Semantic Communication in Earth Observation via Cognitive DT-JSCC Framework for Enhanced Inference

Hong-fu Chou, Vu Nguyen Ha, Prabhu Thiruvasagam, Thanh-Dung Le, Geoffrey Eappen, Ti Ti Nguyen, Duc Dung Tran, †Sean Longyu Ma, Arsham Mostaani, Symeon Chatzinotas
Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg
†School of Computer Science, The University of Auckland, New Zealand

***Abstract*—Earth Observation (EO) systems play a critical role in applications such as mapping, disaster monitoring, and precision agriculture. However, they face ongoing challenges in processing and transmitting large volumes of data efficiently, particularly in time-sensitive scenarios such as real-time disaster response. This paper introduces a novel Cognitive Semantic Augmentation (CSA) framework that addresses these challenges by enabling the selective transmission of task-relevant semantic features instead of raw data. Using the inherent semantic structure of the EO imagery, CSA improves the accuracy of inference at the receiver and facilitates more efficient, purpose-driven communication across the satellite network. The proposed system also integrates inter-satellite communication and on-board deep learning models to enable a fully end-to-end, semantically enhanced communication pipeline. This architecture streamlines the transmission, processing, and analysis of multispectral EO data, resulting in improved object classification, real-time pattern recognition, and faster, more informed decision making. Performance evaluations demonstrate that the CSA Discrete Task-Oriented Joint Source-Channel Coding (DT-JSCC) system significantly outperforms both non-CSA baselines and conventional federated learning methods. It achieves higher Top-1 accuracy and greater robustness, driven by its semantic-aware signal encoding and optimized end-to-end transmission design.**

I. Introduction

The rapid proliferation of satellite and remote sensing data presents significant challenges in data processing and interpretation, despite its critical importance for applications such as environmental monitoring and disaster response. A notable solution to these challenges is the CORSA system [1], which introduces an AI-enabled compression method that maintains data fidelity while producing compact and efficient representations. CORSA effectively compresses Earth Observation (EO) data, such as water detection output, without compromising accuracy, demonstrating its suitability for onboard satellite processing. This approach offers a robust and scalable method for managing the increasing volume of EO data.

Another key advancement in EO data processing is the SAMRS framework [2], which integrates object detection techniques with the Segment Anything Model (SAM) to achieve high-precision semantic segmentation and labeling. Using a large data set of more than 100,000 annotated EO images, SAMRS provides a flexible solution for semantic segmentation in diverse urban environments.

Further progress in EO data analysis is exemplified by the incorporation of semantic knowledge into trajectory simplification. The method proposed in [3] introduces a semantically aware trajectory simplification technique that distinguishes between stationary and movement phases within spatiotemporal trajectories. This segmentation leads to more interpretable and informative representations of movement patterns, which enhances EO-based trajectory analysis.

Table I: Summary of Methodologies and Applications in Semantic Image Processing and Compression

Feature	[1]	[2]	[3]	[4]	[5]	Proposed
EO data processing	✓	✓	✓		✓	✓
AI-based compression	✓		✓	✓		✓
Semantic labeling		✓			✓	
Remote sensing		✓		✓	✓	✓
On-board satellite processing	✓		✓	✓		✓
Multi-spectral Analysis	✓	✓				

The HighDAN framework, as presented in [4], extends semantic compression capabilities to complex urban EO imagery. Using deep learning, HighDAN excels in semantic interpretation in varied urban contexts and demonstrates strong generalization capabilities across different geographic regions, making it a valuable tool for urban geospatial analysis.

An emerging frontier in EO data analysis is the application of deep learning to three-dimensional (3D) semantic segmentation, as explored in [5]. This study applies advanced deep learning models to 3D meshes, enabling an accurate classification of urban structures such as buildings and roads. Moving beyond traditional 2D imagery, this work significantly expands the scope and utility of semantic segmentation in EO data.

Table I highlights key advancements in semantic image processing, compression, and trajectory fusion for EO applications. Innovations such as CORSA, HighDAN, SAMRS, and semantic EO data cubes demonstrate the growing impact of semantic technologies in managing large-scale EO data. These approaches improve data integration, retrieval, and analysis across domains like urban planning, autonomous navigation, and surveillance. However, challenges remain, particularly regarding data heterogeneity and limited training resources, underscoring the need for continued research to enhance model robustness and generalizability.

