Distributed On-Orbit Privacy Preservation for Satellite Imagery via Lightweight Neural Inpainting

Wenbo Wu[1], Cheng Tan[1], Zhishu Shen[1*], Di Huang[1], Qiushi Zheng[2], Jiong Jin[2]
[1] School of Computer Science and Artificial Intelligence, Wuhan University of Technology, China
[2] School of Science, Computing and Engineering Technologies, Swinburne University of Technology, Australia

***Abstract*—Recent advances in Earth observation satellites have enabled high-resolution imagery for Internet of Remote Things (IoRT) applications, but such imagery often contains privacy-sensitive information that may be exposed during transmission or processing. Addressing privacy protection directly on orbit is challenging due to the strict computational, energy, and communication constraints of Low Earth Orbit (LEO) satellite systems and their limited ground-station contact windows. In this paper, we propose a Distributed On-orbit Privacy Protection framework *DOPP* that integrates lightweight neural detection and inpainting with inter-satellite collaborative computing. Each satellite employs a quantized YOLOv9-based detection model to identify privacy-sensitive regions and generate compact binary masks, which enable non-overlapping image slicing and parallel inpainting across neighboring satellites via inter-satellite links. A fine-tuned Mobile Inpainting Generative Adversarial Network (Mi-GAN) reconstructs masked regions with low computational overhead while preserving visual consistency. To ensure timely downlink transmission, a visibility-aware aggregation strategy dynamically selects the satellite with the most imminent ground-station access to assemble and transmit the processed imagery. Simulations on real-world remote sensing datasets demonstrate that *DOPP* achieves effective privacy preservation with high accuracy, while significantly reducing processing latency and energy consumption , resulting in a 15–20% lower Energy-Delay Product (EDP) compared with existing onboard cooperative baselines. These results highlight the feasibility of lightweight neural privacy protection for real-time, resource-constrained satellite imaging systems.**

***Index Terms*—Privacy-preserving learning, satellite image inpainting, distributed neural processing**

I. Introduction

Recent advances in Deep Neural Networks (DNN) have significantly enhanced the semantic understanding capabilities of Earth observation systems, enabling Low Earth Orbit (LEO) satellites to support a wide range of Internet of Remote Things (IoRT) applications, such as disaster monitoring, environmental surveillance, and large-scale infrastructure analysis [1]. High-resolution satellite imagery, combined with powerful vision models, allows automatic detection and interpretation of fine-grained objects and activities over vast geographic areas. However, the same learning-enabled perception capabilities also raise serious privacy concerns, as sensitive information, such as vehicles, facilities, or human activities, may be inadvertently revealed in captured imagery [2].

Protecting privacy in satellite imagery is particularly challenging in Space Information Networks (SINs), where LEO satellites operate under strict constraints on computation, communication, and energy [3]. Unlike ground-based systems, satellites have limited onboard resources and only short communication windows with Ground Stations (GSs) during each orbital pass [4]. As a result, transmitting raw images to the ground for privacy processing not only increases the risk of interception over exposed satellite–ground links, but also introduces substantial delays that violate the real-time requirements of many IoRT applications. These constraints necessitate privacy-preserving processing directly on orbit, where neural inference must be both efficient and timely.

Existing privacy protection approaches typically rely on deep learning pipelines that combine object detection with image inpainting to identify and conceal sensitive regions. Qiu *et al.* [5] proposed a two-stage model integrating an improved Mask R-CNN and PointRend for high-precision detection, followed by an enhanced DeepFill Generative Adversarial Network (GAN) for realistic inpainting. Yao *et al.* [6] proposed PPUP-GAN for Unmanned Aerial Vehicle (UAV) imagery by separating objects of interest from background regions. Although effective in terrestrial and aerial scenarios, these approaches are less suitable for satellite platforms. Their high-capacity detection and generative models incur substantial computational and memory overhead, and their sequential execution on a single device limits inference speed, often exceeding a single satellite–ground contact window. As a result, directly deploying existing neural privacy-preserving pipelines on resource-constrained satellite systems remains impractical.

To address these challenges, we propose a Distributed On-orbit Privacy Protection framework named *DOPP* that restructures the neural privacy-preserving pipeline to enable efficient, parallel inference across multiple satellites. Rather than processing entire images on a single satellite, *DOPP* decomposes the privacy protection task into lightweight neural components that can be executed collaboratively. Each satellite employs a quantized YOLOv9-based detection model [7], using mixed-precision and INT8 post-training techniques to reduce memory usage and inference latency, to identify privacy-sensitive regions and generate compact binary masks. These masks enable non-overlapping image slicing, allowing image segments to be distributed to neighboring satellites via Inter-

The first two authors contributed equally to this work.
* Corresponding author (z_shen@ieee.org).

Satellite Links (ISLs) for parallel neural inpainting.

For image restoration, we adopt a lightweight Mobile Inpainting Generative Adversarial Network (Mi-GAN)-based inpainting network [8] that is fine-tuned on remote sensing imagery to balance reconstruction quality and computational efficiency. Instead of processing full images, the inpainting network operates only on detection-guided masked regions and independently processes image slices, thereby significantly reducing per-satellite inference latency and energy consumption. To ensure timely delivery of processed results, *DOPP* further incorporates a visibility-aware aggregation strategy that dynamically selects the satellite with the most imminent GS access to aggregate the repaired image and perform downlink transmission. This design enables privacy-preserving neural inference to be completed within strict orbital timing constraints while minimizing exposure of raw data.

