}_s + t \cdot \mathbf{l}_{dir}$, $t = \min(t_1, t_2)$, where t_1 and t_2 are the two real roots obtained from solving the corresponding quadratic equation 7. The intersection vector $\mathbf{l}_p = (x, y, z)$ is expressed in the ECEF coordinate system and is subsequently transformed into geographic coordinates (ϕ_p, λ_p) , where ϕ_p and λ_p represent the latitude and longitude, the calculation formula is provided in the Appendix.

The binary cloud mask function $f_{c\text{-mask}}(\phi, \lambda)$ is then used to determine whether (ϕ_p, λ_p) lies within the cloud-covered region. If the intersection point falls outside the cloud-covered area, i.e., $f_{c\text{-mask}}(\phi_p, \lambda_p) = 0$, the current Edge-Sat position \mathbf{l}_s is designated as the Imaging Start Point.

Constraint 2. Satisfying Maximum Off-nadir Angle.

It is essential to account for Edge-Sat's onboard optical constraints. Exceeding the off-nadir angle limitation introduces severe geometric distortion, reducing image usability. We denote the maximum off-nadir angle as θ_{max} . The off-nadir angle between the Edge-Sat and a ground target point $\theta_{\text{off-nadir}}$ is computed as:

$$\theta_{\text{off-nadir}} = \arcsin\left(\frac{d_{\text{ground}}}{d_{\text{slant}}}\right) \times \left(\frac{180}{\pi}\right), \quad (8)$$

where d_{slant} is computed as $d_{\text{slant}} = \|\mathbf{l}_s - \mathbf{l}_t\|$. The ground distance d_{ground} calculation formula is provided in the appendix. The Edge-Sat is considered eligible for

