QSat-SemCom: A Quantum-Enhanced Deep Learning Framework for Semantic Satellite Image Transmission

Khoi T. Nguyen[1], Trung Thanh Pham[1], Ji Su Yoon[2], and Choong Seon Hong[3]*
[1]Department of Artificial Intelligence, Kyung Hee University, Yongin, Gyeonggi-do 17104, Republic of Korea
[2] Department of Computer Engineering, Kyung Hee University, Yongin, Gyeonggi-do 17104, Republic of Korea
[3]School of Computing, Kyung Hee University, Yongin, Gyeonggi-do 17104, Republic of Korea
email:{khoi.nt, trungpt, yjs9512, cshong}@khu.ac.kr

***Abstract*—As sixth-generation (6G) wireless networks evolve, semantic communication (SemCom) improves transmission efficiency by delivering task-relevant information instead of raw bits. This paradigm is particularly attractive for satellite image transmission, where limited bandwidth and high reconstruction quality are critical. Deep Joint Source-Channel Coding (DeepJSCC) provides efficient compression and robust transmission but remains constrained by the representational capacity of classical neural architectures. To address this limitation, we propose QSAT-SEMCOM, a quantum-classical hybrid SemCom framework that integrates a variational quantum circuit into the semantic encoder of a DeepJSCC system to enhance semantic feature representation. Experimental results demonstrate competitive image reconstruction performance and show that quantum-enhanced feature learning can effectively improve semantic image transmission. These findings demonstrate the potential of integrating quantum computing into SemCom for future wireless communication systems.**

***Index Terms*—Semantic communication, hybrid quantum-classical, satellite image transmission.**

I. Introduction

The rapid development of sixth-generation (6G) wireless networks is driving communication systems beyond reliable bit transmission toward semantic information exchange. Unlike traditional communication frameworks that faithfully deliver every transmitted bit, semantic communication (SemCom) conveys only task-relevant information, improving transmission efficiency while preserving information utility. Consequently, SemCom has emerged as a promising technology for 6G networks, where massive multimedia traffic and spectrum constraints demand more efficient communication paradigms [1]–[3].

Among various SemCom applications, image transmission has attracted considerable attention due to the growth of visual data generated by intelligent sensing platforms, particularly in satellite and remote-sensing systems, where high-resolution imagery must be delivered through bandwidth-constrained wireless links. Deep Joint Source-Channel Coding (DeepJSCC) has become the dominant framework for image transmission by jointly optimizing image compression and channel coding, overcoming the limitations of conventional source-channel separation while exhibiting graceful performance degradation under adverse channel conditions [4]. Recent transformer-based approaches, including WITT [5], SwinJSCC [6], and DeepJSCC-MIMO [7], further improve semantic representation learning and transmission robustness.

Despite these advances, existing DeepJSCC systems remain constrained by the representational capacity of classical neural architectures. Under limited bandwidth, semantic encoders must compress rich visual content into compact latent representations while preserving reconstruction-relevant information. This challenge is particularly important for satellite image transmission, where communication resources and reconstruction fidelity are both critical.

Recent developments in quantum computing provide a promising avenue to address this limitation. By exploiting superposition and entanglement, hybrid quantum-classical models can encode information in high-dimensional Hilbert spaces. Such frameworks have demonstrated encouraging results in image classification and wireless communication optimization [8]–[10]. Emerging studies on quantum semantic communication (QSC) also suggest richer semantic embeddings and improved communication efficiency [2], [11]. However, integrating quantum feature learning into wireless semantic image reconstruction remains largely unexplored.

To bridge this gap, this paper proposes QSAT-SEMCOM, a quantum-classical hybrid semantic image transmission framework for satellite communication systems. The main contributions of this work are as follows:

- We propose QSAT-SEMCOM, a novel quantum-classical hybrid SemCom framework that integrates a VQC into a transformer-based DeepJSCC architecture, targeting satellite domain image transmission.

This work was supported by the Institute of Information Communications Technology Planning Evaluation (IITP) grant funded by the Korean government (MSIT) (No. 2019-0-01287, Evolvable Deep Learning Model Generation Platform for Edge Computing); the IITP–ITRC (Information Technology Research Center) grant funded by the Korean government (MSIT) (No. IITP-2026-RS-2023-00258649, 50%); the IITP grant funded by the Korean government (MSIT) (No. RS-2024-00509257, Global AI Frontier Lab); and the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. RS-2024-00352423). Dr. C. S. Hong is the corresponding author.