Unlike prior works that focus on designing new detection or inpainting architectures, this paper emphasizes a system-level neural inference paradigm for privacy preservation under strict on-orbit constraints, where model selection, task decomposition, and inter-satellite coordination jointly determine feasibility. The main contributions of this paper are summarized as follows:

- We propose a distributed on-orbit inference framework that parallelizes neural processing across multiple satellites using compact binary masks, enabling real-time execution without transmitting raw imagery to the ground.
- We design a privacy-preserving vision pipeline for satellite imagery that combines quantized object detection with lightweight generative inpainting, tailored to the computational and energy constraints of on-orbit platforms.
- We introduce a visibility-aware aggregation mechanism that selects the optimal satellite for result fusion and downlink transmission based on orbital predictions, reducing waiting time and end-to-end latency.
- We evaluate the proposed framework on real-world remote sensing datasets, demonstrating significant reductions in processing latency and energy consumption while maintaining high privacy protection accuracy compared with existing onboard cooperative baselines.

II. Related Work

With the widespread application of UAVs and LEO satellites in Earth observation missions, there is growing concern about the leakage of privacy-sensitive data captured during these missions. In recent years, many studies have proposed various methods to address these issues.

Lin *et al.* [9] highlighted fundamental security and privacy challenges in the Internet of Drones, such as location leakage, untrusted cloud storage, and malicious interference, and discuss lightweight cryptographic and identity-based solutions for resource-constrained aerial platforms. Zhang *et al.* [10] proposed a lightweight privacy-preserving system for IoT-based remote sensing images using visual cryptography, enabling efficient encryption of high-resolution imagery without incurring significant computational overhead. To address the inherent noise introduced by visual cryptography, they further integrate denoising neural networks to recover classification accuracy. Zhang *et al.* [11] presented a generalized framework for visual privacy protection, structuring the design process around privacy definition, scenario abstraction, algorithm design, and evaluation. They demonstrate its effectiveness with a diffusion-based model that jointly controls identity anonymization and attribute manipulation through multi-condition inputs. Upenik *et al.* [12] proposed an inpainting-based privacy protection method for omnidirectional images that mitigates geometric distortion by operating in the viewport domain. Specifically, regions to be protected are cropped from the equirectangular projection into planar viewports, inpainted using existing methods, and then mapped back to the panorama. This approach effectively improves inpainting quality in polar regions, as confirmed by subjective evaluations, and supports reversible privacy protection through encryption of the original viewports.

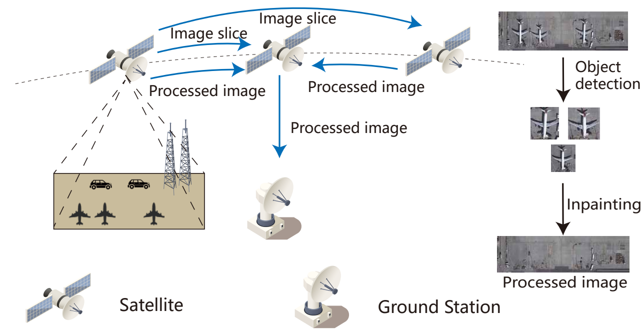


Fig. 1: Network model.

Although existing methods have made progress, they largely overlook the unique constraints of SINs. Compared with UAVs, LEO satellites have much shorter ground-station contact windows, which places strict requirements on the timeliness of privacy-preserving image processing. To address this challenge, we propose a lightweight distributed on-orbit framework that enables parallel privacy protection via inter-satellite collaboration. A visibility-aware aggregation strategy further ensures timely result fusion and downlink transmission, minimizing end-to-end latency.

III. System Model

This section introduces a system model to characterize the constraints that govern on-orbit neural inference and collaborative processing among satellites.

A. Network Model

We consider a LEO satellite system comprising both LEO satellites and GSs. The system comprises P orbital planes. Each orbital plane $p \in \{1, 2, \ldots, P\}$ has S_p satellites, denoted by $\mathcal{S}_p = \{s_{p,1}, s_{p,2}, \ldots, s_{p,S_p}\}$. Hence, the overall set of satellites in the constellation is expressed as $\mathcal{S} = \bigcup_{p=1}^{P} \mathcal{S}_p$, each satellite is equipped with computing resources to remove privacy data from satellite images. Each satellite s moves along a trajectory $\mathbf{r}_s(t)$ around the Earth, whose orbital period is given by $T_p = 2\pi\sqrt{\frac{a_s^3}{\mu}}$, where a_s is the semi-major axis of the orbit, and $\mu = 3.98 \times 10^{14}\ \mathrm{m}^3/\mathrm{s}^2$ is the Earth's standard gravitational parameter. The motion is periodic such that $\mathbf{r}_s(t) \approx \mathbf{r}_s(t + nT_p)$ for any integer n [13]. For circular orbits, the semi-major axis satisfies $a_s = r_E + h_s$, where $r_E = 6371\ \mathrm{km}$ denotes the Earth's radius and h_s represents the altitude above the Earth's surface.

Fig. 1 illustrates an Earth observation mission in which multiple satellites within the same orbital plane collaboratively capture, process, and downlink imagery. Satellites share identical orbital parameters and are distinguished by phase angles, observing targets at the Ground Track Frame Rate (GTFR) to avoid redundant acquisition. Each captured image along the ground track is defined as a Ground Track Frame (GTF), which serves as the basic processing unit. Each satellite processes part of its captured GTF locally and offloads the remaining segments to orbital peers via ISLs for parallel privacy-preserving processing. A visibility-aware aggregation strategy then selects the satellite with the nearest communication window to the GS to fuse and downlink the results.

B. Communication Model

We consider a satellite system that includes two types of communication links: the Satellite–Ground Link (SGL) and ISL. The SGL is responsible for transmitting the inpainted satellite imagery from the satellites to the GS, whereas the ISL enables the distribution of image slices and compact neural masks, making parallel neural inference feasible within a single orbital period.