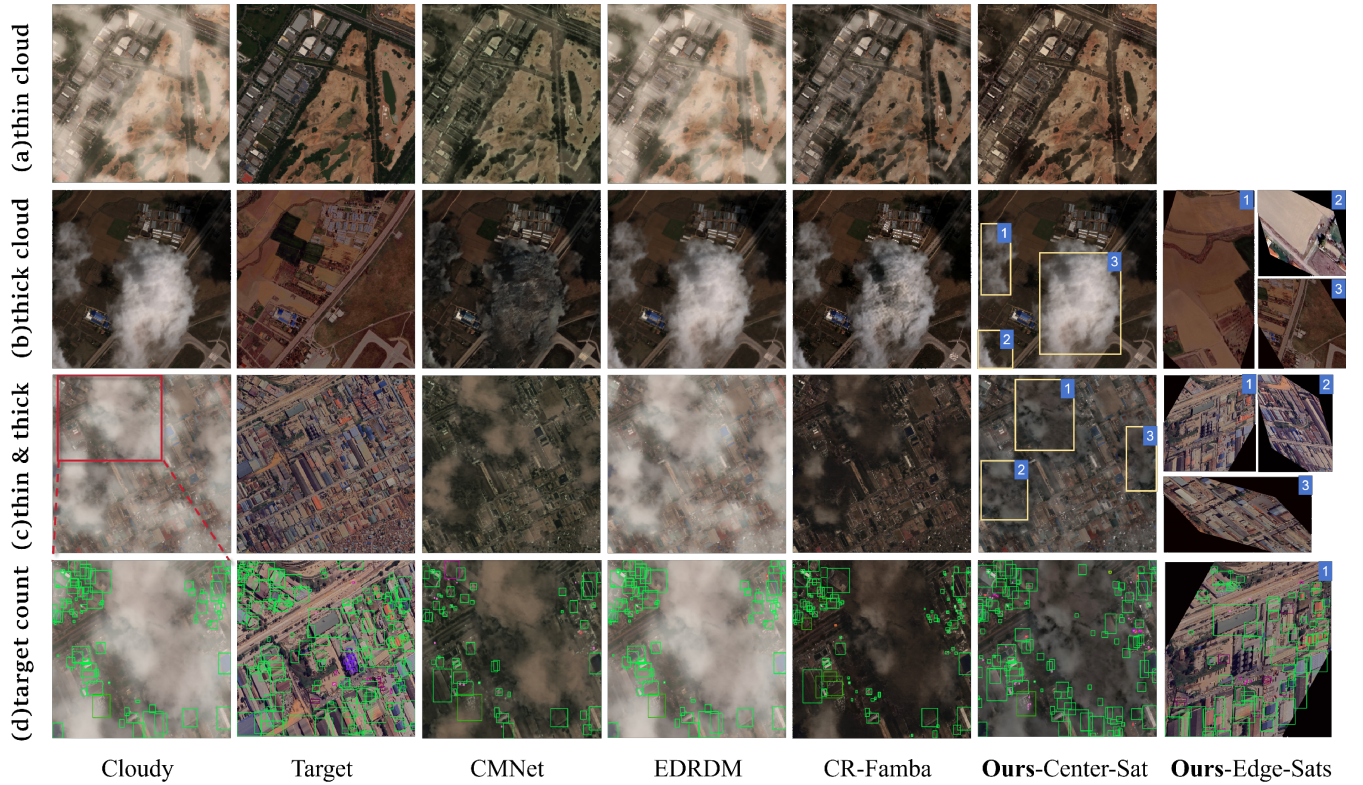


Figure 3: We compare our approach with state-of-the-art cloud removal models for remote sensing images. (a) represents scenes with only thin clouds, (b) shows scenes with only thick clouds, and (c) depicts complex scenarios containing both thin and thick clouds. (d) We selected a location from (c) that contains both thick and thin clouds for zoom-in, and show the number of target detections Across all scenarios, our method demonstrates superior performance.

imaging when the following geometric condition is satisfied $\theta_{\text{off-nadir}} < \theta_{\text{max}}$, indicating that the viewing geometry remains within the allowable off-nadir angle limit. When the off-nadir angle reaches the maximum threshold along the satellite’s imaging trajectory, the corresponding satellite position is defined as the Imaging End Point. Together with the previously determined Imaging Start Point, these two end-points define the imaging window for the Edge-Sat.

In general, when executing the scheduling strategy, Center-Sat first obtains the positions of all Edge-Sats in the constellation through inter-satellite communication and filters out a subset $\mathcal{S}_1 = \{sat_1, sat_2, \dots\}$ that satisfies Condition 1. For each satellite in \mathcal{S}_1 , Center-Sat checks Condition 2. After applying both filtering conditions, the resulting set $\mathcal{S}_2 = \{sat_1, sat_2, \dots\}$ contains the Edge-Sats eligible for imaging. To select the optimal satellite from \mathcal{S}_2 , Center-Sat compares the available imaging window duration. The optimal satellite sat^* is selected by maximizing the imaging window duration:

$$sat^* = \arg \max_{sat_i \in \mathcal{S}_2} (T_{\text{end}}(sat_i) - T_{\text{start}}(sat_i)), \quad (9)$$

where $T_{\text{start}}(sat_i)$ denotes the time when satellite sat_i first meets the imaging condition, and $T_{\text{end}}(sat_i)$ denotes the time when it reaches θ_{max} . Therefore, $T_{\text{end}}(sat_i) - T_{\text{start}}(sat_i)$ represents the length of the imaging window for satellite sat_i .

Content-Aware Selective Transmission

To ensure the completeness of the system, image downlink must also be taken into consideration. The objective of the downlink strategy is to maximize the transmission of cloud-free imagery under fixed bandwidth constraints, thereby minimizing bandwidth waste. After processing by the Thin Cloud module, a satellite image can be represented as $I = I_{\text{eff}} \cup I_{\text{cloud}}$, where I_{eff} denotes the cloud-free regions, and I_{cloud} denotes the thick cloud covered regions. We propose that, during the downlink process, Center-Sat omits I_{cloud} and encodes only the cloud-free region for transmission, $T = \text{Encode}(I_{\text{eff}})$. This significantly saves communication resources. For the obscured regions I_{cloud} , re-imaging is performed by Edge-Sats, guided by the Thick Cloud Detector Module and Edge-Sat Selector. These Edge-Sats encode and transmit the re-imaged regions to the ground station according to imaging commands from the Center-Sat.