We summarize our contributions as follows.

- The cognitive semantic EO system model is proposed to enhance satellite networks by integrating DT-JSCC [6] and Semantic Data Augmentation (SA) [7] techniques. The proposed cognitive semantic augmentation (CSA) technique enables the transmission of task-relevant semantic features, rather than raw data, and leverages the semantic structure of EO data to improve the inference accuracy at the receiver side.
- The framework integrates inter-satellite communication and deep learning models for on-board processing, enabling an end-to-end semantic-augmented communication process. This optimizes the transmission, analysis, and decision-making of multispectral EO data, resulting in improved classification accuracy and overall system efficiency.
- The CSA-enhanced DT-JSCC system significantly outperforms non-CSA and federated learning approaches by achieving higher Top1 accuracy and improved robustness in noisy LEO Rician channels through semantic-aware signal optimization and efficient end-to-end transmission.

The subsequent sections of this paper are structured as follows. Section II provides preliminary concepts and background on semantic joint source-channel coding, laying the foundation for the proposed framework. Section III discusses on-board semantic feature processing and explores the use of deep learning techniques to enhance resource efficiency in EO systems. Section IV presents numerical results that demonstrate the effectiveness and performance of the proposed methods. Section V concludes with a summary of the key findings.

II. Preliminary of Semantic joint source-channel coding

The integration of Joint Source-Channel Coding (JSCC) with deep learning has demonstrated superior performance over traditional methods, particularly in terms of reduced word error rate. This advantage becomes especially evident under constraints on computational resources allocated for syllable-level encoding. However, a key limitation of conventional JSCC approaches lies in the use of fixed-length bit representations to encode variable-length phrases, which leads to inefficiencies in both compression and transmission.

The objective of Deep Learning-based Joint Source-Channel Coding (DJSCC) is to ensure reliable, high-performance communication even under resource-limited conditions and low Signal-to-Noise Ratios (SNRs). To address the challenges of integrating DJSCC into digital communication frameworks, two innovative strategies are proposed in the Digital Task-Oriented JSCC (DT-JSCC) system [6]. The first approach, referred to as uniform modulation, dynamically adjusts the spacing of modulation constellations during transmission. This technique adapts to the statistical properties of the image data, functioning analogously to a flexible spring, and thus improves the efficiency of semantic feature transmission.

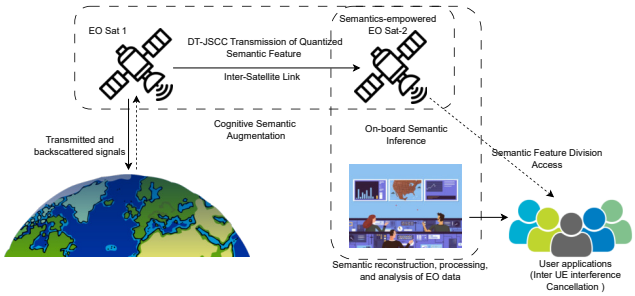

Figure 1: The proposed CSA LEO satellites for enhanced EO systems.

In addition, the study in [9] explores the use of Unequal Error Protection (UEP) in conjunction with semantic encoders/decoders within a JSCC architecture. This approach enhances the preservation of semantically critical information by allocating greater protection to more important data segments. The system utilizes QC-LDPC codes implemented on reconfigurable hardware, simplifying the JSCC architecture while enabling robust semantic communication over noisy channels. A fixed-point implementation prototype further demonstrates its practical feasibility.

III. Cognitive Semantic EO system model

This section presents a cognitive semantic Earth Observation (EO) system model designed to enhance satellite communication networks by incorporating advanced semantic processing methods, including DT-JSCC and semantic augmentation (SA) techniques. The proposed model leverages inter-satellite communication and cognitive processing to optimize the transmission and interpretation of multispectral EO data. By transmitting only essential, task-relevant semantic features instead of raw imagery, the system significantly reduces communication overhead and power consumption while enhancing decision-making efficiency in next-generation (6G) satellite networks.

