EAI Endorsed Transactions
on Tourism, Technology and Intelligence

Review Article EAI.EU

From Bits to Meaning: A Survey of Semantic Communications for 6G Networks

James Adu Ansere[1,2], Quang Nhat Le[1,*] and Tuan Thanh Nguyen[3]

[1]Memorial University, Canada
[2]Sunyani Technical University, Ghana
[3]University of Greenwich, United Kingdom

Abstract

The advent of sixth-generation (6G) networks necessitates a paradigm shift from conventional bit communication to meaning-centric systems capable of semantic understanding, contextual inference, and task-oriented optimization. This survey explores the emerging domain of semantic communications (SemCom), which aims to transmit not just data, but relevant and actionable meaning aligned with user intent. Unlike Shannon's theory that emphasizes bit-level fidelity, SemCom emphasizes the semantic and pragmatic utility of transmitted information. This introduces a unified taxonomy aligned with 6G system layers and highlights applications in latency-sensitive, mission-critical applications such as autonomous vehicles (AVs), extended reality (XR), digital twins, and remote healthcare. The study identifies open challenges in semantic reliability, multi-agent alignment, green SemCom, and integration with quantum and bio-inspired systems—positioning SemCom as a cornerstone for future 6G networks.

Received on 26 December 2025; accepted on 03 March 2026; published on 10 March 2026

Keywords: 6G Networks, Edge Intelligence, Quantum Machine Learning, Semantic Communications, Semantic Metrics, Task-Oriented Transmission

doi:10.4108/eettti.11434

1. Introduction

1.1. Beyond Shannon: Toward Meaning-Centric Communication

Communication systems are at a turning point where conveying meaning, rather than just bits, is crucial [1, 2]. Shannon's communication theory focuses on reliable data transmission by matching sent and received bits, without considering meaning. This framework has supported the digital revolution but overlooks the semantics of information exchange [3]. As we move towards sixth-generation (6G), with its vast connectivity, low latency, reliability, and intelligence, the old bit-focused model falls short for contextually rich, mission-driven applications [4]. Emerging technologies, such as autonomous vehicles (AVs), extended reality (XR), digital twins, and remote surgery need systems that prioritize task-relevant data, not raw volume [5, 6]. For example, an AV should send only important information like "pedestrian detected ahead with high collision risk" rather than every frame. This shift focuses on understanding over data quantity.

In communication systems, "semantics" means the significance of transmitted information in achieving a specific goal [7]. Unlike Shannon's theory of quantifying entropy as uncertainty reduction, semantic communications focus on task utility on how the received information aids the intended purpose. Semantic communication (SemCom) supports this by transmitting crucial information, reducing overhead while enhancing intelligence, adaptability, and resource efficiency [8]. This shift is a major advancement in communication systems, integrating artificial intelligence (AI), knowledge representation, and reasoning into the communication stack. Meaning-centric communication is driven by the need for systems that intelligently perceive, interpret, and act under uncertainty, making 6G not only connected but also contextually aware and cognitively adaptive [9].

*Corresponding author. Email: qnle@mun.ca

1.2. Limitations of Bit-Based Transmission

Conventional communication systems emphasize bit accuracy, ensuring that transmitted bits are reproduced exactly at the receiver, regardless of their contextual significance [10]. This design philosophy treats all data uniformly, whether it encodes vital alerts or redundant background content. In real-time, data-intensive environments, such as autonomous driving, industrial internet of things (IoT), and smart cities, this approach leads to bandwidth congestion [11]. Additionally, it causes computational overload and inefficient energy consumption.

Classical source and channel coding techniques are optimized for fidelity rather than task relevance. As a result, important semantic cues may be lost, while irrelevant data is preserved [12]. This lack of contextual awareness impedes timely decision-making and increases communication overhead. The bit-centric model thus reveals critical limitations in modern systems that must now support intelligent, scalable, and adaptive services.

1.3. Why Semantic Communications in 6G

SemCom redefines the role of networks, from mere data pipelines to intelligent systems that extract, transmit, and interpret meaning [13]. In 6G environments, featuring AVs, XR, digital twins, and eHealth systems must convey only information that directly contributes to task completion or decision-making [14, 15]. SemCom enables this by transmitting high-level semantic abstractions instead of raw data. For example, rather than sending full video frames, a vehicle may report pedestrian detected ahead, reducing bandwidth while enhancing responsiveness. This approach supports energy-efficient, low-latency, and carbon-aware operation, which is essential for massive-scale 6G connectivity [16]. Moreover, semantic-aware networks can dynamically adapt to context, prioritize critical content, and infer user intent key capabilities for emerging use cases where milliseconds matter [17]. By shifting focus from bit correctness to meaning effectiveness, SemCom delivers the cognitive agility required by next-generation intelligent systems. Figure 1 presents a 6G SemCom framework focusing on meaning-centric information aligned with user intent for mission-driven applications. It uses AI for semantic understanding, feature extraction, and context-adaptive reconstruction, enhancing efficiency, reducing overhead, and improving reliability in 6G scenarios like XR, AVs, and reconfigurable intelligent surface (RIS). This work establishes SemCom as essential for future networks, promoting smarter, purpose-aware, and efficient communication systems.

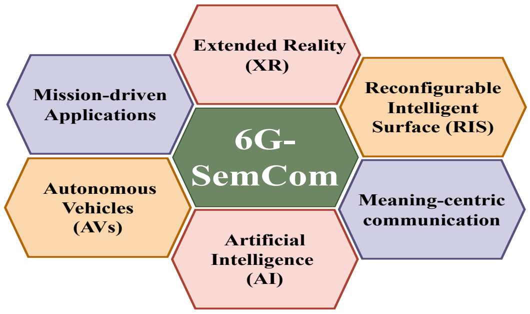

Figure 1. 6G applications with SemCom connectivity.

