From Analog to Digital Semantic Communications: Architectures, Challenges, and Future Directions

Jian Zou, Wenwu Xie, Jian Xiao, Yongsheng Liang, Jules M. Moualeu, and Liang Yang

Abstract

Semantic communication (SemCom) aims to achieve an efficient and intelligent communication model by deeply understanding and accurately transmitting the semantics of information, thus surpassing traditional communication. However, analog semantic communication (A-SemCom) systems exhibit significant shortcomings in compatibility and flexibility, making it difficult to adapt to diverse network environments and rapidly changing demands. This has led to the development of digital semantic communication (D-SemCom) systems. In this article, we first compare the architectures of traditional communication, A-SemCom, and D-SemCom systems by analyzing their respective advantages and disadvantages. Subsequently, we propose a novel D-SemCom system that can accurately transmit semantics of the message by overcoming the limitations of analog systems, providing stronger adaptability and flexibility, and meeting the demands for efficient semantic transmissions in complex network environments. Finally, we briefly outline potential avenues for future research in SemCom that include the further development of theoretical frameworks, the enhancement of security mechanisms, and the integration with advanced communication technologies.

Introduction

With the widespread deployment of fifth-generation (5G) mobile communications and the accelerated development of sixth-generation (6G) technologies, communication systems are fundamentally shifting from "ubiquitous connectivity" to "intelligent interaction." This evolution gives rise to increasingly complex and diverse human–machine interaction scenarios, which demand not only massive machine-type communication but also natural, efficient, and intelligent human–machine interactions. Against this backdrop, the limitations of conventional communication architectures are becoming increasingly evident. In particular, existing source and channel coding methods have approached the Shannon limit, leaving little room for further optimization and making them insufficient to meet the growing demands of next-generation communication systems [1]. Although advances in hardware have continued to improve the performance of communication systems, hardware-driven innovation is approaching its limits and showing signs of unsustainability. Semantic communication (SemCom) emerges as a promising paradigm to address these challenges. Rather than focusing on the transmission of raw bits, SemCom emphasizes the transmission of information meaning. By modeling and optimizing the semantic content of information, this new communication paradigm significantly enhances both the efficiency and robustness of data transmission, aligning more closely with the demands of intelligent and adaptive communication scenarios.

According to the classical communication model proposed by Weaver and Shannon, an information transmission can be divided into three levels: syntax, semantics, and pragmatics. The syntactic level focuses on the structure and arrangement of symbols to ensure correct encoding and decoding; the semantic level emphasizes the actual meaning of the transmitted information to ensure effective communication; and the pragmatic level concerns the use and interpretation of information within specific contexts, aiming to ensure that the information guides the receiver toward accurate understanding and appropriate response [2]. Although these three levels together constitute a complete communication framework, traditional technological limitations have made it difficult for earlier systems to achieve joint optimization across all dimensions. In recent years, the rapid development of artificial intelligence (AI), particularly deep learning, has opened up new opportunities for the construction of SemCom systems. Deep learning has demonstrated strong capabilities in learning semantic representations, modeling complex relationships, and capturing contextual dependencies from large-scale data. These advancements pro-

Jian Zou is with the Shenzhen University, China and is also with Shenzhen Technology University, China; Wenwu Xie is with the Hunan Institute of Science and Technology, China; Jian Xiao is with Central China Normal University, China; Yongsheng Liang is with Shenzhen Technology University, China and with Shenzhen University, China; Jules M. Moualeu is with the University of the Witwatersrand, South Africa and also with Lappeenranta–Lahti University of Technology, Finland; Liang Yang (corresponding author) is with Hunan University, China and Jishou University, China. Jian Zou and Wenwu Xie contributed equally to this work.

Digital Object Identifier: 10.1109/MWC.001.2500052

vide critical support for shifting from conventional bit-level optimization toward meaning-centered transmission, thereby offering a new paradigm to meet the evolving demands of future intelligent communication networks [3].

In recent years, SemCom has made significant progress in several key areas, particularly in improving system robustness and task adaptability. For example, semantic encoding and decoding methods guided by attention mechanisms based on prior knowledge of the signal-to-noise ratio (SNR) have enhanced the semantic representation accuracy under noisy conditions [4]. The gumbel-softmax-based adaptive transmission rate control has improved the system responsiveness to dynamic channel capacities [5]. In addition, generative diffusion model-based image reconstruction modules have achieved high-quality semantic image recovery under constrained rates [6]. Moreover, SemCom has demonstrated excellent performance in various tasks such as multimodal communication [7], channel state information feedback [8], and video transmission [9]. However, despite these initial achievements, current SemCom architectures still face multiple challenges in practical deployment, mainly in the following three aspects:

1. **Adaptability**: Most current SemCom systems rely on idealized channel assumptions, such as simulations based on additive white Gaussian noise models. However, these assumptions fail to accurately reflect the complexity of real-world wireless environments, which are typically affected by multipath fading, temporal variation, and interference. As a result, SemCom systems often exhibit limited robustness and generalization capability under complex and dynamic channel conditions.
2. **Flexibility**: SemCom typically requires an end-to-end joint training between the transmitter and receiver to achieve optimal performance. However, when channel conditions or system architectures change, the entire model often needs to be retrained, resulting in significant computational overhead and latency. The lack of modularity and generalization limits the system's capacity for rapid adaptation in dynamic deployment scenarios.
3. **Compatibility**: Many existing SemCom approaches assume that semantic representations can be directly transmitted as continuous complex-valued vectors, whereas modern communication systems are fundamentally based on discrete digital signals. This fundamental mismatch in signal format and system interfaces poses major challenges to the seamless integration of SemCom with existing communication infrastructures.

In summary, three major breakthroughs are needed to advance SemCom architectures:

- Developing robust SemCom methods tailored to complex and dynamic wireless environments to enhance system adaptability.
- Exploring systematic approaches for decoupling semantic encoding from channel conditions to improve architectural flexibility and transferability.
- Constructing digital semantic communication (D-SemCom) frameworks that are compatible with existing infrastructures to improve both the protocol interoperability and engineering feasibility.

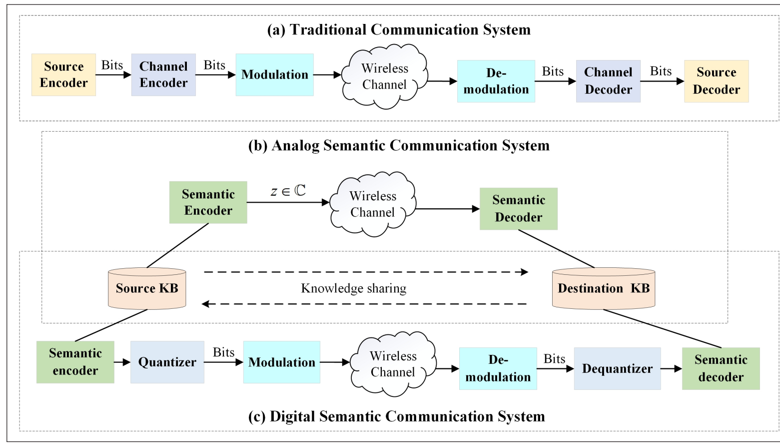

FIGURE 1. A detailed diagram comparing the architectures of traditional communication, A-SemCom, and D-SemCom.

Motivated by the above research challenges, this article proposes a flexible and robust D-SemCom architecture employing orthogonal frequency division multiplexing (OFDM), designed to achieve high adaptability and compatibility in complex channel environments. The main contributions of this article are as follows:

- We analyze the limitations of existing SemCom architectures and outline the evolutionary path from traditional communication to analog semantic communication (A-SemCom), and further to D-SemCom. In addition, we systematically compare the three paradigms in terms of architectural design, technical mechanisms, and key characteristics, thereby providing a theoretical support for the design and optimization of D-SemCom architectures.
- We develop a D-SemCom system based on OFDM, aiming to enhance robustness under complex channel conditions. Building on this system, we propose a bit-level semantic transmission network, SANet, which performs a nonlinear mapping from received signals to bit-level semantic targets, thereby achieving effective decoupling between semantic encoder/decoder and specific channel conditions. Experimental results demonstrate that SANet achieves lower semantic reconstruction error and exhibits clear advantages over traditional decoding methods.
- Finally, we explore the future development of SemCom, with a focus on ongoing challenges in SemCom theory and standardization, system security and privacy protection, as well as deep integration with other communication technologies.

Architecture of Communication

The fundamental distinction between SemCom and traditional communication frameworks lies in its emphasis on the "meaning" of information, rather than the precise transmission of bits. Figure 1 illustrates the architectural differences among traditional communication, A-SemCom, and D-SemCom, which are comparatively analyzed in this section.

Traditional Communication Architecture

Traditional communication is designed to ensure the precise transmission of bit sequences through a layered process, as illustrated in Fig. 1a. The pro-

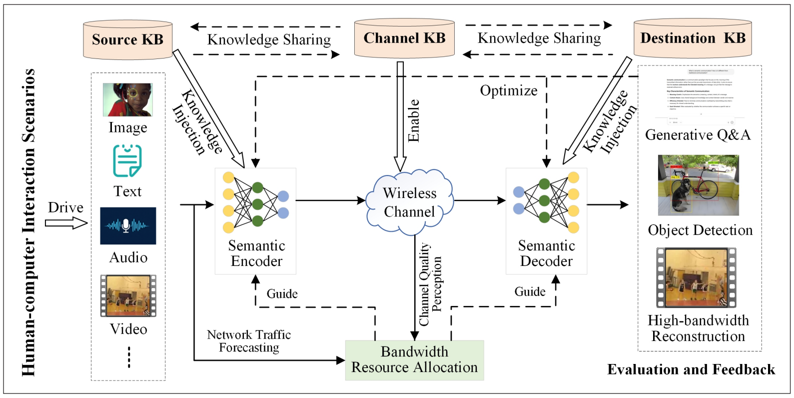

FIGURE 2. Deep learning-enabled SemCom.