The framework supports an end-to-end semantic-aware communication process that ensures efficient data handling and an improved user experience. Furthermore, the deployment of deep learning models for on-board processing is evaluated in terms of complexity and efficiency, with considerations for model size, parameter count, and training/inference latency. As depicted in Fig. 1, the satellite system involves a Low Earth Orbit (LEO) satellite that communicates with a ground-based gateway terminal connected to a server. Additionally, the satellite is part of a broader network of inter-satellite links, enabling real-time data exchange among orbiting satellites. These inter-satellite communications improve data transmission reliability and extend communication coverage beyond the line-of-sight limits of ground terminals.

The system processes multispectral satellite imagery captured at different temporal intervals using a two-satellite configuration. Satellite 1, located near the ground terminal,

acquires a reference image at time t_0, denoted as $U(t_0)$. This image is defined as:

$$U(t_0) = \{u^{t_0}(i,j,k) \mid 1 \leq i \leq H, 1 \leq j \leq W, 1 \leq k \leq D\}, \tag{1}$$

where H is the height, W is the width, and D is the number of spectral bands.

A. Semantic data transmission

Following cognitive processing of the reference image $U(t_0)$, the on-board classification network f_{s1} on Satellite 1 (Sat 1) extracts semantic features, denoted as $U_{\text{sem}}(t_0)$. This can be expressed as:

$$U_{sem}(t_0) = f_{s1}(U(t_0)), \tag{2}$$

where $U_{sem}(t_0) = \{u^{t_0}(i,j,k) \mid 1 \leq i \leq B, 1 \leq j \leq C, 1 \leq k \leq A\}$. We define B as the batch size, C as the number of classes, and A as the feature dimension. This information is then sent across an inter-satellite communication connection to the second satellite, which is located at a further distance and serves as a relay. This semantic representation is transmitted via an inter-satellite communication link to Satellite 2 (Sat 2), which functions as a relay node positioned further in orbit. The transmission is carried out using the DT-JSCC framework [6], which applies quantization and incorporates error correction to enhance reliability over noisy channels.

As depicted in Fig. 2, at time t_1, Sat 2 receives the encoded semantic data and applies its local semantic extractor f_{s2} to align this information with its own captured image $U(t_1)$. This enables Sat 2 to interpret $U(t_1)$ more effectively by leveraging prior semantic context from Sat 1. On-board training of f_{s2} can be used to improve the accuracy of semantic feature extraction based on received guidance. The semantic communication channel is further refined through dynamic adjustment of learning parameters via a covariance prediction network, as will be discussed in subsequent sections.

B. Cognitive semantic augmentation

Sat 2 utilizes the previously received semantic information $U_{\text{sem}}(t_0)$ to enhance feature extraction from the current image $U(t_1)$. The goal of semantic cognitive augmentation is to prioritize task-relevant and semantically invariant features, enabling effective detection of significant changes and relationships between $U(t_0)$ and $U(t_1)$. This process employs the Semantic Augmentation (SA) method introduced in [7].

As illustrated in Fig. 2, the covariance prediction network g_{s2} is trained to minimize the semantic augmentation loss:

$$L_{SA}(f_{s2}; g_{s2} | U_{sem}(t_0)). \tag{3}$$

where g_{s2} is optimized by backpropagating the gradient $\theta_{g_{s2}}$ derived from the cross-entropy loss of the DT-JSCC decoder output l_{s2}.

Processing data in the semantic feature domain $U_{\text{sem}}(t)$—as opposed to the multispectral pixel domain $MP(t)$—demonstrates the potential for efficient cognitive information exchange among neighboring satellites. The semantic cognitive covariance matrix is computed as:

$$\upsilon_{t_1}^{g_{s2}} = g_{s2}(U_{sem}(t_0), \theta_{g_{s2}}), \tag{4}$$

which enhances the system's ability to generalize over temporal and spatial changes in EO data. This process aligns with the optimization objectives of the classification network and extends beyond specific neural architectures, offering a generalized strategy for 6G-enabled semantic communications.