For SGL, communication between the satellite and the GS is possible only when the satellite is visible from the GS's perspective. This visibility condition is satisfied when the satellite's elevation angle meets the threshold requirement $\frac{\pi}{2} - \angle(\mathbf{r}_{GS}, \mathbf{r}_s(t) - \mathbf{r}_{GS}) \geq \alpha_e$, where α_e denotes the minimum elevation angle [13], and $\mathbf{r}_{GS}$ represents the position vector of the GS. When this condition is satisfied, a constant data rate is assumed for the SGL.

We adopt a standard free-space path-loss and SNR model to estimate transmission latency, as the focus of this work is neural inference efficiency rather than physical-layer optimization. The effective path loss between two objects x and y can be represented as

$$L_{xy} = \left(\frac{4\pi d_{xy} f_c}{c}\right)^2, \quad \forall x, y \in \mathcal{S} \cup GS, \tag{1}$$

where c denotes the speed of light, f_c is the carrier frequency of the link, and d_{xy} represents the physical distance between x and y.

The Signal-to-Noise Ratio (SNR) between two objects x and y is expressed as

$$\gamma_{xy} = \frac{P G_{xy}^{tr} G_{xy}^{re}}{K_B T B L_{xy}}, \tag{2}$$

where P denotes the transmit power, G_{xy}^{tr} and G_{xy}^{re} are the transmitting and receiving antenna gains, K_B is the Boltzmann constant, T is the noise temperature at the receiver and B denotes the bandwidth of the link. The achievable data transmission rate between x and y can be calculated as

$$\mathit{TR}_{xy} = B \log_2 (1 + \gamma_{xy}). \tag{3}$$

C. Energy Consumption Model

The Earth observation mission involves a set of N consecutive satellites, denoted as $\mathcal{S}_{p,N} = \{s_{p,i}, s_{p,i+1}, \ldots, s_{p,i+N-1}\}$, which are evenly distributed within the same orbital plane p containing S_p satellites. During an observation cycle, satellite $s_{p,u}$ captures an image of fixed size $D = w \times h$, where w and h represent the image width and height in pixels, respectively. The satellite $s_{p,u}$ then performs a privacy detection task on this image and generates a binary mask of size D. The CPU cycles required for this task can be calculated as $C_{det} = K_{det} D q_{px}$, where K_{det} is the number of CPU cycles required to complete each bit of data in the object detection task and q_{px} is the number of bits per pixel. Therefore, the time for completing the detection task can be expressed as $T_{det} = C_{det}/N_{\text{CPU}} f$, where N_{CPU} is the number of CPU cores, f is the processor's operating frequency. Then the energy consumption required for the task is represented as $E_{det} = \epsilon_0 f^3 T_{det}$, where ϵ_0 is the constant coefficient determined by the hardware architecture.

The captured image and mask are then segmented into multiple slices. Satellite $s_{p,u}$ processes one of the image slices locally to reduce computational effort, while the remaining image slices and corresponding masks are transmitted to corresponding satellites in $\mathcal{S}_{p,N}$ for collaborative processing. Each satellite independently performs the privacy protection operation on its assigned image slice. The transmission delay of image slices between satellites can be expressed as

$$T_{dis} = \frac{(N-1) D (q_{px} + 1)}{N \cdot \mathit{TR}_{ISL}}, \tag{4}$$

where TR_{ISL} is the transmission rate between satellites.

Therefore, the energy consumption can be calculated as

$$E_{dis} = P T_{dis}. \tag{5}$$

Then, the satellites in $\mathcal{S}_{p,N}$ will inpaint privacy data on the image. The CPU cycles required for this task can be calculated as $C_{inp} = K_{inp} D q_{px}/N$, where K_{inp} is the number of CPU cycles required to complete each bit of data in the image inpainting task. The inpainting delay T_{inp} required can be calculated as $T_{inp} = C_{inp}/N_{\text{CPU}} f$, Then the energy consumption is represented as $E_{inp} = \epsilon_0 f^3 T_{inp}$.

After the satellites in $\mathcal{S}_{p,N}$ complete their privacy data processing, they transmit the processing results to $s_{p,v}$ for aggregation. The delay $T_{s,v}$ and energy consumption $E_{s,v}$ for this part can be calculated using Eq. (4) and Eq. (5) as well. Satellite $s_{p,v}$ aggregates all satellite images into a single composite image and transmits it to the GS. The time and energy consumption can be calculated as

$$T_{SG} = \frac{D q_{px}}{\mathit{TR}_{SGL}}, E_{SG} = P T_{SG}, \tag{6}$$

where TR_{SGL} is the transmission rate between satellite and GS.

At this point, the processing of a GTF is completed. The latency T_{GTF} to process a GTF can be defined as

$$T_{GTF} = T_{det} + T_{dis} + T_{inp} + T_{s,v} + T_{wait}(s_{p,v}) + T_{SG}, \quad (7)$$

where $T_{wait}(s_{p,v})$ is the waiting time required for satellite $s_{p,v}$ to communicate directly with GS. And the total energy consumption can be calculated as

$$E_{GTF} = E_{det} + E_{dis} + NE_{inp} + E_{s,v} + E_{SG}. \quad (8)$$

To quantitatively evaluate the efficiency of the proposed collaborative processing framework, we adopt Energy-Delay Product (EDP) [14] as the performance metric. EDP jointly captures the trade-off between processing latency and energy consumption, which is particularly relevant for on-orbit systems operating under strict resource constraints. Specifically, EDP is defined as

$$EDP = E_{GTF} \cdot T_{GTF}. \quad (9)$$

D. Threat Model and Objective

In this work, we consider a visual privacy threat model where sensitive objects may be exposed to human inspection or automated recognition systems during transmission or storage. The goal is to prevent semantic disclosure of sensitive targets rather than provide cryptographic privacy guarantees.