cess begins with the source encoding, where raw data is converted into a bitstream using compression techniques such as Huffman coding or multimedia codecs (e.g., joint photographic experts group (JPEG), better portable graphics (BPG)). To enhance reliability, channel encoding introduces redundancy using Turbo codes or low-density parity-check (LDPC) codes, enabling error detection and correction. The encoded bitstream is then modulated for transmission over a wireless channel using modulation schemes such as quadrature amplitude modulation (QAM), phase-shift keying (PSK). However, the wireless channel introduces challenges such as noise, interference, and fading, which degrade the signal integrity. At the receiver, the signal undergoes demodulation, channel decoding, and source decoding to reconstruct the original data and correct transmission errors. Although this framework ensures accurate bit-level transmission, it neglects the semantic relevance of information, often resulting in redundant data transmission and suboptimal bandwidth utilization. This limitation underscores the need for SemCom, which prioritizes the transmission of meaningful information over raw bit accuracy.

SemCom Architecture

Fundamentals of SemCom: As illustrated in Fig. 2, SemCom systems are typically developed for human-computer interaction scenarios, where multimodal data such as text, speech, and images are generated in response to specific semantic tasks. The system employs a semantic encoder-decoder to perform unified modeling and extract semantic features from the multimodal inputs. This yields semantic representations that are aligned with the task requirements. During this process, semantic encoding not only performs efficient compression but also dynamically adjusts resource allocation based on the data volume and channel conditions. These adjustments help improve the communication efficiency and robustness. Finally, the decoder generates the semantic task outcome and evaluates its effectiveness. The evaluation results are then fed back to the semantic encoder-decoder, enabling adaptive adjustments to semantic extraction and representation in subsequent transmissions. This feedback loop continuously improves the communication performance and task success rate.

Efficient Semantic Encoding and Representation: To support SemCom tasks, the system must extract and encode relevant information from multimodal inputs such as text, speech, and images into compact, task-aware semantic representations. These representations are designed to preserve the semantic meaning rather than the signal-level fidelity, thereby enabling a more efficient and task-oriented communication.

Real-Time Resource Allocation under Channel Dynamics: Given the variability of network conditions and traffic loads, real-time resource allocation is essential to maintaining the communication efficiency and robustness. By adapting transmission strategies dynamically according to the current data volume and channel state, the system ensures efficient bandwidth utilization for semantic information transmissions.

Adaptive Feedback for Task-Oriented Optimization: SemCom systems leverage task outcome feedback to guide subsequent encoding and extraction strategies. This adaptive feedback loop enables the system to continuously refine its semantic representations, ensuring alignment of transmitted content with evolving task requirements and enhancing the overall system performance.

A-SemCom Architecture: A-SemCom represents an emerging communication paradigm that enables an end-to-end transmission of semantic information within a continuous feature domain. As shown in Fig. 1b, A-SemCom employs an architecture that emphasizes direct semantic encoding and transmission, bypassing discrete symbol-level processing. This approach offers significant advantages, including low latency, low computational complexity, and graceful degradation in performance under poor channel conditions. A-SemCom typically utilizes deep neural networks to perform semantic modeling and continuous encoding of multimodal sensory data, effectively bypassing traditional communication procedures including entropy coding, symbol mapping, and discrete channel coding. This leads to a compact, task-oriented transmission pipeline. Owing to the end-to-end joint optimization, A-SemCom achieves superior performance in task-driven applications such as semantic recognition and decision-making. However, due to its reliance on continuous signal processing and task-specific semantic modeling, A-SemCom exhibits fundamental structural incompatibilities with existing digital communication systems in terms of protocol stack design, channel access mechanisms, and standardized network interfaces, making the seamless integration into mainstream infrastructures difficult. This challenge becomes more pronounced in complex scenarios such as human-computer interactions, which demand support for highly diverse and dynamically evolving semantic tasks. In such cases, A-SemCom often requires extensive customization based on task type, data modality, and channel conditions. This high level of task-specific customization limits the system's generalizability and scalability, while also significantly increasing deployment and maintenance costs, thereby posing a major barrier to its large-scale adoption in real-world communication systems.