Definition 1. ***Cognitive Semantic Augmentation*** *is defined as the enhancement of an intelligent agent's ability to interpret and process information by leveraging semantic knowledge. It enables prioritization of relevant features, maintains semantic consistency, and improves the detection of important changes and patterns, thereby enhancing decision-making and situational understanding in complex environments.*

C. End-to-end semantic-augmented communication

As depicted in Fig. 1, the end-to-end communication pipeline between EO Sat 1, EO Sat 2, and the User Terminal (UT) employs advanced cognitive and semantic processing techniques to optimize EO data transmission and analysis. EO Sat 1 first collects raw sensory data (e.g., optical, radar), which are compressed into semantically meaningful representations via the DT-JSCC framework. These semantic features—containing only the most relevant task-specific information—are transmitted to EO Sat 2 through an inter-satellite link, reducing both bandwidth usage and energy consumption.

Upon receiving the features, EO Sat 2 performs on-board semantic inference and augmentation to further refine and analyze the data. This enables advanced tasks such as object detection, pattern recognition, and change detection. The refined semantic features are then transmitted to the UT, providing actionable insights.

Before EO Sat 2 performs the downlink transmission, the UT also receives the initial semantic data $U_{\text{sem}}(t_0)$ from Sat 1 to apply SA techniques [7] and to assist in decoding via the DT-JSCC decoder l_{UT}. As a result, both the classification networks f_{s2} and l_{UT} can be updated and optimized independently, forming a unified semantic-enhanced DT-JSCC system for robust and efficient downlink communication under the CSA paradigm.

As illustrated in Fig.2, we summarize the proposed end-to-end CSA satellite communication in Algorithm 1. In particular, lines 4 and 5 carry out the CSA training of f_{s_2} on Sat2 side and l_{UT} on UT side simultaneously with lines 6 and 7. Following completion of the CSA training, the data are subjected to semantic reconstruction, processing, and analysis to provide useful insights. These insights are then sent to user applications via Semantic Feature Division Multiple Access (SFDMA) [10]. These user applications can use the semantically processed EO data for a variety of purposes, including inter-UE (User Equipment) interference cancellation, which improves communication efficiency and signal

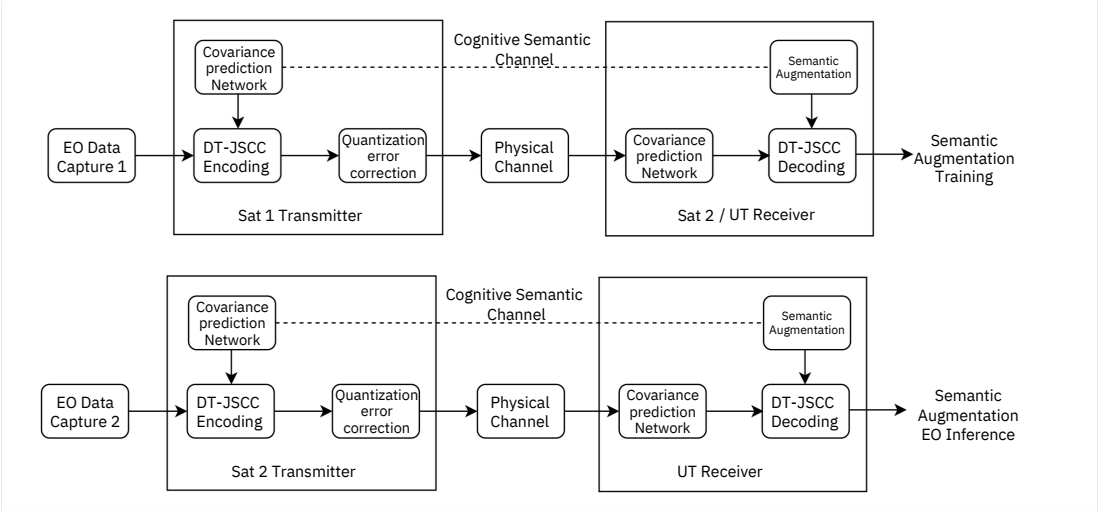

Figure 2: The proposed transmitter and receiver for cognitive semantic satellite networks.