1.4. Related Surveys

Recent surveys in SemCom have identified key enablers in intelligent and goal-oriented 6G networks, covering theory, metrics, cross-layer architectures, and multi-domain integrations [18]. This section highlights their contributions and identifies research gaps. Work in [19] comprehensively surveys semantic communication in "Intellicise" networks, classifying it into knowledge-driven, data-driven, and hybrid paradigms, and discussing network cognition, intent-aware design, and semantic interoperability. However, it underexplores deployment frameworks, standardization, and real-time implementation. In [20], semantic communication is explored as part of the future internet, highlighting context-driven reasoning and semantic compression. It supports cognitive networks, but practical runtime architectures are still mostly theoretical. [21] provides a technical survey on semantic encoding, multi-agent coordination, and resource optimization in 6G, introducing semantic-aware controls and task-driven communication. However, latency, carbon-efficiency, and hardware design are only briefly mentioned. Semantic information measurement is extensively examined in [22] and further developed in [23]. The first offers a comparative survey of over 30 semantic metrics like semantic entropy and task success rates (TSRs), while the second provides a taxonomy of encoder-decoder models and semantic feedback control. These works stress the need to move from accuracy-based to utility-aware metrics but also reveal a lack of standardized benchmarks and datasets.

In [24], a tutorial-survey addresses representation methods, such as symbolic, subsymbolic, hybrid, and their roles in communication systems. It discusses using graph-based embeddings and ontologies for compact, explainable semantic encodings, though integration into dynamic 6G edge-core architectures remains unresolved. In [25], a systemic perspective integrating SemCom with edge learning is explored, focusing on edge inference pipelines, signal processing, and semantic sensing under constrained resources. It offers insights on semantic-awareness at the edge but lacks a taxonomy for managing feedback and control in real-time. From an engineering view, [26] classifies SemCom models by system complexity and feasibility, covering transmitter-side abstraction, modulation, and channel adaptation. Though useful for practitioners, it overlooks co-design with AI, federated learning, and semantics-based security. In the context of immersive connectivity, [27] examines the convergence of edge learning and metaverse applications, focusing on trust management, user intent modeling, and semantic rendering under latency constraints. Performance modeling for XR and digital twin integration remains underdeveloped. [28] investigates the synergy between generative AI (GenAI) and semantic communications, highlighting the role of diffusion models, large language models (LLMs), and cross-modal transformers in information generation and compression. It notes benefits like flexibility and adaptability, but challenges in semantic alignment and trust calibration persist.

While prior surveys on SemCom have provided valuable insights into metrics, architectural visions, and application domains, they often fall short in offering a unified, cross-layer, and deployment-aware perspective. Many focus on isolated components, either metrics, models, or applications, without holistically integrating system-level design, hardware considerations, runtime constraints, and security implications. This survey uniquely bridges these gaps by proposing a layered taxonomy aligned with 6G architectures [29]. It analyzes semantic pipelines from representation to inference and addresses real-time deployment scenarios such as AVs, XR, and RIS-assisted space–air–ground integrated networks (SAGIN). It also incorporates emerging trends like semantic trust, privacy, carbon-aware design, and foundation models. Additionally, it offers reproducible evaluation criteria and standardization insights, positioning it as a comprehensive blueprint for meaning-centric 6G systems. Table 1 provides a comprehensive analysis of the previously mentioned survey works.

1.5. Motivation and Contributions

This survey provides a systematic overview of semantic communications, highlighting its transformative role in enabling intelligent 6G systems [30]. It combines theoretical foundations with system-level insights and outlines future research directions.

The main contributions of this survey are as follows:

- *Unified Taxonomy and Framework:* We introduce a structured classification of SemCom components, including semantic representation, encoding, transmission control, and evaluation within a scalable 6G architecture. This framework provides a conceptual map for designing and analyzing meaning-centric networks.
- *System-Level Perspective:* We offer a comprehensive view of SemCom integration across multiple layers and domains, including edge intelligence, digital twins, and SAGIN. This highlights SemCom's role in enabling cross-layer optimization, context-aware resource allocation, and real-time system orchestration.
- *Practical Insights and Benchmarking:* We survey key datasets, testbeds, simulation tools, and semantic performance metrics (e.g., semantic distortion, TSR) to promote reproducibility and comparability of SemCom research. These insights serve as a foundation for designing benchmarks and evaluating system-level effectiveness.
- *Future Directions and Challenges:* We identify critical research gaps in semantic reliability, semantic alignment in multi-agent networks, energy-aware semantic encoding, and trusted semantic exchange. These open problems inform future research agendas and standardization efforts for 6G semantic-native systems.

2. Foundations of SemCom

SemCom represents a paradigm shift from traditional communication theory. It focuses on the meaning and utility of the transmitted message rather than merely its accurate transmission [31]. This section delves into its foundations, contrasting it with classical information theory and exploring its core concepts and practical implications. Figure 2 gives an overview of the foundational structure of SemCom.

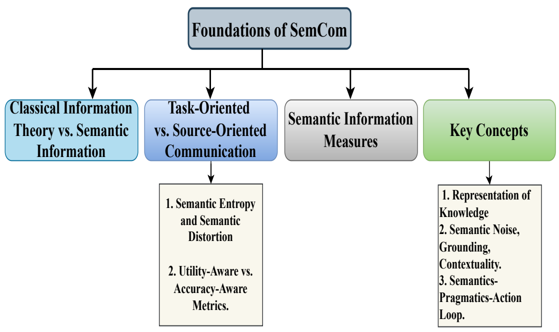

Figure 2. The foundational structure of SemCom.