Experiments

Dataset

The xView dataset (Lam et al. 2018) provides geospatial metadata for each image and contains natural cloud occlusions along with annotated ground targets, making it well suited for evaluating our method under realistic conditions.

Model	Constellation	Target Count(*10 ³) under Different Area Sizes					
		25 km ²	50 km ²	75 km ²	100 km ²	125 km ²	150 km ²
YOLO-v8 (Redmon et al. 2016)		2.48	4.78	6.88	7.21	8.37	10.26
UnDFN (Wang et al. 2023)	✓	2.48	4.79	6.85	7.17	8.34	10.22
Kodan (Denby et al. 2023)		2.49	4.61	6.80	7.01	8.14	9.77
CMNet (Liu, Pan, and Shi 2024)		2.79	5.74	8.94	10.75	12.54	14.96
EDRDM (Liu et al. 2025)		2.56	5.49	8.20	8.86	10.21	12.46
CR-Famba (Liu, Pan, and Shi 2025)		2.64	5.59	8.47	9.34	10.52	12.70
SCo-Cloud-2 (Ours)	✓	3.46	6.90	10.32	13.79	17.25	20.60
SCo-Cloud-4 (Ours)	✓	3.76	7.50	11.23	14.99	18.75	22.39
SCo-Cloud-6 (Ours)	✓	4.93	11.30	16.30	21.27	24.32	27.51

Table 1: Target Count (*10³) under varying satellite observation area sizes. The numbers “2,” “4,” and “6” following SCo-Cloud represent the number of Edge-Sats in the architecture.

We selected all cloud-obscured images from xView, as well as cloud-free images with accurately aligned coordinates sourced from Google Earth. In addition, we collected 1,219 cloud-free images from xView and applied a cloud synthesis algorithm to simulate varying levels of cloud coverage to construct thin cloud removal dataset.

Baselines

To assess the effectiveness of our proposed framework, we compared it with five previous models: (1) UnDFN (Wang et al. 2023), an unsupervised domain factorization network that models multi-temporal remote sensing images to achieve thick cloud removal; (2) Kodan (Denby et al. 2023), a constellation framework that performs onboard cloud detection and filtering; (3) CMNet (Liu, Pan, and Shi 2024), a cascaded memory network that gradually restores remote sensing image details for cloud removal; (4) EDRDM (Liu et al. 2025), an enhanced diffusion model to achieve remote sensing cloud removal; and (5) CR-Famba (Liu, Pan, and Shi 2025), a cloud removal network that integrates frequency-domain detail enhancement and Mamba model. These methods represent the cutting edge in handling thick clouds in images. To evaluate the effectiveness of the Thin Cloud Filter in our framework, we compared it with three state-of-the-art thin cloud removal models: (1) AIDTransformer (Kulkarni and Murala 2023), RSDformer (Song et al. 2023a), (2) Dehazer (Guo et al. 2022), and (3) DehazeFormer (Song et al. 2023b).

Implementation Details

Our implementation is based on PyTorch and is trained and evaluated on NVIDIA A100 GPUs. We use the Adam optimizer with an initial learning rate of 1×10^{-4} , and train the model for 80 epochs with a batch size of 16 and a patch size of 256. Evaluation is performed every 10 epochs. More experimental details are provided in the Appendix.

Evaluation Metrics

We have developed two evaluation metrics, Target Count and Image Utilization Factor. Target Count refers to the number of targets present in the satellite-captured images.

Since our dataset contains a large number of labeled targets (e.g., buildings, oil barrels, playgrounds, etc.), and some of these targets may be obscured by cloud cover, we evaluate model performance by comparing the number of visible targets under certain constraints. A higher value indicates more targets are captured, implying better model performance. Image Utilization Factor is defined as $(g_i - y_i)/g_i$, where g_i denotes the total area of the captured image and y_i represents the invalid portion discarded due to cloud occlusion. This metric reflects the effective utilization of the captured imagery by the system. A higher value indicates less data waste and stronger observation capability.