The objective is to minimize the *EDP* associated with processing privacy-sensitive data in each GTF. Under constrained onboard computation and inter-satellite communication resources, EDP serves as an optimization criterion that characterizes the trade-off between execution latency and energy consumption. The optimization function $F(N)$ can be expressed as

$$F(N) = \min_{N}(\mathit{EDP(N)}), \quad (10)$$

where N denotes the number of collaborating satellites, which governs the trade-off between parallel processing efficiency and communication overhead, thereby determining the resulting EDP.

IV. METHODOLOGY

Guided by the computation, communication, and energy constraints characterized in Section III, this section introduces a distributed on-orbit privacy-preserving processing framework *DOPP*, designed to detect and repair privacy data in satellite imagery within SINs.

A. Overview

DOPP comprises two key components: *Satellite Collaborative Computing*, enabling distributed task allocation and inter-satellite coordination; and *Satellite Imagery Inpainting*, ensuring privacy preservation and visual consistency through contextual reconstruction. The overall procedure is shown in Algorithm 1 and illustrated in Fig. 2. We consider N consecutive satellites $\mathcal{S}_{p,N} = \{s_{p,i}, s_{p,i+1}, \ldots, s_{p,i+N-1}\}$ uniformly distributed in an orbital plane S_p. Each satellite establishes a downlink SGL with the GS during its visible window. Upon capturing a GTF, a satellite partitions the image into slices and distributes them to orbital peers for parallel processing. For each slice, the detection module generates binary masks of privacy-sensitive regions, which guide subsequent inpainting and encryption to prevent privacy leakage. Before completing the computation, each satellite selects the satellite $s_{p,v}$ with the closest upcoming communication window based on visibility-aware aggregation, and forwards its intermediate results to satellite $s_{p,v}$ for aggregation and subsequent downlink transmission.

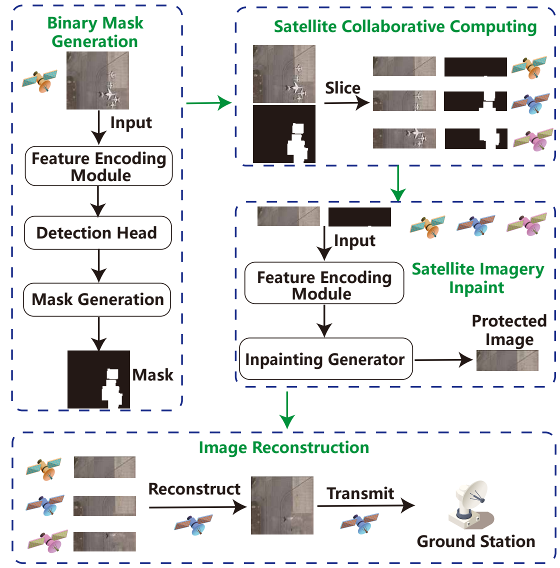


Fig. 2: The overall workflow of *DOPP*.

Algorithm 1 Distributed On-Orbit Privacy Protection Framework (*DOPP*)

Collaborative satellite set $S_{p,N}$
while $s_{p,u} \in S_{p,N}$ captures an image I **do**
    Generate binary mask $M \leftarrow I$;
    Slice $(I, M) \rightarrow \{(I_1, M_1), (I_2, M_2), \ldots, (I_N, M_N)\}$;
    for $s_{p,i}$ in $S_{p,N}$ **do**
        Transmit (I_i, M_i) to $s_{p,i}$;
    end for
end while
for $s_{p,j} \in S_{p,N}$ **do**
    Estimate processing time and select aggregator $s_{p,v}$;
    Inpaint image $\hat{I}_j \leftarrow I_j$;
    Transmit $\hat{I}_j$ to $s_{p,v}$;
end for
$s_{p,v}$ aggregates $\{\hat{I}\} \rightarrow I_{final}$;
$s_{p,v}$ transmits I_{final} to GS.

B. Satellite Collaborative Computing

For privacy data detection, we adopt a lightweight YOLOv9-based model configured with a reduced backbone. To meet on-orbit constraints, the detection network is deployed with post-training INT8 quantization, reducing inference latency and memory footprint while maintaining sufficient detection accuracy for privacy-sensitive objects in remote sensing imagery.

After capturing an image I of size D, the image I is processed by a convolutional backbone to extract features and predict object information

$$F_{det} = Conv(I; \theta), \tag{11}$$

where θ represents the network parameters. After feature extraction, the segmentation branch generates K global mask prototypes from the feature maps

$$P = \Psi(F_{det}; \theta_P), \tag{12}$$

where $\Psi(\cdot)$ denotes the prototype generation network and θ_P represents the prototype generation network parameters. For each detected object belonging to the target class c^p, the final mask is obtained by combining the prototypes P with the corresponding predicted mask coefficients $\boldsymbol{\alpha}_{c^p}$

$$M_i = \mathcal{D}_k^t(\sigma(P\boldsymbol{\alpha}_{c^p}) \cdot \mathbb{I}[c = c^p]), \tag{13}$$

where $\sigma(\cdot)$ is the sigmoid function producing the binary segmentation mask M, $\mathcal{D}_k^t$ is a morphological dilation operator and t is dilate iterations and k is dilate kernel size. We apply morphological dilation to the predicted masks to compensate for boundary uncertainty inherent in lightweight object detectors. This conservative expansion reduces the risk of incomplete privacy coverage at object edges, at the cost of a marginal increase in inpainting region size.