2.1. Classical Information Theory vs. Semantic Information

The foundational distinction lies in what each theory attempts to quantify and optimize. The shift from classical to semantic communication is defined by the three levels of communication addressed in the Shannon-Weaver framework:

Table 1. Summary of key semantic communication surveys

Ref	Focus Area	Contributions	Gaps
[19]	Architecture & vision	Proposes intellicise framework, treating semantics as knowledge. Discusses network intelligence and ambient cognition.	Lack of concrete metric design; unclear system-level runtime feasibility.
[20]	Internet evolution	Introduces task-oriented semantics and deep context reasoning for future internet.	Does not detail edge/cloud orchestration or adaptive runtime support.
[21]	Technical implementation	Covers semantic encoding, intelligent resource allocation, and layered SemCom protocols.	Energy-aware and delay-constrained mechanisms not fully explored.
[22]	Metrics & goals	Semantic distortion, goal-completion rate, and alignment-based metrics discussed.	No standard datasets or benchmarking frameworks provided.
[24]	Knowledge representations	Explores hybrid symbolic/sub-symbolic semantics and AI integration.	Lacks system-wide application context and runtime profiling.
[25]	Edge learning	Discusses edge inference, semantic fusion, and wireless sensing in beyond fifth-generation (B5G).	No runtime-aware control taxonomy or scheduling design.
[26]	Engineering perspective	Analyses trade-offs in model complexity, deployment feasibility, and semantic fidelity.	No detailed integration with AI/machine learning (ML)-driven network control.
[27]	Metaverse and XR	Explores SemCom for metaverse including trust, latency, and immersive personalization.	Absence of quantitative performance benchmarks.
[28]	GenAI and SemCom	Highlights LLMs, diffusion models, and generative semantics for 6G.	Challenges remain in semantic calibration and adversarial robustness.

- *Classical Communication Focus: The Syntactic Layer (Level A)*: Shannon's information theory addresses the Syntactic Layer, prioritizing accurate symbol transmission from sender to receiver, independent of meaning [32]. The technical problem involves determining how accurately communication symbols can be transmitted. It is measured by syntactic reliability, focusing on minimizing bit error rate (P_e) or maximizing channel capacity, with equal importance given to all bits or symbols.

- *SemCom Focus: Semantic and Pragmatic Layers (Levels B & C)*: SemCom enhances the classical model by tackling the Semantic Layer (Level B) and Pragmatic/Effectiveness Layer (Level C). The Semantic Layer ensures symbols accurately convey meaning, relying on shared knowledge to boost semantic fidelity [33]. The Pragmatic Layer evaluates how well the meaning influences behavior, focusing on maximizing message utility or task success rather than symbol accuracy, making communication goal-oriented.

Classical theory focuses on the syntactic layer (Level A: Accurate symbol transmission), while semantic theory deals with both the semantic layer (Level B: Precise meaning conveyance) and the pragmatic layer (Level C: Effective influence on behavior). Table 2 compares the key characteristics of classical and semantic information theories in terms of focus, measure, optimization, receiver, and channel model.

2.2. Task-Oriented vs. Source-Oriented Communication

- *Source-Oriented Communication (Classical):* The goal is to reconstruct the source signal/data as faithfully as possible at the receiver, aiming for high fidelity in terms of mean squared error (MSE) or symbol error rate (SER) [34].

Table 2. Comparison of Classical and Semantic Information Theories

Feature	Classical (Shannon) Information Theory	Semantic Information Theory
Focus	*Syntactic reliability* (accurate symbol transmission).	*Meaning and utility* (successful interpretation and impact).
Measure	*Bits* (reduction of uncertainty at the symbol level).	*Semantics* (reduction of uncertainty about the *meaning*).
Optimization	Maximize *channel capacity*; minimize *error probability* (P_e).	Maximize *semantic fidelity*; maximize *task success probability*.
Receiver	*Decodes symbols* (recovers the sent bit sequence).	*Interprets meaning* and *executes tasks* (uses the meaning).
Channel Model	Bit-level noise (e.g., additive white gaussian noise (AWGN)).	*Semantic noise* (misinterpretation, lack of context).

- *Task-Oriented Communication (Semantic/Pragmatic):* The goal is to enable the successful execution of a specific task at the receiver by transmitting only the minimum necessary semantic information [35].

The task-oriented approach uses TSR or task utility instead of fidelity metrics. In automated driving, conveying key semantic features like "Obstacle detected at coordinates (x, y)" is more efficient than fully reconstructing raw sensor data [36].

2.3. Semantic Information Measures

SemCom necessitates new metrics to quantify meaning and utility, moving beyond the Shannon entropy measure ($H(X)$).

Semantic Entropy and Semantic Distortion.

- *Semantic Entropy (H_S):* Quantifies the uncertainty regarding the meaning or intent of the source message. It weighs symbols based on their relevance to the underlying meaning or task [37].
- *Semantic Distortion (D_S):* Measures the gap between the source's intended meaning/utility and what the receiver extracts [38]. Distortion can remain low despite high symbol distortion (D) if key semantic features are preserved.

Utility-Aware vs. Accuracy-Aware Metrics.

- *Accuracy-Aware Metrics:* Focus on the syntactic accuracy of the reconstructed signal (MSE, SER).
- *Utility-Aware Metrics (Pragmatic/Task-Oriented):* Focus on the value of the received information to the receiver's specific goal, often defined as

$$U(\hat{X}) = f(\text{task success}, \text{resource cost}) \qquad (1)$$

  For a real-time control loop, utility U might be defined as: $U = \text{Accuracy in control action} - \lambda \cdot \text{Delay}$. System design becomes an end-to-end optimization maximizing U [39].

2.4. Key Concepts

Representation of Knowledge (Symbolic vs. Subsymbolic). The knowledge base (KB) is shared between the transmitter and receiver.

- *Symbolic Representation:* Uses discrete symbols and logical rules (e.g., knowledge graphs). Example: Car $\rightarrow$ Vehicle $\wedge$ Has_Wheels [40].
- *Subsymbolic Representation:* Uses continuous vectors (embeddings) derived from deep learning models (e.g., transformer models) to represent complex, nuanced meanings implicitly.

Semantic Noise, Grounding, Contextuality.

- *Semantic Noise:* The loss or corruption of meaning due to differences in the KBs, context, or interpretation models between the transmitter and receiver [26].
- *Grounding:* The process of associating abstract symbols/representations with perceptual reality or physical actions. A message is grounded if its meaning can be unambiguously translated into a real-world state or action [41].
- *Contextuality:* The meaning of a message is dependent on the shared context (time, location, prior events). Systems must dynamically adapt encoding/decoding based on context [42].

Semantics-Pragmatics-Action Loop. This loop transforms the communication system into a goal-directed intelligence component:

1. *Semantics (Meaning):* The receiver extracts the *meaning* ($\hat{M}$) from the received symbols.
2. *Pragmatics (Utility/Action Selection):* Based on $\hat{M}$ and context, the receiver determines the optimal action ($\hat{A}$) to maximize utility [43].
3. *Action (Execution):* The receiver executes the chosen action ($\hat{A}$).