Main Experiments

As shown in Table 1, our method outperforms the baseline across satellite observation areas of varying sizes. This indicates that, for images captured during the same orbital duration of the satellite, our approach enables more comprehensive target acquisition within the designated observation regions. Furthermore, we conclude that the advantage of SCo-Cloud becomes more pronounced as the observation area increases. Since satellites typically operate in orbit for several decades after launch, the long-term impact of our method is substantial. The performance advantage is most significant when the number of Edge-Sats is six. Even when the number of Edge-Sats is reduced to two, there is still a notable improvement in the number of detected targets. As shown in Figure 3, it presents the target recognition performance of our method and the baseline under cloud-covered scenarios.

Ablation Study

Thin Cloud Filter. We conducted an ablation study to verify the necessity of each module in the proposed system. In Table 2, under the same observation area, the absence of the thin cloud filter leads to an increase in the number of regions identified for reacquisition. Given a fixed number of Edge-Sats, this increase causes certain regions to remain unassigned, resulting in missed targets.

Edge-Sat Selector. We modified the Edge-Sat Selector so that it no longer follows the proposed hierarchical con-

Module	Method	Target Count(*10 ³) under Different Area Sizes				
		50km ²	80km ²	110km ²	150km ²	200km ²
Thin Cloud Filter	Delete	14.89	23.94	33.05	44.72	59.73
	Add	15.50	24.69	34.08	46.42	61.61
Edge-Sat Selector	Random selection	4.86	7.85	10.59	12.98	19.93
	Algorithmic selection	15.27	24.36	33.89	40.51	62.01

Table 2: Ablation studies for the SCo-Cloud framework. Target Count(*10³) under varying satellite observation areas with modifications to key framework modules: Thin Cloud Filter and Edge-Sat Selector.

Model	Thin Haze		Thick Haze	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑
AIDTransformer	17.60	0.80	15.67	0.75
RSDformer	<u>18.42</u>	0.86	18.02	0.87
Dehazer	18.17	<u>0.88</u>	13.59	0.60
DehazeFormer	17.94	0.74	<u>19.11</u>	<u>0.76</u>
Ours	28.32	0.93	32.85	0.97

Table 3: Quantitative comparisons of various advanced thin cloud removal models.

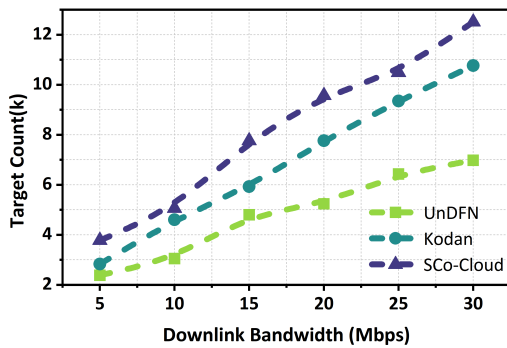


Figure 4: Quantitative comparisons of our approach with advanced models by evaluating the Target Count (*10³) under varying downlink bandwidths.

strained scheduling algorithm; instead, it randomly selects Edge-Sat. In Table 2, under this circumstance, the number of observable targets has dropped significantly, to approximately 30% of the original.

Motivation Study

We evaluate different methods based on the number of targets contained in the transmitted images under such constraints. Considering bandwidth is usually limited to a few tens of Mbps, we focus on a representative bandwidth range from 5 Mbps to 30 Mbps. As shown in Figure 4, our proposed method surpasses baseline approaches by 20% to 70%, demonstrating superior task utility under the low-bandwidth conditions commonly encountered in real-world satellite-to-ground communication scenarios. In the Thin Cloud Filter, we fine-tune DehazeFormer (Song et al. 2023b) on our thin-cloud removal dataset and compare it with sev-

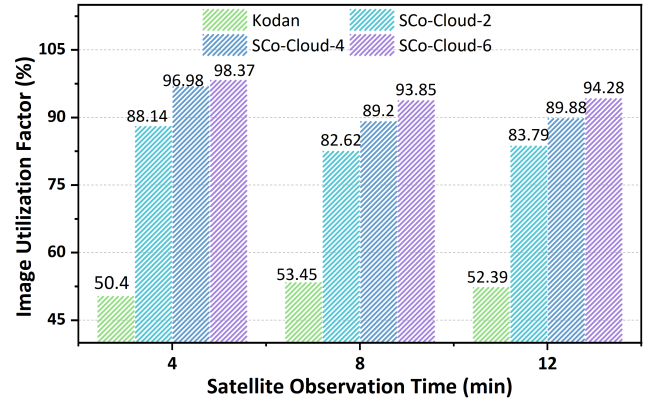


Figure 5: Image Utilization Factor (%) under varying satellite observation time. The numbers “2,” “4,” and “6” following SCo-Cloud represent the number of Edge-Sats in the architecture.