Then, the satellite $s_{p,u}$ divides image I and mask M into N equal slices $\{(I_1, M_1), (I_2, M_2), \ldots, (I_N, M_N)\}$. Since each satellite only needs to inpaint its assigned slice according to the mask, overlapping regions between adjacent slices are unnecessary. This design avoids redundant computation while maintaining consistent repair quality across slice boundaries. The satellite processes one slice locally and distributes the remaining $N-1$ slices to other satellites for parallel processing. Each satellite's assigned portion of the GTF is determined sequentially based on the total number of satellites N and their respective orbital indices.

Each satellite then performs image inpainting on its assigned image slice based on the binary mask. Details are described in Section IV-C. Prior to completing all slices, a satellite $s_{p,v}$ is selected for result aggregation based on orbital predictions, ensuring that the aggregation process does not introduce idle waiting. Satellite $s_{p,v}$ is chosen such that, upon completion of all slice processing, its communication link to the GS is available and the remaining time in the communication window is sufficient to transmit the aggregated results. If the link is currently unavailable, $s_{p,v}$ establishes a connection to the GS at the earliest possible opportunity. After all satellites have completed their processing, the selected satellite $s_{p,v}$ must satisfy the following requirements $s_{p,v} = \arg\min_{s_{p,k} \in \mathcal{S}_{p,N}} T_{wait}(s_{p,k})$. Since each satellite processes an image slice of identical size and possesses the same computational capability, the processing time T_{inp} for each assigned slice is uniform across all satellites. As the satellites in $\mathcal{S}_{p,N}$ operate nearly in parallel, the latency required for all satellites to process their respective portions of the image can be expressed as $T_{ex} = T_{dis} + T_{inp}$.

Owing to the deterministic nature of satellite motion, the state of each satellite after a time interval T_{ex} can be predicted with high accuracy. Based on the predicted positions, the most suitable satellite $s_{p,v}$ is selected for result aggregation. Once the privacy-preserving operations on the image slices are completed, the remaining satellites in $\mathcal{S}_{p,N}$ transmit their processed slices to $s_{p,v}$ via ISLs. Satellite $s_{p,v}$ then aggregates the received slices and forwards the consolidated result to GS.

In our framework, the object detection is performed on the full image to preserve global semantic context, whereas inpainting is inherently local to masked regions. This asymmetric design enables reliable detection while allowing inpainting to be efficiently parallelized across satellites.

C. Satellite Imagery Inpaint

We adopt Mi-GAN instead of diffusion-based or large transformer inpainting models due to its low inference latency and memory footprint, which are critical for on-orbit deployment.

After satellite $s_{p,i}$ receives the image slice I_i and mask M_i for which it is responsible, the satellite $s_{p,i}$ employs an inpainting network to construct the masked regions in its assigned image slice. The inpainting network first encodes the known context and mask information

$$F_{inp} = \Phi(I_i \odot (1 - M_i), M_i), \tag{14}$$

where $\Phi(\cdot)$ denotes the convolutional encoder and $\odot$ represents element-wise multiplication. Instead of reconstructing the missing content in a single stage, the decoder generates a set of intermediate repair outputs at multiple scales. Specifically, at scale level l, the decoder predicts a repair component

$$x_i^{(l)} \in \mathbb{R}^{\frac{h}{2^l} \times \frac{w}{2^l} \times 3}, \quad l = 0, \ldots, L, \tag{15}$$

where lower-resolution outputs model global structure, while higher-resolution outputs refine local details. These multi-scale repair predictions are then recursively aggregated in a coarse-to-fine manner

$$\hat{x}_i^{(l)} = \mathrm{Up}\left(\hat{x}_i^{(l+1)}\right) + x_i^{(l)}, \tag{16}$$

where $\mathrm{Up}(\cdot)$ denotes bilinear upsampling. The final reconstructed image slice is obtained as

$$\tilde{I}_i = \hat{x}_i^{(0)}. \tag{17}$$

To ensure that the restoration does not alter the visible regions, the output image is composed under mask constraints

$$\hat{I}_i = I_i \odot (1 - M_i) + \tilde{I}_i \odot M_i. \tag{18}$$

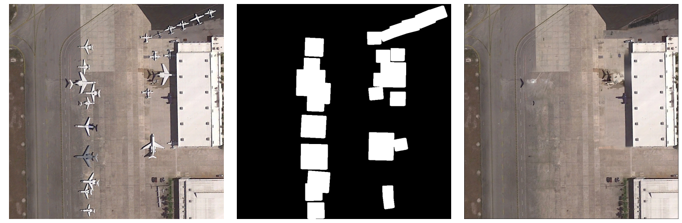
At this point, the satellite $s_{p,i}$ has completed the restoration of the privacy data in satellite image I_i and transmitted the image $\hat{I}_i$ to satellite $s_{p,v}$ for image aggregation.

To ensure visual continuity between adjacent repaired image slices, a boundary-smoothing operation is performed after all slices have been independently restored. For each repaired slice $\hat{I}_i$ with spatial range (y_1^i, y_2^i), a position-dependent weight map W_i is assigned along its boundaries to gradually reduce the intensity difference between neighboring slices. The weights follow a linear transition from 1 to 0 across a narrow boundary zone, enabling smooth blending at the interfaces. The final reconstructed image I_{final} is obtained through weighted accumulation and normalization

$$I_{\text{final}} = \frac{\sum_{i=1}^{N} W_i \odot \hat{I}_i}{\sum_{i=1}^{N} W_i}. \tag{19}$$

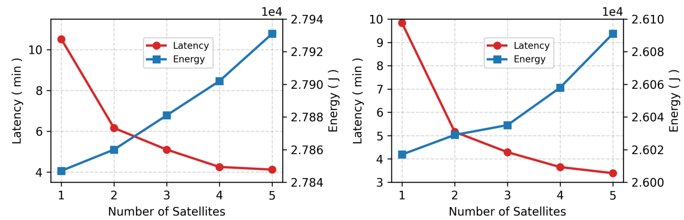
This process effectively eliminates visible seams or brightness inconsistencies along the borders of adjacent repaired regions. At this stage, the detection and restoration of privacy data within the GTF are completed.