Algorithm 1 End-to-End CSA Satellite Communication

Input: EO images $U(t_i)$ from Sat1 and Sat2 with size $H \times W \times D$; semantic parameters: batch size B, class count C, feature dim. A
Output: EO inference at UT

1: **for** $i = 0, 1, \ldots, n-1$ **do**
2: **Sat1:** Encode and transmit semantic features:

$$U_{\text{sem}}(t_i) = f_{s_1}(U(t_i))$$

3: **Sat2, UT:** Receive $U_{\text{sem}}(t_i)$, process:

$$U_{\text{sem}}(t_{i+1}) = f_{s_2}(U(t_{i+1}))$$

4: **Meta-learning at Sat2:** Minimize $\mathcal{L}_{SA}(f_{s_2}; g_{s_2}|U_{\text{sem}}(t_i))$
5: Update g_{s_2} via backprop on $\theta_{g_{s_2}}$
6: **Meta-learning at UT (in parallel with line 4):** Minimize $\mathcal{L}_{SA}(f_{UT}; g_{UT}|U_{\text{sem}}(t_i))$
7: Update g_{UT} via backprop on $\theta_{g_{UT}}$
8: **Sat2:** Transmit $U_{\text{sem}}(t_{i+1})$ to UT for final inference
9: **end for**

clarity. The SFDMA network reduces interference through approximate orthogonal transmission and introduces the Alpha-Beta-Gamma formula for semantic communications, with simulations validating its effectiveness. This overall process demonstrates how the integration of semantic cognition into satellite systems can drastically improve the efficiency of EO data transmission, interpretation, and utilization, minimizing unnecessary data loads and focusing on task-relevant information.

IV. Numerical Results

A. Channel model

To simulate propagation and attenuation, the channel model incorporates small- and large-scale fading; the path loss components are the large-scale fading between the satellite and UT [11]:$P_{L_{\text{tot}}} = P_{L_b} + P_{L_g} + P_{L_s}$, in which the attenuation resulting from atmospheric gases is denoted by P_{L_g}, the attenuation resulting from ionospheric or tropospheric scintillation by P_{L_s}, and the basic path loss is represented by P_{L_b}. Furthermore, the ISL of the free space may be understood as $P_{L_{\text{ISL}}} = P_{L_{ib}}$, in which $P_{L_{ib}}$ denotes Sat1 and Sat2's fundamental path loss. All of these loss components on the route are expressed in decibels (dB). More specifically, free space propagation and shadow fading of the signal are taken into account using the fundamental path loss model P_{L_b}. Thereby, the loss of the free space path ($P_{L_{\text{FS}}}$) can be calculated as: $P_{L_{\text{FS}}}(d, f_c) = 32.45 + 20\log_{10}(f_c) + 20\log_{10}(d)$, Shadow fading (SF) is modeled by a log-normal distribution as $\mathcal{N}(0, \sigma_{SF}^2)$ with zero mean and σ_{SF} standard deviation. The values of σ_{SF}^2 can be extracted from 3GPP Release-15. The path loss with and without shadow fading in dB units is then modeled as $P_{L_b} = P_{L_{\text{FS}}}(d, f_c) + SF$ and $P_{L_{ib}} = P_{L_{\text{FS}}}(d_{ib}, f_c)$, where d_{ib} represent the distance of Sat1 and Sat2. We consider the Rician model to express the channel between the LEO satellite $s1$ and the UT z and ISL between the LEO satellite $s1$ and $s2$. The channel from the LEO satellite $s1$ to the UT z and ISL are given as

$$f_{s1,z} = \sqrt{\frac{R_z \zeta_{s1,z}}{R_z + 1}}\bar{f}_{s1,z} + \sqrt{\frac{\zeta_{s1,z}}{R_z + 1}}\tilde{f}_{s1,z}, \tag{5}$$

$$f_{s1,s2} = \sqrt{\frac{R_{s2} \zeta_{s1,s2}}{R_{s2} + 1}}\bar{f}_{s1,s2}, \tag{6}$$

where $\bar{f}_{s1,z}$, $\bar{f}_{s1,s2}$ and $\tilde{f}_{s1,z}$ are the line-of-sight (LoS) and non-line-of-sight (NLoS) paths components, respectively, R_z is the Rician factor, and the ISL does not have NLoS component. After accounting for the gain of the LEO satellite antenna G_T, the large-scale fading effects of the LEO satellite $s1$ on UT z and ISL can be modeled as follows: $\zeta_{s1,z}(\text{dB}) = PL_{\text{tot}} - G_T$ and $\zeta_{s1,s2}(\text{dB}) = PL_{\text{ISL}} - G_T$. Furthermore, the complex Gaussian distribution with zero mean and unit variance is followed by the independent and identically distributed (i.i.d.) entries of $\tilde{f}_{s1,z}$ data.