eral state-of-the-art thin cloud removal models. Our model achieves the best performance in two commonly used metrics: PSNR and SSIM, as shown in Table 3. As shown in Figure 5, Kodan maintains a utilization rate of around 50%, indicating that nearly half of the captured image area is unusable due to cloud cover, leading to significant observational inefficiency. In contrast, SCo-Cloud with only two Edge-Sats, the utilization already exceeds 88%; when scaling up to six Edge-Sats, the utilization reaches over 94%.

Discussion of Response Time

Satellites communicate via laser links (10–20 Gbps) transmitting just 200 KB of task and status data with 3.33–6.67 ms latency. The baseline constellation requires several days for a revisit, whereas our approach completes thin-cloud removal in 208.70 seconds, re-imaged localization in 130.74 seconds, and Edge-Sat selection in 5.08 ms.

Conclusion

This work introduces a constellation system designed to address the challenge of cloud occlusion. SCo-Cloud leverages onboard cloud detection and a hierarchical constraint-based scheduling algorithm to enable inter-satellite collaboration for re-imaging cloud-covered regions. Extensive experiments show that, compared to baseline methods, SCo-Cloud can increase the number of captured targets.

Acknowledgments

We thank the anonymous reviewers for their insightful comments and suggestions. The corresponding author is Qian Li. The authors of this paper were supported by the NSFC through grant No.62425203 and No.62032003, No.62402054, the China Postdoctoral Science Foundation through grant 2024M760279, and the Postdoctoral Fellowship Program and China Postdoctoral Science Foundation under grant BX20250390.