In our framework, the detection and inpainting models are trained offline using remote sensing datasets and deployed on orbit for inference only. The inpainting network is optimized using a combination of pixel-level reconstruction loss and adversarial loss to balance visual fidelity and computational efficiency.



(a) Original image (b) Binary mask (c) Protected image

Fig. 3: Visual example of distributed privacy-preserving inpainting.



(a) DOTA (b) DIOR

Fig. 4: The performance of *DOPP* under different numbers of participating satellites.

V. Performance Analysis

A. Simulation Settings and Baselines

We consider an orbital plane p with S_p = 20 satellites and an orbital altitude h_s of 550 km. Therefore, the orbital period T_p of a satellite is about 95 minutes, and the window for communication with the GS within one cycle is about 9 minutes. Moreover, we set the maximum communication power as 5 W, the maximum computing frequency as 5 GHz, the link bandwidth as 20 MHz, $\gamma_{uv} = 10^3$, $\epsilon_0 = 10^{-28}$, and $\phi = 10^5$ cycles/sample [15]. YOLOv9 is initialized with COCO-pretrained weights and fine-tuned on the target dataset with a batch size of 16 and an input resolution of 640×640. MI-GAN is fine-tuned using detection-guided masks generated by YOLOv9 with a batch size of 8.

We evaluate our method on two datasets: **DOTA [16]** is a large-scale high-resolution optical remote sensing dataset for object detection, containing oriented bounding box annotations for multi-class objects with arbitrary orientations. **DIOR [17]** is a large-scale optical remote sensing dataset covering diverse scenes and imaging conditions for object detection.

All simulations are implemented in Python 3.9.23 using PyTorch 2.7.1. The fine-tuning process is conducted on a GPU server equipped with an NVIDIA GeForce RTX 4090 with 24 GB memory, while inference is performed on a Raspberry Pi device. To evaluate the effectiveness of our proposed ***DOPP***, we compare it with the following scenarios:

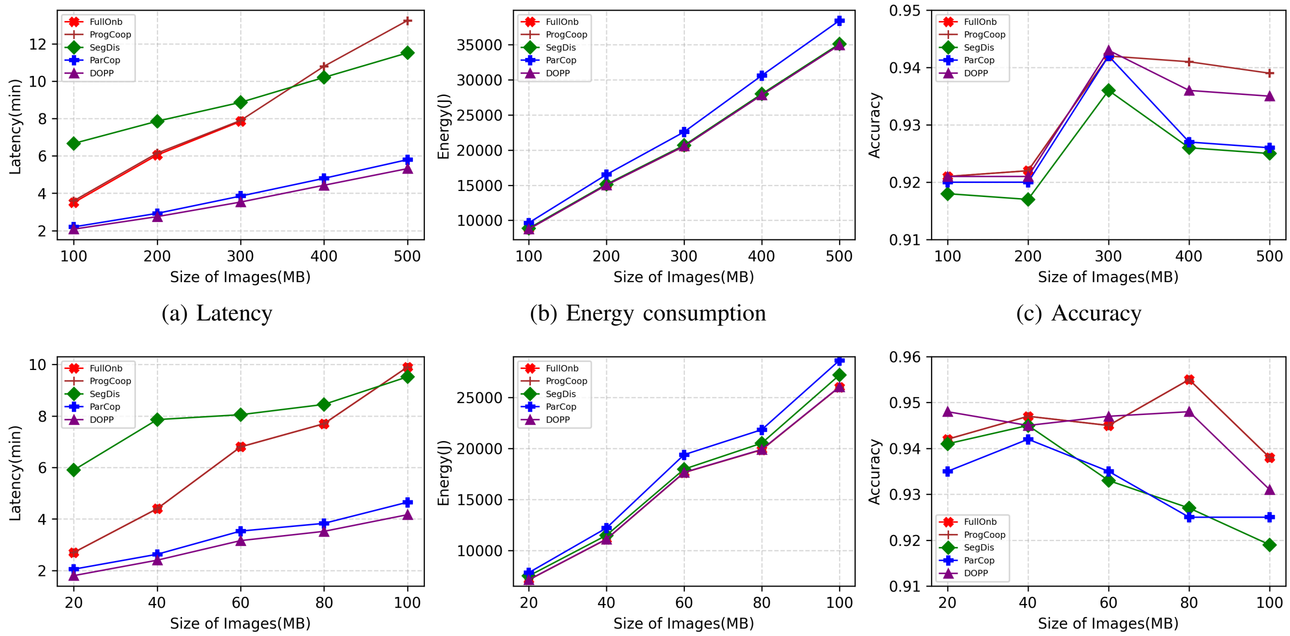
- **Full Onboard (FullOnb)**: After capturing a GTF, the satellite detects and repairs the privacy data of the entire GTF and transmits it directly to the GS in the next communication window.
- **Progressive Cooperative (ProgCoop)**: The satellite detects and inpaints the privacy data of the entire GTF. If the task cannot be completed within the communication window, intermediate results are sent to subsequent satellites, which continue processing and transmit the final results to the GS.
- **Segmented Distributed (SegDis)** [18]: The satellite captures each GTF but processes only part of it. Subsequent satellites complete the processing when they capture the GTF.
- **Partitioned Cooperative (ParCop)**: The satellite slices the GTF after capture, keeping one slice and transmitting the rest to other satellites. To handle privacy objects near slice boundaries, we set an overlap between adjacent slices, with a width of 1/10 of the original image.