The channel of the complex baseband downlink space domain at a time instance t and frequency f can be expressed

using a ray-tracing-based channel modeling method. We have $H_{s1,z}(t,f) = f_{s1,z}\exp(2\pi[tv_z - f\tau_z])$ and $H_{s1,s2}(t,f) = f_{s1,s2}\exp(2\pi[tv_{s2} - f\tau_{s2}])$, where v_z, v_{s2} are the Doppler shift and τ_z, τ_{s2} are the fading gain distributed by the Rician propagation delay, accounting for scintillation and atmospheric losses summarized in [11].

B. Experimental Setup

The EuroSAT dataset is a publicly available remote sensing benchmark derived from Sentinel-2 satellite imagery. It consists of 27,000 labeled images across 10 land use and land cover classes, with each image having a size of 64×64 pixels and 13 spectral bands. This dataset is widely used for evaluating image classification and transmission systems due to its diversity, balanced classes, and relevance to Earth observation tasks. The EuroSAT dataset was split into training/validation/test sets with an 80/10/10 ratio. Transmission experiments were conducted under PSNR training value $\in \{8, 12, 16\}$ dB and testing values ranging from 2 dB to 20 dB using a 16-APSK modulation scheme. The fading parameter was varied with $K \in \{32, 64, 128\}$ to represent different channel conditions. Training used a batch size of 512 for 100 epochs with the Adam optimizer. The class count C and the feature dimension A are set to 10 and 64, respectively. The results were averaged over multiple runs to ensure stability and statistical reliability.

C. Physical layer transmission

Fig. 3 presents the Top-1 accuracy using DT-JSCC [6] under different channel conditions for EuroSAT data. This graph depicts performance across different PSNR (Peak Signal-to-Noise Ratio) levels for the 16APSK modulation scheme. We show accuracy for the AWGN (Additive White Gaussian Noise) and Rician channels, with different K-factor values (32, 64, 128), where a higher K-factor generally corresponds to better channel quality. Furthermore, we compared accuracy across the LEO Rician and LEO Rayleigh channels for certain K values, showing a higher accuracy in the Rician channel. Across all scenarios, the graphs indicate that increasing PSNR improves accuracy, with Rician channels generally performing better than Rayleigh, and higher K factors leading to more stable performance, especially in LEO environments.

D. CSA satellite network

Table II shows the CSA DT-JSCC system with high accuracy and evidently with most classes, such as annual crops (94.51%), herbaceous vegetation (91.43%) and sea / lake (98.85%), showing high Top1 accuracy, indicating correct classifications. Some misclassifications can be improved, such as between Industrial and Residential areas. Non-CSA shows slightly reduced accuracy, with some increased misclassifications, particularly in Herbaceous Vegetation (89.64%) and Sea/Lake (96.63%). In general, the CSA-enhanced system demonstrates a 16% improvement in accuracy and fewer misclassifications than the non-CSA system.

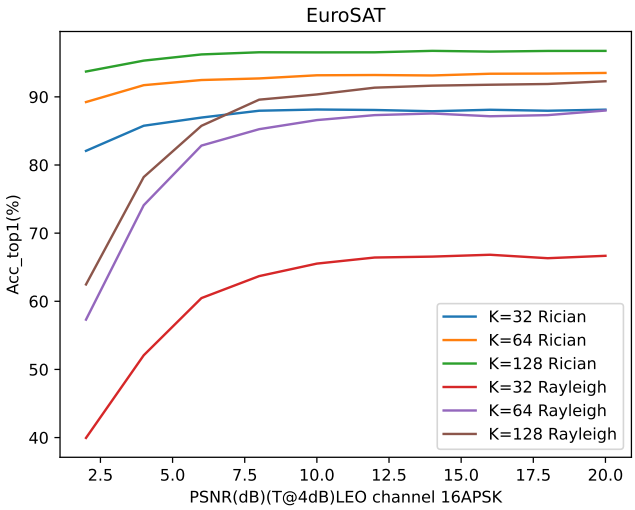

Figure 3: Top1 accuracy using DT-JSCC based on 16APSK LEO Rician and LEO Rayleigh channel.