- We design a quantum-enhanced semantic encoder that maps image features into a quantum latent representation, improving feature expressiveness while maintaining compatibility with conventional wireless transmission.
- We evaluate the proposed framework under satellite wireless channel conditions using Peak Signal-to-Noise Ratio (PSNR) and Multi-Scale Structural Similarity (MS-SSIM), providing insights into quantum feature representations in SemCom and future opportunities enabled by advances in quantum hardware.

II. Related Work

SemCom has emerged as a key technology for next-generation wireless networks by transmitting task-relevant information rather than every transmitted bit. Building upon DeepJSCC [4], transformer-based frameworks such as WITT [5] and SwinJSCC [6] have improved semantic representation learning and transmission performance. SemCom has also attracted increasing interest in satellite image delivery and space-air-ground integrated networks, where bandwidth constraints make efficient image transmission essential [12], [13]. In parallel, hybrid quantum-classical learning has emerged as a promising approach for image representation learning. Sebastianelli *et al.* [14] and Fan *et al.* [8] demonstrated that integrating VQCs into deep neural networks can improve image classification and remote-sensing representation learning. These studies suggest that VQCs provide expressive feature transformations for visual tasks. The intersection of quantum computing and SemCom has only recently emerged. QSC [11] introduced quantum representations for semantic communication, highlighting the potential of quantum embeddings. However, integrating quantum learning modules into semantic image reconstruction remains largely unexplored.

Motivated by these observations, we integrate a VQC into the semantic encoding pipeline of WITT to investigate whether quantum feature representations improve semantic image transmission under bandwidth-constrained wireless channels. Unlike existing studies that focus on communication optimization, our work targets end-to-end satellite image transmission with reconstruction fidelity as the primary objective.

III. Methodology

A. Preliminary

Quantum computing and quantum circuit. Quantum computing operates on quantum states in a complex Hilbert space. Unlike classical systems, an n-qubit system is represented by a state vector in a 2^n-dimensional space, enabling superposition. A single qubit is expressed as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1}$$

where $\alpha, \beta \in \mathbb{C}$ satisfy $|\alpha|^2 + |\beta|^2 = 1$ [14]. Multi-qubit systems are represented using tensor products, and certain multi-qubit states exhibit entanglement, producing non-classical correlations between qubits [14].

Quantum gates. Quantum information is manipulated through unitary transformations satisfying

$$U^\dagger U = UU^\dagger = I. \tag{2}$$

Quantum gates include single-qubit and multi-qubit operations. Parameterized rotation gates are commonly used single-qubit gates. For example, the R_y rotation gate is defined as

$$R_y(\theta) = e^{-i\theta Y/2}, \tag{3}$$

where θ denotes the rotation angle. In contrast, controlled gates such as the controlled-NOT (CNOT) gate flip a target qubit based on the state of a control qubit, creating multi-qubit entanglement. Rotation and entangling gates together form expressive variational quantum circuits through trainable transformations and qubit interactions.

Quantum measurement. The output of a quantum system is obtained through measurement, collapsing the quantum state to a basis state according to its probability amplitudes [14].

B. System Model

As illustrated in Fig. 1, the proposed framework transmits an image over a noisy wireless channel. Let $\mathbf{x} \in \mathbb{R}^{H\times W\times C}$ denote the input image. At the transmitter, a semantic encoder consisting of a patch embedding layer followed by multiple Swin Transformer blocks extracts semantic features and compresses them into a latent representation:

$$\mathbf{s} = E_\theta(\mathbf{x}), \tag{4}$$

where $E_\theta(\cdot)$ denotes the encoder parameterized by θ, and $\mathbf{s}$ is the transmitted semantic feature vector. The latent representation is transmitted through an Additive White Gaussian Noise (AWGN) channel:

$$\hat{\mathbf{s}} = \mathbf{s} + \mathbf{n}, \tag{5}$$

where $\mathbf{n} \sim \mathcal{N}(0, \sigma^2)$ and $\sigma = \sqrt{\frac{1}{2\cdot 10^{\mathrm{SNR_{dB}}/10}}}$. At the receiver, a semantic decoder composed of hierarchical Swin Transformer layers reconstructs the image:

$$\hat{\mathbf{x}} = D_\phi(\hat{\mathbf{s}}), \tag{6}$$

where $D_\phi(\cdot)$ denotes the decoder parameterized by ϕ.