4. *Feedback (New Context/State):* The action's result changes the environment and establishes a new context for subsequent communication, closing the loop.

3. Taxonomy of Semantic Communications in 6G

SemCom in 6G introduces a layered and function specific taxonomy. This taxonomy dictates the encoding, processing, transmission, and interpretation of meaning, context, and task relevance in complex wireless networks [44, 45]. This section introduces a classification based on five axes: representation, modeling, control, learning, and deployment. Table 3 presents a taxonomy of SemCom in 6G, highlighting the main categories, core components, and representative use cases.

3.1. Semantic Representation

Semantic information in 6G networks is modeled through diverse representations:

- *Embeddings:* Low-dimensional vector representations capture latent semantic features [46]. Techniques include word2vec-style encoding, contextualized transformer embeddings, and vision-language representations for multimodal signals [47].
- *Knowledge Graphs:* Symbolic, structured knowledge bases that describe real-world entities and their semantic relationships [48]. Ideal for goal-driven reasoning, dynamic task inference, and explainability.
- *Latent World Models:* Probabilistic generative structures such as variational autoencoders (VAEs), diffusion models, and neural world models enable predictive understanding of task environments with uncertainty quantification [49].

3.2. Encoder–Decoder Models

Encoders and decoders in SemCom systems perform task-aware compression and semantic abstraction:

- *Semantic Compression:* Eliminates irrelevant signal components, retaining only task-relevant meaning (e.g., key object classes or intentions in AVs) [7].
- *Uncertainty Modeling:* Encoders may quantify semantic confidence (entropy, distributional divergence) to prioritize critical information and adapt to noise or ambiguity.
- *Generative Semantics:* Decoders increasingly leverage GenAI (e.g., LLMs, VAEs) to reconstruct or predict intended meaning, bridging gaps in incomplete or noisy input [50].

3.3. Semantic Control

Effective semantic communication depends on intelligent control strategies that manage when, what, and how to transmit:

- *Task Scheduling:* Prioritization based on semantic importance, urgency, and downstream impact (e.g., safety-critical vs. informational updates) [51].
- *Resource Allocation:* Dynamic spectrum and energy assignment guided by semantic utility, not raw bit-rate efficiency [52].
- *Feedback Loops:* Real-time semantic feedback (e.g., task status, confidence scores) drives adaptive modulation, retransmission, or knowledge refinement.

3.4. Learning Feedback Loops

Learning mechanisms are central to adapting SemCom systems to dynamic environments:

- *Reinforcement Learning:* Agents learn semantic policies for compression, control, and decoding based on task rewards [13].
- *Federated and Continual Learning:* Distributed agents collaboratively update semantic models while preserving local context and privacy.
- *Active Perception:* Systems selectively acquire the most informative observations guided by learned semantic utility functions [53].

3.5. Deployment Topologies

The implementation of semantic systems varies across network architectures:

- *Edge/Cloud-Native Topologies:* SemCom modules co-exist with edge inferencing, caching, and orchestration frameworks for low-latency applications [54].
- *SAGIN Integration:* SAGIN uses SemCom to minimize bandwidth and delay in delay-tolerant or mission-critical environments [13].
- *Quantum-Assisted SemCom Architectures*: explores quantum-assisted SemCom systems, highlighting quantum technologies' role in boosting semantic transmission [55]. The system uses semantic encoders/decoders handle content, with a quantum processor uses a semantic channel for learning, optimization, and inference. This hybrid model aims to surpass traditional models by using quantum machine learning (QML) for context-aware message reconstruction and better communication in dynamic networks [56].

Table 3. Taxonomy of SemCom in 6G: Categories, Core Components, and Use Cases

Category	Core Components	Practical Use Cases
Semantic Representation	Embeddings, knowledge graphs, latent world models	Digital twins, AV reasoning, smart home agents
Encoder–Decoder Models	Semantic compression, uncertainty modeling	XR streaming, anomaly detection, task-driven inference pipelines
Semantic Control	Scheduling, prioritization, semantic feedback loops	Edge orchestration, emergency alerts in vehicular networks
Learning Loops	Reinforcement learning agents, federated learning, continual adaptation	Industrial IoT, healthcare monitoring, adaptive XR experiences
Deployment Topologies	Edge/cloud-native deployments, SAGIN, RIS-assisted architectures	Remote monitoring, resilient 6G networks, metaverse and XR integration

Figure 3 shows how classical SemCom operates in parallel with a quantum-enhanced semantic channel, forming a hybrid system that improves intelligence and efficiency in communication.

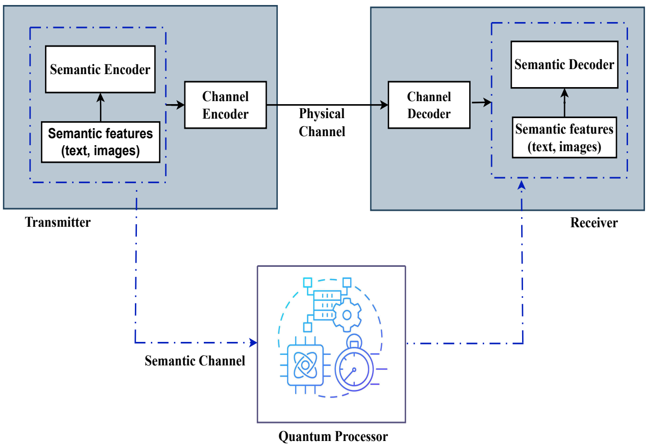

Figure 3. The SemCom with QML layers and connectivity.

4. Semantic Models and Protocol Architectures

SemCom in 6G requires models and protocols that jointly capture meaning, context, and task relevance across the network stack [57]. This section outlines five major pillars: knowledge-driven models, deep generative architectures, foundation models, protocol layering, and cross-layer co-optimization mechanisms [50].