D-SemCom Architecture: D-SemCom represents and transmits semantic information in a discrete form, in clear contrast to A-SemCom, which operates on continuous signal processing. As shown in Fig. 1c, D-SemCom adopts a digital encoding mechanism that expresses semantic

Characteristic	Traditional Communication	A-SemCom	D-SemCom
Core Objective	Reliable bitstream transmission	Transmission of continuous semantic information	Transmission of discrete semantic information
Technology Basis	Modular design (coding, modulation, error correction	End-to-end learning	End-to-end learning, quantization, modulation
Noise Resistance	Strongly depends on channel coding	Strong noise robustness	Strong noise robustness
Spectral Efficiency	Low (due to bitstream and redundancy)	High (no redundancy)	High (no redundancy)
Flexibility	/	Relatively fixed, limited adaptability	High flexibility, strong scalability
Computational Cost	Low, modular and hardware-efficient	Medium, end-to-end continuous mapping is lightweight	High, multiple intermediate steps (quantization, modulation) add complexity
Typical Applications	Telephony, file transfer, video streaming	Supports semantic task scenarios but requires customized design	Suitable for task-oriented scenarios such as intelligent question answering, autonomous driving, and vehicular networks

TABLE 1. Comparison of traditional communication, A-SemCom, and D-SemCom.

information through bitstreams and enables end-to-end transmission via standardized communication protocols. This architecture makes D-SemCom inherently compatible with existing digital communication infrastructures such as 5G, and also scalable to future networks like 6G. Moreover, by leveraging mature mechanisms in digital systems, including channel coding, error correction, and interference mitigation, D-SemCom exhibits strong reliability and robustness in transmission.

A fundamental distinction between D-SemCom and A-SemCom lies in the use of discrete quantization, where continuous semantic features are mapped into a finite set of symbols. Common quantization methods include uniform quantization and vector quantization (VQ). The former divides the feature space into equal intervals for a simple and efficient encoding but may result in large errors when handling high-dimensional features. In contrast, VQ learns an optimized codebook to minimize the quantization loss and improve the fidelity of semantic reconstruction [10]. This discretization process is not only a core technical difference from A-SemCom but also serves as a foundational step for integrating semantic information into digital communication pipelines.

At the physical layer, D-SemCom employs standard digital modulation techniques to convert the quantized bitstream into transmittable signals, unlike A-SemCom, which directly modulates continuous semantic features. D-SemCom commonly uses classical schemes such as QAM and PSK, often combined with OFDM to enhance the spectral efficiency and resistance to interference. In recent years, AI-based intelligent modulation approaches, such as neural network modulation, have attracted considerable attention. These techniques can automatically select optimal modulation strategies based on channel conditions, data modalities, and task requirements, significantly improving system adaptability and communication reliability [11].

In summary, by incorporating discrete quantization and digital modulation mechanisms, D-SemCom not only ensures effective semantic transmissions but also significantly enhances the system compatibility, deployability, and engineering feasibility. This marks a key distinction from A-SemCom.

Comparison Between Traditional Communication, A-SemCom, and D-SemCom

Table 1 provides a comprehensive, multi-dimensional comparison of traditional communication, A-SemCom, and D-SemCom, covering aspects such as core objectives, technological foundations, noise robustness, spectral efficiency, system flexibility, computational complexity, and typical application scenarios. Traditional communication systems emphasize reliable bit-level transmissions based on modular designs involving source coding, channel coding, and modulation. However, they often suffer from low spectral efficiency and limited semantic awareness. In contrast, A-SemCom leverages end-to-end deep learning techniques to directly transmit continuous semantic information, such as speech and image features, without explicit quantization or modulation. It offers some advantages in terms of low latency and graceful performance degradation under adverse channel conditions. Despite these attractive features, it exhibits limited adaptability and often requires task-specific customization.

D-SemCom adopts a hybrid architecture that integrates end-to-end learning with discrete quantization and digital modulation. It supports the transmission of discrete semantic information through bitstreams, providing high compatibility with existing communication protocols and infrastructures such as 5G, and offering scalability for future 6G networks. Although its multi-stage processing increases in computational complexity, D-SemCom demonstrates an improved scalability, a greater task flexibility, and a robust performance under noisy channel conditions, making it a compelling alternative to A-SemCom.