The encoder and decoder are optimized by minimizing the reconstruction loss between $\mathbf{x}$ and $\hat{\mathbf{x}}$, typically using the mean squared error (MSE). Reconstruction quality is evaluated using the PSNR for pixel-wise reconstruction accuracy and the MS-SSIM metric [15] for perceptual image quality.

C. Quantum-Enhanced MLP

In the proposed QSat-SemCom framework, the conventional multilayer perceptron (MLP) block within the Swin Transformer architecture, denoted as MLP* in Fig. 1, is replaced by a quantum circuit module. The transmitter-side MLP maps latent features to the target bandwidth ratio (CBR). This compression and representation-learning process aligns naturally with the expressive capacity of quantum circuits.

As shown in Fig. 2, the proposed quantum circuit consists of three stages: quantum state encoding, variational feature

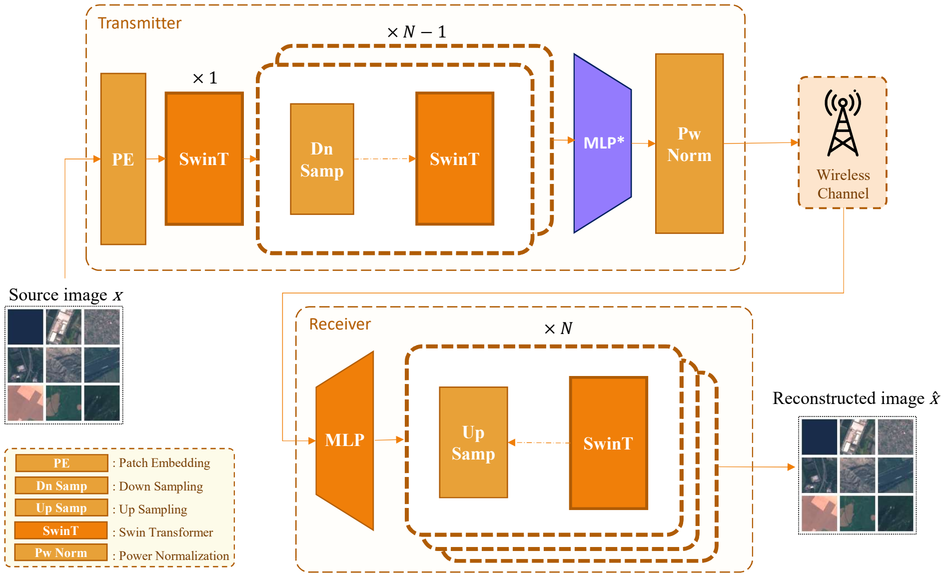

Fig. 1. Overview of the Proposed QSAT-SEMCOM Framework

transformation, and measurement. Given an input feature vector $\mathbf{z} = [z_1, \ldots, z_Q]$ from the preceding transformer layer, each feature is encoded into a quantum state through rotational gates acting on individual qubits, as defined in (3). Specifically, by setting the rotation angle of the R_y gate to $\theta = z_i$, each input feature is encoded onto a quantum wire through rotation about the y-axis, mapping classical feature values into quantum amplitudes.

The encoded quantum state is then processed by N variational entanglement layers. Each layer consists of trainable $R_y(\theta)$ gates followed by controlled-NOT gates connecting adjacent qubits. The rotation gates serve as trainable parameters, while the CNOT gates introduce entanglement between qubits. Through repeated parameterized rotations and entangling operations, the circuit learns expressive feature representations in a high-dimensional Hilbert space.

Finally, the transformed quantum state is projected back into the classical domain through a measurement layer. For each qubit, the expectation value of the Pauli-Z observable is measured as

$$m_i = \langle \psi | Z_i | \psi \rangle, \tag{7}$$

where $|\psi\rangle$ denotes the final quantum state and Z_i is the Pauli-Z operator acting on the i-th qubit. The resulting measurement vector $\mathbf{m}$ forms the output of the quantum-enhanced MLP and is processed by the remaining transformer components.