4.1. Knowledge-Driven Models: Ontologies and Reasoning Engines

Knowledge-driven models form the conceptual backbone of semantic communication, emphasizing interpretable and domain-specific reasoning [7]. Ontologies define structured vocabularies of entities, relationships, and contexts that enable consistent understanding between communicating agents [48]. For instance, in vehicular networks, an ontology may formalize terms such as pedestrian, traffic signal, and collision risk to ensure unified interpretation among vehicles and infrastructure nodes.

Reasoning engines use symbolic logic, rule-based systems, or hybrids to infer uncertain meanings [3]. Operating at the edge, they offer context-aware decisions like warnings or route adjustments. Future dynamic ontologies could evolve via federated learning or cooperative updates, ensuring adaptability and interoperability across devices and environments [58].

4.2. Deep Generative Models: GANs, VQ-VAE, and Diffusion

Deep generative models transform how semantics are encoded, compressed, and reconstructed [59]. Instead of transmitting raw data, these models learn latent representations that preserve essential meaning:

- *GANs:* Generative Adversarial Networks (GANs) efficiently create realistic reconstructions with low bit budgets [60]. This is ideal for XR and metaverse streaming where minimal bandwidth is needed.
- *VQ-VAE:* Vector Quantized VAE (VQ-VAE) discretizes latent spaces into compact tokens, allowing real-time semantic encoding of visual, audio, or multimodal data in distributed inference systems [61].

- *Diffusion Models:* These models iteratively refine noisy signals into meaningful outputs, providing robust semantic recovery under lossy channels or interference [47]. They are particularly useful in applications requiring resilient content restoration.

By jointly training generative encoders and channel models, future 6G systems can achieve semantic-channel co-adaptation, where encoding strategies evolve alongside network conditions.

4.3. Foundation Models: LLMs and VLMs for Multimodal SemCom

Foundation models, including LLMs and vision-language models (VLMs), are pivotal to multimodal semantic communication [62]. They integrate diverse modalities, such as text, audio, images, and sensor data, into unified embedding spaces, facilitating shared understanding across heterogeneous devices [63].

LLMs as Semantic Engines: Large models deployed at the edge or core can perform summarization, translation, and intent extraction, drastically reducing redundant transmission [64]. Instead of raw sensor feeds, an edge LLM may send a compressed summary such as “no anomalies detected in turbine 3”.

VLMs and Multimodal Reasoning: Models like CLIP and Flamingo connect text and visuals for XR, autonomous driving, and surveillance [65]. Fine-tuned versions, like MobileLLMs, operate on edge processors or RIS for ultra-low-latency decisions. These models create a new standard for semantic interoperability, exchanging information from various sensors as unified, context-rich embeddings [66].

4.4. Protocol Layering: Where Semantics Reside in the Stack

Semantic functions permeate multiple protocol layers, reshaping traditional network operations:

- *Physical Layer:* Semantics may guide symbol mapping or power allocation, prioritizing meaning-bearing signals (e.g., emergency alerts) with higher transmission reliability [67].
- *MAC Layer:* Scheduling policies evolve from throughput maximization to semantic relevance, transmitting the most meaningful information first.
- *Network Layer:* Semantic-aware routing enables context-optimized paths, reducing end-to-end delay for critical intent-driven communication [47].
- *Application Layer:* High-level semantic processing, knowledge integration, and goal inference occur here, but cross-layer application programming interfaces (APIs) ensure lower layers are informed by semantic context.

Emerging concepts such as semantic overlay networks and semantic APIs allow decoupling between knowledge control and data transport, paving the way for adaptive, meaning-centric communication stacks.

4.5. Cross-Layer Optimization: Joint Codec-Inference-Scheduler Design

True semantic efficiency emerges from cross-layer integration of encoding, inference, and scheduling mechanisms:

- *Joint Codec-Inference Optimization:* Semantic encoders dynamically adapt compression levels based on downstream inference requirements or task success probability [68].
- *Semantic-Aware Scheduling:* Scheduler decisions prioritize packets with higher expected task utility or goal completion probability, moving beyond traditional quality of service (QoS)-based metrics [69].
- *Feedback Integration:* Semantic acknowledgment signals confirming meaning comprehension or action success enabled closed-loop adaptation of encoding and modulation strategies.
- *Hardware-In-Loop Co-Design:* Neuromorphic and neural processing unit (NPU)-based co-processors can execute real-time semantic inference while influencing radio-layer behavior, leading to a new class of semantic software-defined networking (SDN) controllers.

6G networks will become autonomous ecosystems, dynamically balancing accuracy, latency, and energy while aligning with human and machine intent [70]. Semantic 6G transforms communication to focus on meaning rather than data by using symbolic reasoning, generative representation, multimodal models, and cross-layer adaptability, enabling devices to share understanding [71].

5. Semantic Metrics and Evaluation Benchmarks

5.1. Beyond BER/BLER: Toward Meaning-Aware Performance Metrics

Traditional metrics like bit error rate (BER) and block error rate (BLER) don’t capture the performance of meaning-centric systems [72]. Semantic communication systems need new metrics to assess how well the intended meaning is reconstructed for tasks. Emerging metrics include:

- *TSR*: Evaluates whether the communication outcome enables successful completion of the target task, such as object detection in AVs or medical alerting in healthcare [73].
- *Semantic Fidelity*: Measures the alignment between the transmitted and received meanings, potentially through embedding similarity, concept graph overlap, or entailment scores.
- *Utility-Aware Distortion*: Quantifies performance degradation based on its practical utility impact, such as semantic loss leading to incorrect robotic actuation [49].
- *Safety Risk Index (SRI)*: Introduces weighted penalties for semantic errors that could lead to harmful or unsafe outcomes in critical domains like autonomous driving or industrial control [74].
- *Contextual Relevance*: Assesses the alignment of received data with environmental context or user intent, important in dynamic tasks like XR or remote collaboration [75].

Figure 4 gives the summary of the SemCom metrics and evaluation benchmarks for real-time testing.

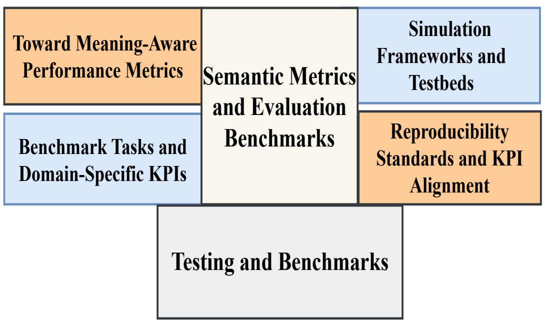

Figure 4. SemCom metrics and evaluation benchmarks.