In addition to system-level metrics (latency, energy, EDP), we also evaluate the visual effectiveness of privacy protection (See Fig. 3) to verify that sensitive objects are realistically concealed without introducing visible artifacts. Specifically, we adopt accuracy as a task-oriented metric, defined as the proportion of privacy-sensitive objects whose entire spatial regions are correctly covered by the generated masks and subsequently reconstructed through the inpainting process.

B. Simulation Results and Analysis

Fig. 4 illustrates the performance of *DOPP* under different numbers of participating satellites. In both datasets, increasing the number of satellites N initially reduces latency due to finer-grained parallelism, while energy consumption rises with additional communication overhead. *DOPP* achieves the best trade-off between latency and energy consumption when N=3.

Table I presents the sensitivity analysis of the dilation parameters on the DOTA and DIOR datasets, where the kernel size k and dilation iteration number t jointly control the



(a) Latency (b) Energy consumption (c) Accuracy

(d) Latency (e) Energy consumption (f) Accuracy

Fig. 5: Performance for different method under two datasets, (a)-(c): DOTA; (d)-(f): DIOR.

TABLE I: Parameter sensitivity analysis on different datasets. (The best result is bolded, and the second best is underlined)

Dataset	Parameters: Kernel size k	Parameters: Dilation iterations t	Metrics: Latency (min)	Metrics: Energy (J)	Metrics: Accuracy
DOTA	5	1	**4.22**	**27827**	92.5%
	5	3	4.39	27849	93.6%
	10	1	4.32	27831	93.2%
	10	3	4.43	27863	**94.1%**
DIOR	5	1	**4.02**	**25997**	89.7%
	5	3	4.13	26019	93.5%
	10	1	4.12	26012	92.1%
	10	3	4.17	26035	**95.5%**

TABLE II: Performance comparisons on different datasets. (The best result is bolded, and the second best is underlined)

Dataset	Methods	Latency (min)	Energy consumption (J)	EDP	Accuracy
DOTA	FullOnb	95.1	**27835**	N/A	**94.1%**
	ProgCoop	10.8	27845	1	**94.1%**
	SegDis	10.21	28002	0.951	92.6%
	ParCop	4.8	30624	0.489	92.7%
	DOPP (N=2)	5.7	27851	0.528	93.4%
	DOPP (N=3)	**4.43**	27863	**0.411**	93.6%
DIOR	FullOnb	95.1	**26014**	N/A	**95.5%**
	ProgCoop	9.9	26024	1	**95.5%**
	SegDis	9.52	27193	1	92.7%
	ParCop	4.65	28630	0.517	92.5%
	DOPP (N=2)	5.67	26057	0.573	94.8%
	DOPP (N=3)	**4.17**	26035	**0.421**	94.8%

spatial expansion of the detection masks. Overall, increasing either k or t consistently improves the inpainting accuracy on both datasets, indicating that a larger and smoother mask can more effectively cover object boundaries. However, this improvement in accuracy is accompanied by a slight increase in processing latency and energy consumption due to the enlarged inpainting region. The configuration with (k=10, t=3) achieves the highest inpainting accuracy on both the DOTA and DIOR datasets, reaching 94.1% and 95.5%, respectively. Compared to lighter configurations, such as (k=5, t=1), this setting incurs only a marginal increase in latency (less than 5%) and a negligible rise in energy consumption (under 0.15%), while delivering a substantial accuracy gain of 1.5 to 6.0%. These results demonstrate that the additional computational overhead associated with stronger dilation remains well within acceptable limits. In contrast, configurations with smaller kernel sizes or fewer dilation iterations achieve lower latency and reduced energy consumption; however, they often produce incomplete object boundaries in the mask. This compromises the effectiveness of the inpainting process and ultimately limits achievable accuracy. Based on these results, the configuration (k=10, t=3) is selected as the default for *DOPP*, as it achieves the optimal trade-off between accuracy, latency, and energy efficiency across all evaluated datasets.

Table II compares different methods in terms of processing time, energy consumption, EDP, and accuracy, where accuracy denotes the proportion of correctly repaired privacy targets and EDP is normalized to ProgCoop. On both datasets, *DOPP* achieves fast and energy-efficient privacy protection through non-overlapping image slicing, where only binary masking introduces minor additional computation. Since target detection takes much less time than image restoration, it achieves efficient inter-satellite parallelism and a trade-off between time and energy consumption, resulting in the best EDP results. For example, for the DOTA dataset, when N=3, the latency is significantly reduced with only a 12 J increase in energy consumption, and its EDP is lower; thus, N=3 is adopted in subsequent comparisons. Although ParCop also supports fast privacy protection, overlapping slices cause extra computation and transmission, resulting in much higher energy consumption. SegDis employs a pipelined strategy that avoids slice transmission between satellites, but its performance is limited by inter-satellite distances and waiting for subsequent image captures, which reduces parallel efficiency. Moreover, it neglects privacy continuity across adjacent slices, lowering restoration accuracy. ProgCoop forwards intermediate results to mitigate delays but remains serial overall, causing high latency. Finally, FullOnb performing all processing on a single satellite, suffers from long turnaround times and orbital delays, making it unsuitable for real-time usage.