Case Study: Image Transmission in Different Communication Architectures

Traditional communication systems adopt a modular design that effectively decouples source encoding from physical-layer transmission, providing strong system generality and implementation efficiency. However, conventional receiver algorithms such as least squares (LS) and linear minimum mean square error (LMMSE) rely heavily on linear channel models and statistical assumptions

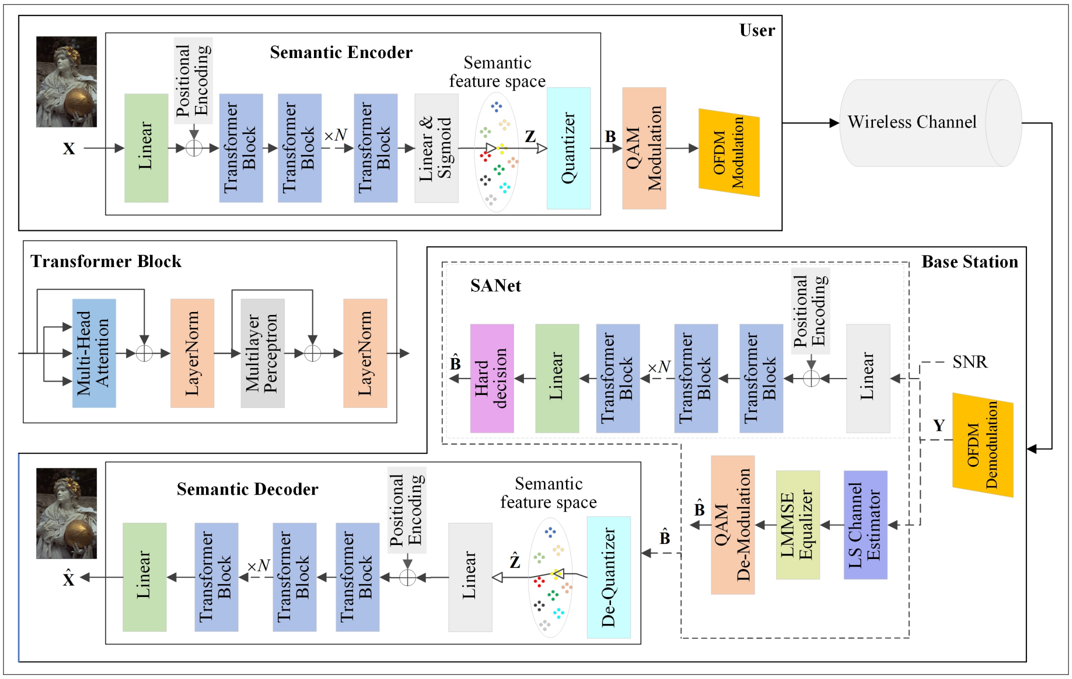

FIGURE 3. Overview of the proposed OFDM-based D-SemCom framework for image transmission.

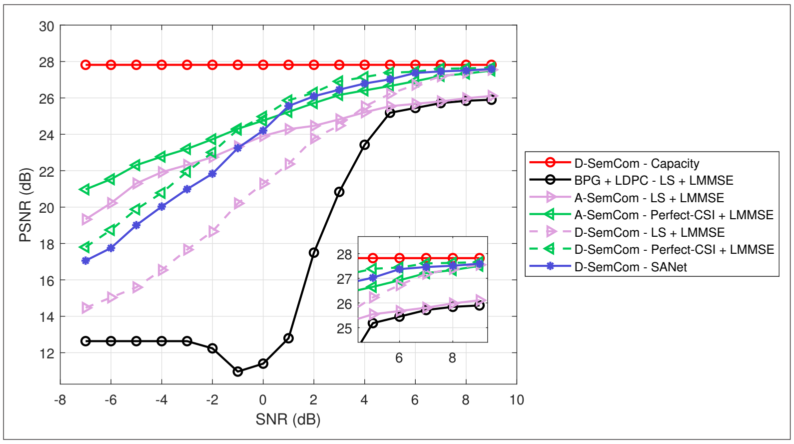

FIGURE 4. CIFAR-10 image reconstruction PSNR under different communication architectures.

of noise. Consequently, they are inadequate for handling the nonlinear interference and dynamic variations commonly encountered in real-world wireless communication environments. To address these limitations, this article designs a D-SemCom system based on OFDM and proposes a bit-level semantic transmission network named SANet. This network utilizes deep learning techniques to learn the nonlinear mapping between received signals and semantic bit-level targets, effectively mitigating the strong coupling between the semantic encoder and specific channel conditions. This approach enhances the system's adaptability and semantic recovery capability under complex channel environments. The detailed network structure is shown in Fig. 3.

To evaluate the effectiveness of the proposed system, we assess the image transmission quality of several communication architectures in the multi-user multiple-input multiple-output (MU-MIMO) OFDM uplink UMi scenario provided by the Sionna library [12]. The evaluation metric used is the peak signal-to-noise ratio (PSNR). The communication system is configured with 12 OFDM symbols, 96 subcarriers, and adopts the "2P" pilot pattern. The system consists of four users, each equipped with a single transmit antenna, and a base station equipped with eight receive antennas. In the digital system, 4-QAM modulation is used along with a 2-bit uniform quantization. In the conventional separate system, BPG is adopted for source coding, and a rate-1/2 LDPC code is used for channel encoding. Both semantic communication systems employ the same encoder and decoder architectures, which are based on a six-layer Transformer backbone with 192-dimensional embeddings, also using a multi-head self-attention with 4 heads, and a multilayer perceptron (MLP) expansion factor of 4 [13]. For SANet, the embedding dimension is set to 256, using a multi-head self-attention with 4 heads, and an MLP expansion factor of 2. All systems operate under the same channel bandwidth ratio of R = 0.25 to ensure a fair comparison. Under this configuration, each image in the analog semantic communication system is encoded into a 1536-dimensional semantic vector, corresponding to 768 complex-valued symbols. In contrast, the digital semantic communication system generates 1536 bits per image, which are mapped to the same number of complex symbols via 4-QAM modulation.