5.2. Benchmark Tasks and Domain-Specific KPIs

To ensure consistent and meaningful evaluation, domain-aligned benchmark tasks must be established, each linked with its own set of semantic key performance indicators (KPIs):

- *AVs*: Tasks include semantic map sharing, perception fusion, and cooperative maneuvering. KPIs: inference latency, goal completion rate, collision avoidance success [76].
- *XR*: Evaluation of semantic compression and stream fidelity during real-time interactions. KPIs: perceptual accuracy, interaction latency, frame prediction robustness.
- *Healthcare and Remote Monitoring*: Focus on semantic alert generation and multimodal data fusion (e.g., ECG + NLP reports). KPIs: diagnostic correctness, alert precision, delay sensitivity [77].
- *Collaborative Robotics and Industrial IoT*: Benchmarks include semantic negotiation, shared task planning, and swarm control. KPIs: alignment success, coordination score, semantic drift index [78].

5.3. Simulation Frameworks and Testbeds

To support repeatable and realistic evaluation, simulation testbeds and frameworks should offer:

- *SemCom-SIM*: A modular, open-source simulation framework integrating 6G channel models, semantic encoding modules, and task-specific evaluators [74].
- *S3Bench (Semantic Simulation and Sensing Benchmark)*: Includes annotated multimodal datasets, pre-built encoding stacks, and unified scripts for performance evaluation [79].
- *Digital Twin Testbeds*: Edge-to-cloud platforms modeling real-world SemCom deployments in AV fleets, industrial plants, or smart homes.
- *Integration with AI SDKs*: Plug-and-play compatibility with platforms like NVIDIA Clara (healthcare), Unity/Unreal (XR), or multimodal agents [80].

5.4. Reproducibility Standards and KPI Alignment

For robust progress in semantic communication research, the following guidelines are proposed:

- *Evaluation Protocol Templates*: Provide standardized configurations, models, and test routines for fair comparison across publications [81].
- *KPI Taxonomy*: Define and classify semantic KPIs across domains by criticality (e.g., safety-critical vs non-critical), modality (text/image/sensor), and interpretability [67].
- *Model Zoo and Task Sharing*: Host pretrained semantic encoders, semantic utility functions, and evaluation harnesses for public reproducibility.
- *Semantic Fidelity Audits*: Implement semantic validation pipelines, possibly using knowledge reasoning, to ensure grounded of received outputs [82].

5.5. Testing and Benchmarks

To further enrich benchmarking:

- *Semantic Fuzz Testing*: Inject adversarial but semantically plausible perturbations to assess model robustness [83].
- *Federated Semantic Benchmarks*: Evaluate SemCom generalization across distributed, non-independent and identically distributed (IID) semantic contexts in federated environments.
- *Green SemCom KPIs*: Introduce energy-to-utility ratios or carbon cost per successful semantic task as sustainability-aware metrics [84].

6. System-Level Design and Implementation

SemCom at the system level creates end-to-end pipelines to turn data acquisition into task-oriented feedback [85]. It integrates algorithms with hardware, optimizes energy-latency, and ensures sustainable deployment across terrestrial, aerial, and orbital infrastructures.

6.1. End-to-End Pipelines: Acquisition to Task Feedback

The semantic pipeline includes five stages: (1) Data acquisition from diverse sensors (e.g., cameras, LiDAR, EEG); (2) Semantic encoding, compressing data into latent variables via advanced models [3]; (3) Task-optimized transmission; (4) Inference and action for autonomous decisions; and (5) Feedback for adaptive model retraining [86]. Unlike traditional systems, these pipelines focus on goals, learning from feedback to enhance semantic efficiency and adapting to context and priorities.

6.2. Hardware Integration: NPUs, Edge Chips, RIS, and SemCom Radios

Implementing semantic pipelines needs specialized hardware like NPUs and AI accelerators for efficient on-device semantic compression and feature extraction [13]. Edge AI chips, like NVIDIA Jetson or Google Coral, facilitate real-time data processing, reducing delays [87]. RIS enhances channel intelligence by dynamically managing critical streams. SemCom-aware radios, a new hardware class, optimize bandwidth and coding based on semantic scores [88]. As 6G evolves, future chips may integrate "semantic control units" for intent-driven packet management.

6.3. Latency–Energy–Utility Tradeoffs

Balancing latency, energy efficiency, and semantic utility is crucial. High-fidelity models boost performance but use more power, while lightweight encoders save energy but risk losing information [89]. Adaptive frameworks that adjust model complexity, bit allocation, and offloading strategies based on context are needed. The optimal balance depends on the task: XR streaming prioritizes low latency, while industrial monitoring focuses on energy efficiency [90]. Multi-objective optimization can adjust transmission, resolution, and channel allocation with real-time feedback.

6.4. Carbon-Aware Design

6G networks aim for sustainability, making carbon-aware semantic design vital. Systems should consider carbon footprints in computation, transmission, and model inference [91]. Carbon-aware scheduling can prioritize low-emission tasks, delaying non-urgent transmissions during peak energy demand. Future SemCom orchestrators might use carbon intensity forecasts, dynamic semantic throttling, carbon-tagged packets, and energy certificates for eco-friendly data exchange [92].

6.5. Deployment Case Studies: Edge, RIS, UAVs, and LEO Satellites

Semantic pipelines can be deployed across multiple network domains:

- *Edge-native Vehicular SemCom*: Cooperative AVs share high-level semantic maps and driving intentions via edge servers, reducing bandwidth while improving safety [17].
- *RIS-assisted Smart Hospitals*: RIS panels prioritize critical telemetry signals and emergency alerts within dynamic hospital networks [93].
- *UAV-Relayed Emergency Networks*: unmanned aerial vehicles (UAVs) equipped with SemCom relays enable situational awareness and coordination among rescue units with limited spectrum resources [94].
- *LEO Satellite SemCom*: Satellite constellations facilitate global semantic synchronization for environmental monitoring, disaster management, and maritime safety.