Next, we evaluate the performance of different methods under varying data volumes. Figs. 5a to 5c show that *DOPP* (N=3) achieves the lowest latency with minimal energy overhead, benefiting from efficient inter-satellite parallelism and lightweight mask transmission. ParCop incurs higher latency and energy due to overlapping slices. When the data volume is less than 300 MB, a single satellite can complete processing within one communication window, allowing ProgCoop to operate without forwarding data to subsequent satellites. As the data volume increases, a single satellite can no longer complete the task within one window, forcing FullOnb to reschedule transmission to the next orbital cycle, adding about 95 minutes of delay per cycle (not shown in Fig. 5a). SegDis performs worse under small data volumes due to limited parallel efficiency. As shown in Fig. 5c, the accuracy of FullOnb and ProgCoop is slightly higher than that of other methods. This is because the accuracy of other methods may introduce minor accuracy loss at slice boundaries. However, the overall accuracy difference is less than 1%. Similar trends are observed on the DIOR dataset (Figs. 5d–5f) , demonstrating that DOPP maintains superior efficiency and scalability across different data sizes.

In summary, *DOPP* achieves efficient and scalable privacy-preserving satellite image processing by exploiting inter-satellite parallelism, delivering low latency and energy–delay product with negligible accuracy loss.

VI. Conclusion

This paper presents a distributed on-orbit framework for privacy-preserving neural processing of satellite imagery under strict computational and energy constraints. By enabling privacy detection and restoration directly on orbit, *DOPP* reduces the risk of sensitive data exposure during transmission while satisfying real-time requirements of IoRT observation missions. *DOPP* combines a quantized YOLOv9-based detection model with a lightweight Mi-GAN inpainting network fine-tuned for remote sensing imagery, achieving an effective trade-off between inference accuracy, latency, and energy–delay product. Through mask-based image slicing and inter-satellite parallel processing, the proposed framework significantly reduces per-satellite workload, while a visibility-aware aggregation strategy ensures timely result fusion and downlink transmission within limited satellite–ground contact windows. Simulation results on real-world datasets demonstrate that *DOPP* consistently outperforms existing onboard cooperative baselines in processing latency and energy consumption while maintaining high privacy restoration accuracy. Our future work will address the challenges of heterogeneous satellite constellations and intermittent connectivity by developing predictive link management and resource-aware task orchestration.

References

[1] Z. Shen *et al.*, "A survey of next-generation computing technologies in space-air-ground integrated networks," *ACM Computing Surveys*, vol. 56, no. 1, 2023.

[2] R. McAmis, M. Sim, M. Bennett, and T. Kohno, "Over fences and into yards: Privacy threats and concerns of commercial satellites," *Proceedings on Privacy Enhancing Technologies Symposium (PETS)*, pp. 379–396, 2024.

[3] Z. Bao *et al.*, "Blockchain-based secure communication for space information networks," *IEEE Network*, vol. 35, no. 4, pp. 50–57, 2021.

[4] Z. Liu *et al.*, "FedHC: A hierarchical clustered federated learning framework for satellite networks," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, pp. 6232–6237, 2025.

[5] T. Qiu *et al.*, "Techniques for the automatic detection and hiding of sensitive targets in emergency mapping based on remote sensing data," *ISPRS International Journal of Geo-Information*, vol. 10, no. 2, p. 68, 2021.

[6] Z. Yao *et al.*, "PPUP-GAN: A GAN-based privacy-protecting method for aerial photography," *Future Generation Computer Systems*, vol. 145, pp. 284–292, 2023.

[7] C.-Y. Wang, I.-H. Yeh, and H.-Y. Mark Liao, "YOLOv9: Learning what you want to learn using programmable gradient information," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 1–21, 2024.

[8] A. Sargsyan, S. Navasardyan, X. Xu, and H. Shi, "Mi-GAN: A simple baseline for image inpainting on mobile devices," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 7335–7345, 2023.

[9] C. Lin, D. He, N. Kumar, K.-K. R. Choo, A. Vinel, and X. Huang, "Security and privacy for the internet of drones: Challenges and solutions," *IEEE Communications Magazine*, vol. 56, no. 1, pp. 64–69, 2018.

[10] D. Zhang, L. Ren, M. Shafiq, and Z. Gu, "A lightweight privacy-preserving system for the security of remote sensing images on IoT," *Remote Sensing*, vol. 14, no. 24, p. 6371, 2022.

[11] Y. Zhang, J. Ji, W. Wen, Y. Zhu, Z. Xia, and J. Weng, "Understanding visual privacy protection: A generalized framework with an instance on facial privacy," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 5046–5059, 2024.

[12] E. Upenik, P. Akyazi, M. Tuzmen, and T. Ebrahimi, "Inpainting in omnidirectional images for privacy protection," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2487–2491, 2019.

[13] N. Razmi, B. Matthiesen, A. Dekorsy, and P. Popovski, "On-board federated learning for dense LEO constellations," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 4715–4720, 2022.

[14] T. Karimi and A. Kamran, "Energy-delay efficient segmented approximate adder with smart chaining," *IEEE Transactions on Computers*, vol. 74, no. 2, pp. 597–608, 2025.

[15] Z. Zhai *et al.*, "FedLEO: An offloading-assisted decentralized federated learning framework for low earth orbit satellite networks," *IEEE Transactions on Mobile Computing*, vol. 23, no. 5, pp. 5260–5279, 2023.

[16] G.-S. Xia *et al.*, "DOTA: A large-scale dataset for object detection in aerial images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3974–3983, 2018.

[17] K. Li *et al.*, "Object detection in optical remote sensing images: A survey and a new benchmark," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 296–307, 2020.

[18] B. Denby and B. Lucia, "Orbital edge computing: Nanosatellite constellations as a new class of computer system," in *Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, pp. 939–954, 2020.