The proposed quantum-enhanced MLP preserves the original role of the dense network in adapting feature representations and controlling latent dimensionality. Meanwhile, the VQC provides an expressive feature transformation mechanism that exploits quantum superposition and entanglement to learn compact semantic representations for bandwidth-constrained image transmission.

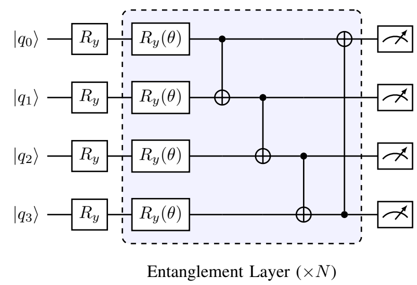

Fig. 2. Quantum Circuit Architecture for the MLP Replacement

IV. EXPERIMENTS

A. Settings

The proposed QSAT-SEMCOM framework is implemented by augmenting the WITT architecture [5] with a VQC module. For the backbone configuration, the encoder and decoder use embedding dimensions of $[16, 16]$, depths of $[2, 2]$, attention heads of $[4, 4]$, and an input size of $64 \times 64 \times 3$. The VQC employs two or four qubits with one, two, or four variational layers, resulting in six model variants. The framework is trained end-to-end using the Adam optimizer with a learning rate of 10^{-4} and the MSE reconstruction loss. Experiments are conducted on the EuroSAT dataset [16], which contains approximately 27,000 Sentinel-2 satellite images. Following stratified sampling to accommodate the compact model size and reduce training complexity, 20% of the dataset is used for training and 2% for testing. During training, latent features are transmitted through an AWGN channel with the SNR randomly sampled from $\{1, 3, 7, 10, 13\}$ dB for each mini-batch to improve robustness across channel conditions. Reconstruction performance is evaluated using PSNR and MS-SSIM.

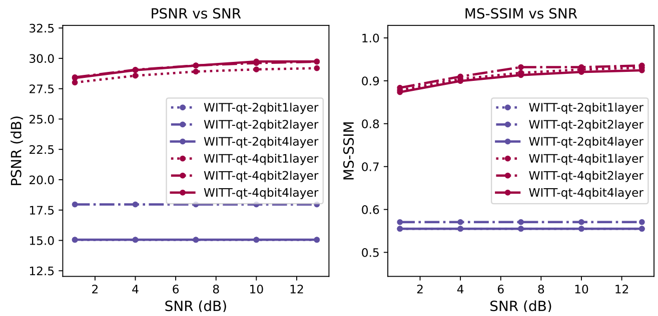

Fig. 3. Image reconstruction performance under AWGN channel conditions

TABLE I
MODEL PARAMETERS AND PERFORMANCE AT 7 dB SNR

Num qubits	Num layers	Num params	PSNR	MS-SSIM
2	1	47400	15.05	0.555
2	2	47406	17.96	0.571
2	4	47418	15.05	0.555
4	1	47472	28.91	0.919
4	2	47484	**29.42**	**0.932**
4	4	47508	29.40	0.913

Bold indicates the best candidate for each metric.

B. Results

Fig. 3 and Table I present the reconstruction performance and model complexity of the proposed quantum-enhanced SemCom framework under different quantum circuit configurations. Fig. 3 shows that reconstruction quality improves with increasing SNR for all models, demonstrating robust transmission across varying channel conditions. Across the SNR range, the 4-qubit models consistently outperform their 2-qubit counterparts in both PSNR and MS-SSIM, indicating that increasing the number of qubits enhances semantic feature representation.

Table I confirms these observations. At an SNR of 7 dB, the best-performing 4-qubit model achieves 29.42 dB PSNR and 0.932 MS-SSIM, compared with 17.96 dB and 0.571 for the best-performing 2-qubit model. These gains are obtained with only 72–102 additional trainable parameters, corresponding to less than 0.25% of the total model size, demonstrating the parameter efficiency of the proposed framework.

The results also show that circuit depth should be considered together with the number of qubits. Increasing the depth from one to two layers improves reconstruction quality for both qubit settings, whereas further increasing the depth to four layers provides little additional benefit. Consequently, the best performance is achieved by the 4-qubit, 2-layer configuration. These results suggest that increasing the number of qubits is the primary contributor to performance gains, while deeper circuits mainly increase optimization complexity without improving reconstruction quality.