As shown in Fig. 4, the PSNR performance achieved by the three OFDM-based communication architectures on the image-transmission task is presented, where "Capacity" denotes the peak performance that the semantic encoder/decoder can achieve under ideal conditions without channel transmission.

Cliff Effect: In Fig. 4, the conventional "BPG + LDPC" separate scheme experiences a classic cliff collapse: once the SNR drops below -1 dB, decoding fails and the PSNR falls to 12.445 dB. In contrast, both A-SemCom and D-SemCom architectures avoid this cliff effect. By jointly optimizing the source and channel through their semantic encoder-decoder, they maintain a usable image quality even in the same low-SNR regime.

Low Spectral Efficiency: Conventional separate architectures rely on channel coding redundancy to combat noise. For example, a rate-1/2 LDPC code requires 1 parity bit for every 1 bit of information, effectively halving the usable bandwidth and significantly reducing the spectral efficiency. In contrast, A-SemCom and D-SemCom systems leverage the strong nonlinear fitting capability of deep neural networks to learn noise-robust representations through the end-to-end training, eliminating the need for explicit channel coding. Under the same bandwidth, all symbols can carry meaningful semantic information, resulting in a net information rate roughly twice that of the separate scheme and a significant improvement in spectral efficiency.

PSNR Performance: Both A-SemCom and D-SemCom outperform the traditional separate coding scheme in terms of PSNR for image reconstruction. Specifically, A-SemCom demonstrates stronger robustness under low SNR conditions, while D-SemCom achieves better reconstruction quality at high SNR. Moreover, the proposed SANet significantly enhances the PSNR performance of D-SemCom, further highlighting its advantages.

Additionally, Fig. 5 presents the reconstructed images under different SNR conditions for both A-SemCom and D-SemCom systems with various receiver configurations. For comparison, both systems adopt traditional receivers "LS + LMMSE" and ideal receivers "Perfect CSI + LMMSE" as baselines. Additionally, the D-SemCom system includes the proposed semantic-aware receiver network SANet. The results show that under low SNR conditions , the reconstruction quality of traditional receiver methods degrades significantly, with severe blurring and distortion, particularly when D-SemCom uses "LS + LMMSE." In contrast, the proposed SANet consistently demonstrates stronger robustness and higher reconstruction quality across all SNR levels. Notably, even under low and moderate SNRs, SANet is able to preserve clear structural details and semantic contours, verifying its high adaptability and strong robustness in non-ideal channel environments.

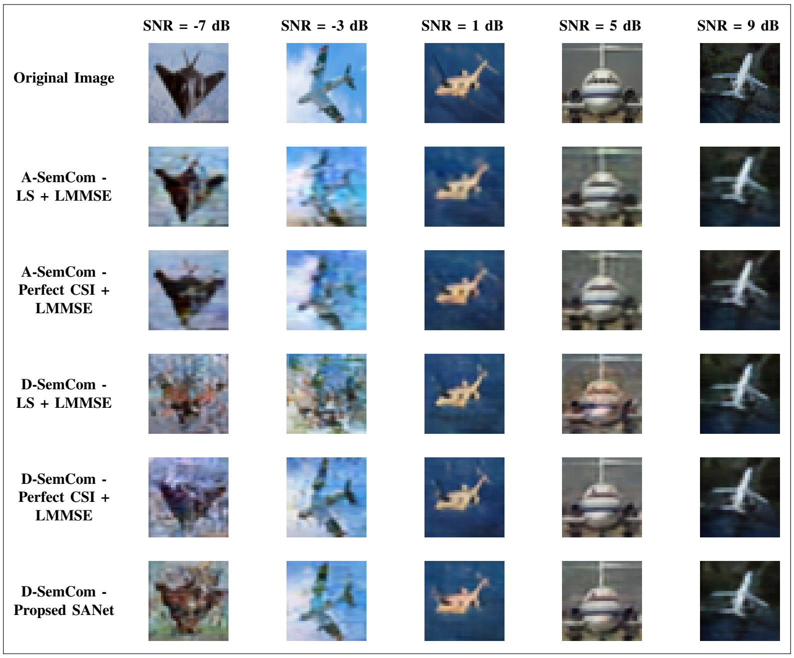

FIGURE 5. Performance comparison of different transmission schemes under varying SNR conditions.