References

- Cheng, Z.; Denby, B.; McCleary, K.; and Lucia, B. 2024. EagleEye: Nanosatellite constellation design for high-coverage, high-resolution sensing. In *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 1*, 117–132.
- Cheng, Z.; and Lucia, B. 2025. Nanosatellite Constellation and Ground Station Co-design for Low-Latency Critical Event Detection. *arXiv preprint arXiv:2503.01756*.
- Czerkawski, M.; Atkinson, R.; Michie, C.; and Tachtatzis, C. 2023. Satellitecloudgenerator: controllable cloud and shadow synthesis for multi-spectral optical satellite images. *Remote Sensing*, 15(17): 4138.
- Denby, B.; Chintalapudi, K.; Chandra, R.; Lucia, B.; and Noghabi, S. 2023. Kodan: 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.
- Du, K.; Cheng, Y.; Olsen, P.; Noghabi, S.; and Jiang, J. 2025. Earth+: On-board satellite imagery compression leveraging historical earth observations. In *Proceedings of the 30th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 1*, 361–376.
- Ebel, P.; Garnot, V. S. F.; Schmitt, M.; Wegner, J. D.; and Zhu, X. X. 2023. UnCRtainTS: Uncertainty quantification for cloud removal in optical satellite time series. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2086–2096.
- Ebel, P.; Xu, Y.; Schmitt, M.; and Zhu, X. X. 2022. SEN12MS-CR-TS: A remote-sensing data set for multimodal multitemporal cloud removal. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–14.
- Furutanpey, A.; Zhang, Q.; Raith, P.; Pfandzelter, T.; Wang, S.; and Dustdar, S. 2025. Fool: Addressing the downlink bottleneck in satellite computing with neural feature compression. *IEEE Transactions on Mobile Computing*.
- Ghildiyal, S.; Goel, N.; Singh, S.; Lal, S.; Kawsar, R.; El Saddik, A.; and Saini, M. 2024. SSGAN: Cloud removal in satellite images using spatio-spectral generative adversarial network. *European Journal of Agronomy*, 161: 127333.
- Gómez, P.; and Meoni, G. 2024. Tackling the satellite downlink bottleneck with federated onboard learning of image compression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6809–6818.
- Guo, C.-L.; Yan, Q.; Anwar, S.; Cong, R.; Ren, W.; and Li, C. 2022. Image dehazing transformer with transmission-aware 3d position embedding. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 5812–5820.
- Hao, Z.; Haseeb, M.; Xiangtian, Z.; Tahir, Z.; Mahmood, S. A.; Tariq, A.; Aslam, R. W.; Abdullah-Al-Wadud, M.; and El-Askary, H. 2024. Multi-Temporal Analysis of Urbanization-Driven Slope and Ecological Impact Using Machine-Learning and Remote Sensing Techniques. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
- Kuang, X.; Xiang, S.; and Guo, J. 2025. Soil moisture retrieval and trend prediction using multi-temporal remote sensing data: An interpretable deep regression approach. *Expert Systems with Applications*, 128172.
- Kulkarni, A.; and Murala, S. 2023. Aerial image dehazing with attentive deformable transformers. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 6305–6314.
- Lam, D.; Kuzma, R.; McGee, K.; Dooley, S.; Laielli, M.; Klaric, M.; Bulatov, Y.; and McCord, B. 2018. xview: Objects in context in overhead imagery. *arXiv preprint arXiv:1802.07856*.
- Li, C.; Liu, X.; and Li, S. 2023. Transformer meets GAN: Cloud-free multispectral image reconstruction via multisensor data fusion in satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1–13.
- Li, X.; Zhao, X.; Wang, F.; and Ren, P. 2024. HF-T2CR: High-fidelity thin and thick cloud removal in optical satellite images through SAR fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1–13.
- Liu, H.; Ma, Y.; Yan, M.; Chen, Y.; Peng, D.; and Wang, X. 2024. Dida: Disambiguated domain alignment for cross-domain retrieval with partial labels. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, 3612–3620.
- Liu, J.; Pan, B.; and Shi, Z. 2024. Cascaded memory network for optical remote sensing imagery cloud removal. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1–11.
- Liu, J.; Pan, B.; and Shi, Z. 2025. CR-Famba: A Frequency-Domain Assisted Mamba for Thin Cloud Removal in Optical Remote Sensing Imagery. *IEEE Transactions on Multimedia*, 1–10.
- Liu, Y.; Li, W.; Guan, J.; Zhou, S.; and Zhang, Y. 2025. Effective cloud removal for remote sensing images by an improved mean-reverting denoising model with elucidated design space. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 17851–17861.
- Long, C.; Li, X.; Jing, Y.; and Shen, H. 2023. Bishift networks for thick cloud removal with multitemporal remote sensing images. *International Journal of Intelligent Systems*, 2023(1): 9953198.
- Luo, G.; Weng, L.; Li, Y.; Sun, Y.; Hong, Y.; Wu, Y.; Luo, R.; Wang, L.; Wang, C.; and Chen, L. 2025. FireExpert:

- Fire Event Identification and Assessment Leveraging Cross-Domain Knowledge and Large Language Model. *IEEE Transactions on Mobile Computing*, 24(6): 4794–4810.
- Oquab, M.; Darcet, T.; Moutakanni, T.; Vo, H.; Szafraniec, M.; Khalidov, V.; Fernandez, P.; Haziza, D.; Massa, F.; El-Nouby, A.; et al. 2023. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*.
- Pan, J.; Xu, J.; Yu, X.; Ye, G.; Wang, M.; Chen, Y.; and Ma, J. 2024. HDRSA-Net: Hybrid dynamic residual self-attention network for SAR-assisted optical image cloud and shadow removal. *ISPRS Journal of Photogrammetry and Remote Sensing*, 218: 258–275.
- Redmon, J.; Divvala, S.; Girshick, R.; and Farhadi, A. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 779–788.
- Song, T.; Fan, S.; Li, P.; Jin, J.; Jin, G.; and Fan, L. 2023a. Learning an effective transformer for remote sensing satellite image dehazing. *IEEE Geoscience and Remote Sensing Letters*, 20: 1–5.
- Song, Y.; He, Z.; Qian, H.; and Du, X. 2023b. Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 32: 1927–1941.
- Sui, J.; Ma, Y.; Yang, W.; Zhang, X.; Pun, M.-O.; and Liu, J. 2024. Diffusion enhancement for cloud removal in ultra-resolution remote sensing imagery. *arXiv preprint arXiv:2401.15105*.
- Tao, B.; Chabra, O.; Janveja, I.; Gupta, I.; and Vasisht, D. 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.
- Tao, C.; Fu, S.; Qi, J.; and Li, H. 2022. Thick cloud removal in optical remote sensing images using a texture complexity guided self-paced learning method. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–12.
- Wang, J.; Olsen, P. A.; Conn, A. R.; and Lozano, A. C. 2016. Removing clouds and recovering ground observations in satellite image sequences via temporally contiguous robust matrix completion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2754–2763.
- Wang, J.-L.; Zhao, X.-L.; Li, H.-C.; Cao, K.-X.; Miao, J.; and Huang, T.-Z. 2023. Unsupervised domain factorization network for thick cloud removal of multitemporal remotely sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1–12.
- Wang, M.; Song, Y.; Wei, P.; Xian, X.; Shi, Y.; and Lin, L. 2024. IDF-CR: Iterative diffusion process for divide-and-conquer cloud removal in remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1–14.
- Wu, Y.; Wang, Z.; Yuan, Y.; Gong, M.; Li, H.; Zhang, M.; Ma, W.; and Miao, Q. 2025. MUCD: Unsupervised Point Cloud Change Detection via Masked Consistency. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 8505–8513.
- Xiang, X.; Tan, Y.; and Yan, L. 2024. Cloud-guided fusion with SAR-to-optical translation for thick cloud removal. *IEEE Transactions on Geoscience and Remote Sensing*.
- Xie, Y.; Li, Z.; Bao, H.; Jia, X.; Xu, D.; Zhou, X.; and Skakun, S. 2023. Auto-CM: Unsupervised deep learning for satellite imagery composition and cloud masking using spatio-temporal dynamics. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 14575–14583.
- Xu, M.; Deng, F.; Jia, S.; Jia, X.; and Plaza, A. J. 2022. Attention mechanism-based generative adversarial networks for cloud removal in Landsat images. *Remote sensing of environment*, 271: 112902.
- Yuan, S.; Lin, G.; Zhang, L.; Dong, R.; Zhang, J.; Chen, S.; Zheng, J.; Wang, J.; and Fu, H. 2024. FUSU: A multi-temporal-source land use change segmentation dataset for fine-grained urban semantic understanding. *Advances in Neural Information Processing Systems*, 37: 132417–132439.
- Zhang, Q.; Yuan, X.; Xing, R.; Zhang, Y.; Zheng, Z.; Ma, X.; Xu, M.; Dustdar, S.; and Wang, S. 2024. Resource-efficient in-orbit detection of earth objects. In *IEEE INFOCOM 2024-IEEE Conference on Computer Communications*, 551–560. IEEE.
- Zhang, Z.; Angelov, P.; Soares, E.; Longepe, N.; and Mathieu, P. P. 2022. An interpretable deep semantic segmentation method for earth observation. In *2022 IEEE 11th International Conference on Intelligent Systems (IS)*, 1–8. IEEE.
- Zheng, Z.; Tian, S.; Ma, A.; Zhang, L.; and Zhong, Y. 2023. Scalable multi-temporal remote sensing change data generation via simulating stochastic change process. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 21818–21827.
- Zhou, H.; Kao, C.-H.; Phoo, C. P.; Mall, U.; Hariharan, B.; and Bala, K. 2024. AllClear: A Comprehensive Dataset and Benchmark for Cloud Removal in Satellite Imagery. *arXiv preprint arXiv:2410.23891*.
- Zi, Y.; Xie, F.; Song, X.; Jiang, Z.; and Zhang, H. 2021. Thin cloud removal for remote sensing images using a physical-model-based CycleGAN with unpaired data. *IEEE Geoscience and Remote Sensing Letters*, 19: 1–5.