V. CONCLUSION

This paper proposed QSAT-SEMCOM, a quantum-classical hybrid SemCom framework for satellite image transmission by integrating a VQC into the semantic encoder of a DeepJSCC system. Experimental results demonstrated robust image transmission under AWGN channels and competitive reconstruction performance across different SNR levels. The ablation study further showed that increasing the number of qubits consistently improves PSNR and MS-SSIM with only a negligible increase in model parameters, while moderate circuit depth is sufficient to achieve the best performance. These results demonstrate the potential of VQCs for enhancing semantic image transmission in DeepJSCC systems. Future work will evaluate the proposed framework on more challenging remote sensing tasks, including high-resolution satellite and hyperspectral image transmission, alongside larger hybrid quantum-classical architectures, advanced quantum circuit designs, and improved training strategies as quantum hardware matures.

REFERENCES

[1] L. X. Nguyen *et al.*, "A contemporary survey on semantic communications: Theory of mind, generative ai, and deep joint source-channel coding," *IEEE Commun. Surv. Tutor.*, vol. 28, pp. 2377–2417, 2026.
[2] A. Andreou, C. X. Mavromoustakis, G. Mastorakis, A. Bourdena, and E. K. Markakis, "Quantum computing in semantic communications: Overcoming optimization challenges with high-dimensional hilbert spaces," *IEEE Access*, vol. 13, pp. 157 942–157 964, 2025.
[3] Y. Qiao *et al.*, "Deepseek-inspired exploration of rl-based llms and synergy with wireless networks: A survey," *ACM Comput. Surv.*, vol. 58, no. 7, Dec. 2025. [Online]. Available: https://doi.org/10.1145/3776745
[4] E. Bourtsoulatze, D. Burth Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 3, pp. 567–579, 2019.
[5] K. Yang *et al.*, "Witt: A wireless image transmission transformer for semantic communications," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2023, pp. 1–5.
[6] ——, "Swinjscc: Taming swin transformer for deep joint source-channel coding," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 1, pp. 90–104, 2025.
[7] H. Wu, Y. Shao, C. Bian, K. Mikolajczyk, and D. Gündüz, "Deep joint source-channel coding for adaptive image transmission over mimo channels," *IEEE Trans. Wirel. Commun.*, vol. 23, no. 10, pp. 15 002–15 017, 2024.
[8] F. Fan, Y. Shi, T. Guggemos, and X. X. Zhu, "Hybrid quantum-classical convolutional neural network model for image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 12, pp. 18 145–18 159, 2024.
[9] B. Narottama, Z. Mohamed, and S. Aïssa, "Quantum machine learning for next-G wireless communications: Fundamentals and the path ahead," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 2204–2224, 2023.
[10] B. Narottama and S. Aïssa, "Quantum machine learning for performance optimization of RIS-assisted communications: Framework design and application to energy efficiency maximization of systems with RSMA," *IEEE Trans. Wirel. Commun.*, vol. 23, no. 10, pp. 12 830–12 843, 2024.
[11] M. Chehimi, C. Chaccour, C. K. Thomas, and W. Saad, "Quantum semantic communications for resource-efficient quantum networking," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 803–807, 2024.
[12] S. Meng *et al.*, "Semantics-empowered space-air-ground-sea integrated network: New paradigm, frameworks, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 1, pp. 140–183, 2025.
[13] H. Cao, R. Meng, X. Xu, S. Han, and P. Zhang, "Importance-aware robust semantic transmission for leo satellite–ground communication," *IEEE Internet Things J.*, vol. 13, no. 5, pp. 9665–9681, 2026.
[14] A. Sebastianelli, D. A. Zaidenberg, D. Spiller, B. L. Saux, and S. L. Ullo, "On circuit-based hybrid quantum neural networks for remote sensing imagery classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 15, pp. 565–580, 2022.
[15] Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems Computers, 2003*, vol. 2, 2003, pp. 1398–1402.
[16] P. Helber, B. Bischke, A. Dengel, and D. Borth, "Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 7, pp. 2217–2226, 2019.