Open Issues and Future Directions

As SemCom continues to advance, the improvement of its theoretical foundations, enhancement of security mechanisms, and effective integration with existing communication technologies have become critical challenges that hinder its practical deployment and sustainable development. This section focuses on these core issues and analyzes the significance of unified theoretical frameworks, strengthened security mechanisms, and cross-technology integration in future research.

Theoretical Foundations and Standardization of SemCom

Current SemCom systems primarily rely on deep learning models to perform semantic representation, compression, and transmission. Although they have shown promising performance across various tasks, the lack of a solid theoretical foundation remains a critical bottleneck. The inherent opacity and poor interpretability of deep models make it difficult to establish clear system models, performance boundaries, and a unified theoretical guidance for SemCom. To address this, it is essential to develop a unified and systematic theoretical framework that covers key components such as semantic modeling, representation compression, transmission mechanisms, and task-oriented decoding, while ensuring compatibility with fundamental disciplines such as communication theory, information theory, and language processing. In parallel, standardization is urgently needed to define protocol structures, functional interfaces, and performance evaluation metrics, thereby enhancing the generality, consistency, and scalability of SemCom systems.

Security and Privacy Protection in SemCom

SemCom leverages deep learning models to construct end-to-end semantic encoding and decoding processes. However, the inherent opacity and high sensitivity of these models introduce new challenges in security and privacy. Deep models are vulnerable to adversarial perturbations, where attackers can manipulate transmission outcomes through subtle semantic-level modifications, leading to misclassification or semantic deception. Moreover, the highly condensed semantic features may inadvertently expose sensitive information such as user identity or behavioral intent, posing risks of model inversion and privacy inference. Future research should explore how to enhance model robustness against semantic interference, implement context-aware semantic consistency verification at the receiver side, and dynamically regulate the extraction and exposure of sensitive semantics based on task context, thereby enabling secure and trustworthy SemCom in open environments.

Integration of SemCom with Emerging Technologies

SemCom focuses on task-oriented information transmission, emphasizing semantic meaning over bit-level fidelity. However, it still faces practical challenges in dynamic environment adaptability, wide-area accessibility, multi-user resource allocation, and efficient spectrum utilization, which hinder its large-scale deployment. To address these limitations, future research should explore the integration of SemCom with emerging technologies such as non-orthogonal multiple access (NOMA) [14], reconfigurable intelligent surface (RIS), and non-orthogonal superimposed pilot (NOSIP) [15]. NOMA improves multi-user concurrency by enabling efficient resource sharing across overlapping users and tasks. RIS allows for dynamic reconfiguration of the wireless propagation environment, thereby enhancing signal quality and semantic decoding accuracy. In addition, NOSIP increases the utilization efficiency of resource grids, effectively improving system goodput in bandwidth-constrained scenarios. Through coordinated optimization with these technologies, SemCom is expected to overcome current performance bottlenecks and support scalable deployment in complex and heterogeneous network environments.

Conclusions

This article investigates the development trajectory of SemCom and systematically compares the architectural designs and key characteristics of traditional communication, A-SemCom, and D-SemCom. To address the limitations of A-SemCom in terms of system compatibility and flexibility, a D-SemCom system architecture based on OFDM is proposed to enhance adaptability in practical communication environments. Building on this architecture, a bit-level semantic transmission network, SANet, is introduced to establish a nonlinear mapping from received signals to bit-level semantic targets, thereby achieving effective decoupling between semantic encoder/decoder and specific channel conditions. Experimental results show that SANet significantly outperforms traditional decoding methods in terms of semantic reconstruction error, demonstrating a strong transmission performance and robustness. Finally, this article presents a forward-looking discussion on the future development of SemCom, focusing on ongoing challenges including the unification of theoretical frameworks, the enhancement of security and privacy mechanisms, and the deep integration with advanced communication technologies.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62031013, in part by the Guangdong Province Key Construction Discipline Scientific Research Capacity Improvement Project under Grant 2022ZDJS117, and in part by the Natural Science Foundation of Hunan Province under Grant 2023JJ50045, Grant 2025JJ70287, Grant 2024JJ7218, and Grant 2024JJ7219.

References

[1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, Oct. 1948, pp. 379–423.

[2] W. Weaver, "Recent Contributions to the Mathematical Theory of Communication," *ETC Rev. Gen. Semant.*, vol. 10, no. 4, 1953, pp. 261–81.

[3] E. Bourtsoulatze, D. Burth Kurka, and D. Gündüz, "Deep Joint Source-Channel Coding for Wireless Image Transmission," *IEEE Trans. Cognit. Commun. Net.*, vol. 5, no. 3, Sept. 2019, pp. 567–79.