Each deployment highlights the transformative potential of meaning-centric communication across heterogeneous, distributed, and resource-constrained environments.

7. Applications and Use Cases

SemCom supports 6G applications by enhancing performance through context, meaning, and task intent. Below are key use cases and their practical implications.

7.1. Vehicular Networks (AV/IoV)

In AVs and internet of vehicles (IoV), semantic communication uses situational data like "pedestrian approaching at 4 km/h from right" instead of raw sensor data, enhancing latency, interpretability, and cooperative decisions [95]. Semantic reinforcement learning (RL) agents focus on critical data in traffic. Federated SemCom frameworks allow vehicles to improve encoders with aggregated feedback while maintaining data privacy [96].

7.2. Metaverse, XR, and Perceptual SemCom

The metaverse uses synchronized multimodal perception among users. Semantic compression transmits perceptual streams as structured data instead of raw pixels [97]. In XR, salient features are shared to lower rendering demands. Perceptual SemCom combines gaze tracking, gestures, and emotions to predict user intent, offering responsive experiences with less bandwidth and energy use [98]. Neuro-symbolic encoding, akin to human cognition, may underpin real-time semantic exchange in 6G immersive applications.

7.3. Industrial IoT and Predictive Maintenance

SemCom revolutionizes industrial IoT (IIoT) by shifting from data collection to sharing insights. Smart factories use semantic sensors to identify and report patterns like "motor vibration anomaly" instead of raw data [99]. Semantic encoders, trained on diverse states, support predictive maintenance, reducing downtime. Federated learning enables secure cross-factory semantic sharing, improving generalization across machines. Thus, task-driven SemCom is key to autonomous, resilient industry 5.0 ecosystems [100].

7.4. Disaster Response and eHealth

In disaster response, SemCom enables efficient coordination with semantic mapping. UAVs and robots send concise "mission semantics" like "blocked road ahead, rerouting needed," ensuring timely updates despite unstable connections [101]. In healthcare, SemCom shares critical patient indicators (e.g., "arrhythmia detected") instead of full data streams, reducing data exposure and ensuring real-time telemedicine responses. Semantic reasoning allows adaptive alert prioritization and context-sensitive triage in varying network conditions [102].

7.5. Privacy-Conscious SemCom

SemCom enhances privacy by sending task-relevant meanings instead of raw data. In surveillance, saying "object-of-interest detected" protects privacy more than video feeds. In eHealth, semantic summaries like "risk level: high, condition: irregular heart rhythm" avoid sharing sensitive data [103]. Semantic watermarking tracks shared meanings for accountability and compliance. Future systems may verify intent to ensure ethical and legal meaning exchange. System-level SemCom integrates meaning-driven data processing, hardware co-design, energy optimization, and ethical deployment across domains [104]. Using semantic pipelines, intelligent edge devices, and sustainable architectures, 6G communication will advance from data transfer to goal realization, fostering resilient, efficient, and ethically responsible information ecosystems.

8. Open Challenges and Research Directions

Despite advances in SemCom, realizing its full potential in 6G networks requires foundational breakthroughs and system-level innovation.

8.1. Semantic Reliability and Theoretical Foundations

Current SemCom literature lacks a solid theoretical basis for semantic reliability. Traditional metrics like BER/BLER don't adequately assess meaning preservation in task-focused scenarios [105]. A clear definition of semantic fidelity, based on probabilistic reasoning and formal semantics, is essential to measure meaning retention through transmission, especially when facing noise, compression, and contextual ambiguity.

8.2. Multi-Agent Semantic Negotiation

6G scenarios like cooperative AVs, swarm robotics, and digital twins require multi-agent semantic negotiation for dynamic alignment of meanings [106]. This requires adaptive protocols for real-time negotiation, conflict resolution, and consensus in shared semantic spaces [107]. Research should focus on decentralized knowledge alignment, context inference, and negotiation policies using game theory and RL.

8.3. Feedback Loops and Convergence Stability

SemCom systems employ closed-loop designs where inference adjusts encoder adaptation [108]. Ensuring stable convergence in these feedback loops is challenging, especially in dynamic, continually learning environments [109]. Future research should examine theoretical guarantees, control mechanisms, and performance limits of feedback-augmented semantic pipelines.

8.4. Green SemCom and Carbon-Aware Prioritization

SemCom should emphasize energy efficiency and sustainability. Current research often ignores the carbon footprint of semantic inference and transmission [110].

Future studies should explore carbon-aware scheduling, prioritize low-impact tasks, and adjust semantic model selection dynamically based on energy, hardware, and social policies.

8.5. Integration with Quantum and Bio-Inspired Networks

Emerging paradigms like quantum communication and bio-inspired nanonetworks present intriguing possibilities for SemCom [111]. Quantum-enhanced encoding, entangled meaning transfer, and hybrid feedback remain largely unexplored. Bio-semantic interfaces mimicking neural or genetic communication could enable efficient and self-organizing semantic systems.

8.6. Semantic Knowledge Lifecycle and Interoperability

SemCom knowledge representations like ontologies, embeddings, and graphs must evolve, yet the semantic lifecycle is often overlooked [112, 113]. Research should tackle version control, decay, provenance, and adaptive forgetting. Interoperability among vendors, models, and platforms is a key challenge, necessitating standard interfaces, semantic APIs, and cross-domain mapping protocols.

9. Conclusion

SemCom shifts from symbol transmission to meaning-focused interaction, aligning protocols with tasks, context, and user intent. This survey covers theories, models, architectures, control, learning strategies, deployments, security, and evaluation for semantic-native 6G networks. SemCom integrates semantic intelligence across the wireless stack, enabling perceptual, proactive, and goal-oriented messaging beyond traditional bitstream logic. The future vision involves creating scalable, secure, carbon-aware ecosystems where devices share meaning, using symbolic reasoning, deep generative models, edge-native intelligence, and quantum/bio-inspired frameworks. SemCom transforms communication by making meaning central, paving the way for resilient and context-adaptive networks.