As illustrated in Fig. 4, the SA technique struggles significantly under the conditions of a low PSNR value due to the inherent noise and interference present in LEO Rician channels. This adverse environment leads to diminished signal quality, which negatively impacts the performance of the SA method. In contrast, a notable improvement in Top1 accuracy is evident with K = 32 and K = 128 DT-JSCC transmissions. This improvement can be attributed to the end-to-end CSA architecture, which employs advanced processing techniques to optimize the semantic representation and transmission of signals, resulting in greater robustness under varying channel conditions. By intelligently aggregating and interpreting semantic features, the system is able to preserve critical information, thereby improving accuracy even in the presence of noise and channel distortion.

However, at low PSNR levels, where the signal is significantly degraded by noise and distortion, the accuracy of CSA-based solutions may decline. This is primarily due to the fact that semantic representations become increasingly vulnerable to corruption during transmission, especially when quantization noise and channel errors affect higher-level features more severely than raw pixel values. As a result, the semantic decoder may struggle to reconstruct meaningful content, leading to reduced classification or inference accuracy in these low-quality regimes. Unlike pixel-based compression techniques that can retain low-level textures despite distortion, semantic features are abstract and task-specific—thus more sensitive to information loss.

Although federated learning typically aggregates local updates from distributed clients on a central server without transmitting raw data, thereby preserving privacy, our proposed end-to-end CSA scenario offers a more communication-efficient and context-aware alternative. By directly analyzing the semantic structure of source data from a neighboring satellite, the CSA-based method demonstrates superior per-

formance. Specifically, using a 100-round communication process between the UT server and the client satellites, as outlined in [12], our system demonstrates improved adaptability and reduced communication overhead compared to a federated learning baseline operating under a fixed communication budget of B=512 bits across 100 rounds in space communications The results highlight the importance of

Table II: Top1 accuracy comparison of CSA and non-CSA DT-JSCC K=32 systems while PSNR=12dB and 16APSK over LEO Rician channel.

Class	Accuracy (%) CSA	Accuracy (%) non-CSA
AnnualCrop	94.50	88.38
Forest	98.41	98.84
HerbaceousVegetation	91.63	89.64
Highway	92.30	77.27
Industrial	95.95	90.95
Pasture	94.59	90.53
PermanentCrop	90.03	83.55
Residential	98.02	95.95
River	92.33	72.63
SeaLake	98.95	98.33

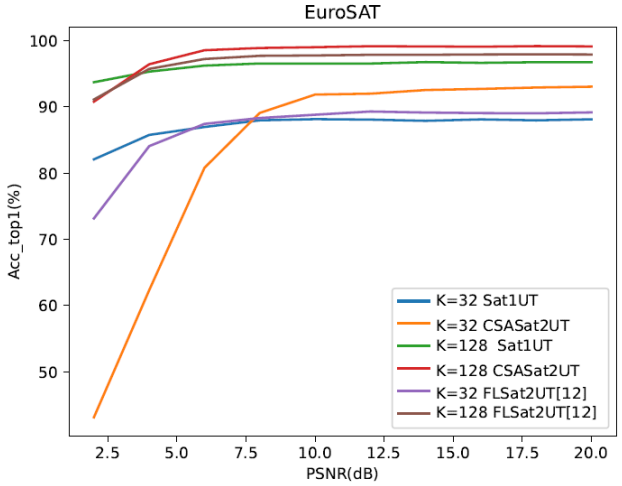

Figure 4: Top1 accuracy of CSA satellite networks using DT-JSCC over 16APSK LEO Rician channel while DT-JSCC training at 4dB and CSA/federated learning training at 12dB.

employing sophisticated transmission strategies, particularly in environments characterized by high levels of noise and interference. The substantial gains in accuracy achieved with the DT-JSCC and CSA approaches underscore the efficacy of incorporating semantic-aware mechanisms in the design of communication systems, suggesting that further exploration and development in this area could yield even more significant benefits in future applications.