[4] J. Xu *et al.*, "Wireless Image Transmission Using Deep Source Channel Coding with Attention Modules," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 4, Apr. 2022, pp. 2315–28.

[5] M. Yang and H. -S. Kim, "Deep Joint Source-Channel Coding for Wireless Image Transmission with Adaptive Rate Control," *Proc. IEEE Int'l. Conf. Acoust., Speech Signal Process. (ICASSP)*, Singapore, Singapore, 2022, pp. 5193–97.

[6] C. Liang *et al.*, "Selection-Based Image Generation for Semantic Communication Systems," *IEEE Commun. Lett.*, vol. 28, no. 1, Jan. 2024, pp. 34–38.

[7] G. Zhang *et al.*, "A Unified Multi-Task Semantic Communication System for Multimodal Data," *IEEE Trans. Commun.*, vol. 72, no. 7, July 2024, pp. 4101–16.

[8] J. Zou *et al.*, "Distributed Learning-Based Channel Estimation and Feedback for RIS-Aided Wireless Communications," *IEEE Wireless Commun. Lett.*, vol. 14, no. 2, Feb. 2025, pp. 460–64.

[9] T. -Y. Tung and D. Gündüz, "DeepWiVe: Deep-Learning-Aided Wireless Video Transmission," *IEEE JSAC*, vol. 40, no. 9, Sept. 2022, pp. 2570–83.

[10] Q. Hu *et al.*, "Robust Semantic Communications with Masked VQ-VAE Enabled Codebook," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 12, Dec. 2023, pp. 8707–22.

[11] Y. Bo *et al.*, "Joint Coding-Modulation for Digital Semantic Communications via Variational Autoencoder," *IEEE Trans. Commun.*, vol. 72, no. 9, Sept. 2024, pp. 5626–40.

[12] J. Hoydis *et al.*, "Sionna: An Open-Source Library for Next-Generation Physical Layer Research," Mar. 2022, arXiv:2203.11854.

[13] A. Vaswani *et al.*, "Attention is All You Need," *Proc. NeurIPS*, vol. 30, Nov. 2017.

[14] X. Mu *et al.*, "Heterogeneous Semantic and Bit Communications: A Semi-NOMA Scheme," *IEEE JSAC*, vol. 41, no. 1, Jan. 2023, pp. 155–69.

[15] J. Zou *et al.*, "Non-Orthogonal Superimposed Pilot for Goodput Enhancement in Digital Semantic Communication," *IEEE Trans. Cognit. Commun. Netw.*. DOI: 10.1109/TCCN.2025.3573407.

Biographies

Jian Zou (zoujian250@gmail.com) received the B.Eng. and M.Sc. degrees from Hunan Institute of Science and Technology, Yueyang, China, in 2020 and 2023, respectively. He is currently pursuing the Ph.D. degree with the College of Applied Technology, Shenzhen University. His research interests include semantic communication, intelligent communication, and channel estimation.

Wenwu Xie received the B.S., M.S., and Ph.D. degrees in communication engineering from Huazhong Normal University, Wuhan, China, in 2004, 2007, and 2017, respectively. He is currently an Associate Professor with the School of Information Science and Engineering, Hunan Institute of Science and Technology, Yueyang, China. His research interests include communication and control algorithms.

Jian Xiao received the B.Eng. and M.Sc. degrees from Hunan Institute of Science and Technology, Yueyang, China, in 2019 and 2022, respectively. He is currently pursuing the Ph.D.degree with Central China Normal University. His research interests include electromagnetic information theory, reconfigurable intelligent surface, and machine learning.

Yongsheng Liang received the Ph.D. degree from Harbin Institute of Technology in 1999. He has been a Professor at the School of Electronic and Information Engineering, Harbin Institute of Technology, since 2005. He is currently a Chair Professor at Shenzhen Technology University and a Distinguished Professor at Shenzhen University. His research interests include video coding and transmission, machine learning and applications. He was a recipient of the First Prize of the Wu Wenjun Artificial Intelligence Technical Invention Award and the Second Prize of the Guangdong Provincial Science and Technology Award.

Jules M. Moualeu is an Associate Professor with the School of Electrical and Information Engineering, University the Witwatersrand, Johannesburg, South Africa. He received the Ph.D. degree in electronic engineering from the University of KwaZulu-Natal, Durban, South Africa, in 2013. His current research interests include 5G and 6G enabling technologies.

Liang Yang was born in Hunan, China. He received the Ph.D. degree in electrical engineering from Sun Yat-sen University, Guangzhou, China, in 2006. From 2006 to 2013, he was a Teacher with Jinan University, Guangzhou. He joined the Guangdong University of Technology, Guangzhou, in 2013. He is currently a Professor with Hunan University, Changsha, China. His current research interest includes the performance analysis of wireless communication systems.