V. Conclusion

This work presents an advanced framework for semantic communication networks, with a focus on the development and optimization of cognitive semantic EO systems. The integration of semantic data augmentation, intersatellite links, and cognitive processing capabilities ensures that only the most relevant, task-specific data is transmitted, significantly reducing communication overhead and improving system performance. Furthermore, by leveraging semantic inference and cognitive augmentation techniques, the system facilitates better image analysis, change detection, and decision-making in real-time applications. The proposed end-to-end model demonstrates how integrating semantic cognition into satellite systems can drastically improve data transmission, interpretation, and overall system efficiency, thereby addressing the demands of 6G networks and beyond. This research lays the foundation for future advancements in semantic communication, with promising applications in a wide range of fields, including remote sensing, autonomous systems, and smart communication networks.

Acknowledgment

This work was funded by the Luxembourg National Research Fund (FNR), with the SENTRY project granted corresponding to the grant reference C23/IS/18073708/SENTRY.

References

[1] X. Ivashkovych, L. Landuyt, and T. Van Achteren, "CORSA Deep Earth Observation Semantic Compression applied to Flood Detection," in *OBPDC Conference*, 2022.

[2] D. Wang, J. Zhang, B. Du, M. Xu, L. Liu, D. Tao, and L. Zhang, "Samrs: Scaling-up remote sensing segmentation dataset with segment anything model," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[3] M. Liu, G. He, and Y. Long, "A semantics-based trajectory segmentation simplification method," *Journal of Geovisualization and Spatial Analysis*, vol. 5, pp. 1–15, 2021.

[4] D. Hong, B. Zhang, H. Li, Y. Li, J. Yao, C. Li, M. Werner, J. Chanussot, A. Zipf, and X. X. Zhu, "Cross-city matters: A multimodal remote sensing benchmark dataset for cross-city semantic segmentation using high-resolution domain adaptation networks," *Remote Sensing of Environment*, vol. 299, p. 113856, 2023.

[5] G. Weixiao, L. Nan, B. Boom, and H. Ledoux, "Pssnet: Planarity-sensible semantic segmentation of large-scale urban meshes," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 196, pp. 32–44, 2023.

[6] S. Xie, S. Ma, M. Ding, Y. Shi, M. Tang, and Y. Wu, "Robust information bottleneck for task-oriented communication with digital modulation," *arXiv preprint arXiv:2209.10382*, 2023.

[7] Y. Pu, Y. Han, Y. Wang, J. Feng, C. Deng, and G. Huang, "Fine-grained recognition with learnable semantic data augmentation," *IEEE Transactions on Image Processing*, 2024.

[8] A. Mostaani, T. X. Vu, S. K. Sharma, V.-D. Nguyen, Q. Liao, and S. Chatzinotas, "Task-oriented communication design in cyber-physical systems: A survey on theory and applications," *IEEE Access*, vol. 10, pp. 133842–133868, 2022.

[9] X. Zhong, C.-W. Sham, S. L. Ma, H.-F. Chou, A. Mostaani, T. X. Vu, and S. Chatzinotas, "Joint Source-Channel Coding System for 6G Communication: Design, Prototype and Future Directions," *IEEE Access*, vol. 12, pp. 17708–17724, 2024.

[10] S. Ma, C. Zhang, B. Shen, Y. Wu, H. Li, S. Li, G. Shi, and N. Al-Dhahir, "Semantic feature division multiple access for multi-user digital interference networks," *IEEE Transactions on Wireless Communications*, 2024.

[11] A. B. Adam, M. Samy, C. E. García, E. Lagunas, and S. Chatzinotas, "Diffusion model-based signal recovery in coexisting satellite and terrestrial networks," in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2024, pp. 1–5.

[12] N. Razmi, B. Matthiesen, A. Dekorsy, and P. Popovski, "On-board federated learning for satellite clusters with inter-satellite links," *IEEE Transactions on Communications*, 2024.