References

[1] Z. Qin, F. Gao, B. Lin, X. Tao, G. Liu, and C. Pan, "A generalized semantic communication system: From sources to channels," *IEEE Trans. Wireless Commun.*, vol. 30, no. 3, pp. 18–26, Jun. 2023.

[2] L. Yan, Z. Qin, R. Zhang, Y. Li, and G. Y. Li, "Resource allocation for text semantic communications," *IEEE Wireless Commun. Lett.*, vol. 11, no. 7, pp. 1394–1398, Jul. 2022.

[3] Z. Weng and Z. Qin, "Semantic communication systems for speech transmission," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2434–2444, Aug. 2021.

[4] P. Porambage, G. Gür, D. P. M. Osorio, M. Liyanage, A. Gurtov, and M. Ylianttila, "The roadmap to 6G security and privacy," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1094–1122, May 2021.

[5] J. A. Ansere, S. C. Prabhashana, N. Simmons, O. A. Dobre, H. Shin, and T. Q. Duong, "Quantum deep deterministic policy gradient for digital twin-enabled semantic IoV networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 13, pp. 54–66, Jun. 2025.

[6] Y. Liu, H. Chen, and L. Wang, "Physical layer security for next generation wireless networks: Theories, technologies, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 347–376, Jan. 2017.

[7] X. Luo, H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Trans. Wireless Commun.*, vol. 29, no. 1, pp. 210–219, Feb. 2022.

[8] D. Huang, F. Gao, X. Tao, Q. Du, and J. Lu, "Toward semantic communications: Deep learning-based image semantic coding," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 55–71, Jan. 2023.

[9] X. Mu and Y. Liu, "Exploiting semantic communication for non-orthogonal multiple access (NOMA)," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2563–2576, Aug. 2023.

[10] H. Zhang, S. Shao, M. Tao, X. Bi, and K. B. Letaief, "Deep learning-enabled semantic communication systems with task-unaware transmitter and dynamic data," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 170–185, Jan. 2023.

[11] H. Xie, Z. Qin, G. Y. Li, and B. Juang, "Deep learning enabled semantic communication systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2663–2675, Apr. 2021.

[12] T. Tung, D. B. Kurka, M. Jankowski, and D. Gündüz, "DeepJSCC-Q: Constellation constrained deep joint source-channel coding," *IEEE J. Sel. Areas Inf. Theory*, vol. 3, no. 4, pp. 720–731, Dec. 2022.

[13] Z. Weng, Z. Qin, X. Tao, C. Pan, G. Liu, and G. Y. Li, "Deep learning enabled semantic communications with speech recognition and synthesis," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 6227–6240, Sep. 2023.

[14] V. Nguyen, P. Lin, B. Cheng, R. Hwang, and Y. Lin, "Security and privacy for 6G: A survey on prospective technologies and challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 4, pp. 2384–2428, Dec. 2021.

[15] S. M. A. Rizvi, U. Khalid, S. Chatzinotas, T. Q. Duong, and H. Shin, "Controlled quantum semantic communication for industrial CPS networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 13, pp. 996–1009, Jul. 2025.

[16] H. Xie and Z. Qin, "A lite distributed semantic communication system for Internet of things," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 142–153, Jan. 2021.

[17] P. Singh, W.-J. Huang, B. Hazarika, K. Singh, and T. Q. Duong, "Semantic-aware priority-based resource allocation for C-V2X platoons using transformer encoding," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Taipei, Taiwan, Dec. 2025, pp. 1–6.

[18] C. D. Alwis, "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," *IEEE Open J. Commun. Soc.*, vol. 2, no. 4, pp. 836–886, Apr. 2021.

[19] P. Zhang, W. Xu, and Y. L. et al., "Intellicise wireless networks from semantic communications: A survey, research issues, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 27, no. 3, pp. 2051–2084, Aug. 2025.

[20] W. Yang, H. Du, Z. Q. Liew, W. Y. B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Shen, and C. Miao, "Semantic communications for future internet: Fundamentals, applications, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 213–250, Nov. 2023.

[21] Y. Wang, H. Han, Y. Feng, J. Zheng, and B. Zhang, "Semantic communication empowered 6G networks: Techniques, applications, and challenges," *IEEE Access*, vol. 13, pp. 28 293–28 314, Jan. 2025.

[22] T. M. Getu, G. Kaddoum, and M. Bennis, "Making sense of meaning: A survey on metrics for semantic and goal-oriented communication," *IEEE Access*, vol. 11, pp. 45 457–45 492, Apr. 2023.

[23] ——, "A survey on goal-oriented semantic communication: Techniques, challenges, and future directions," *IEEE Access*, vol. 12, pp. 51 225–51 274, Mar. 2024.

[24] Z. Lu, R. Li, K. Lu, X. Chen, E. Hossain, L. Song, D. Niyato, X. Shen, and Z. Han, "Semantics-empowered communications: A tutorial-cum-survey," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 1, pp. 41–79, Nov. 2024.

[25] W. Xu, Z. Yang, D. W. K. Ng, M. Levorato, Y. C. Eldar, and M. Debbah, "Edge learning for B5G networks with distributed signal processing: Semantic communication, edge computing, and wireless sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 1, pp. 9–39, Jan. 2023.

[26] D. Wheeler and B. Natarajan, "Engineering semantic communication: A survey," *IEEE Access*, vol. 11, pp. 13 965–13 995, Feb. 2023.

[27] M. Z. Aloudat, A. Aboumadi, A. Soliman, H. A. Al-Mohammed, M. Al-Ali, A. Mahgoub, M. Barhamgi, and E. Yaacoub, "Metaverse unbound: A survey on synergistic integration between semantic communication, 6G, and edge learning," *IEEE Access*, vol. 13, pp. 58 302–58 350, Mar. 2025.

[28] A. Celik and A. M. Eltawil, "At the dawn of generative AI era: A tutorial-cum-survey on new frontiers in 6G wireless intelligence," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 2433–2489, Feb. 2024.

[29] B. Narottama, A. U. Haq, J. A. Ansere, N. Simmons, B. Canberk, S. L. Cotton, H. Shin, and T. Q. Duong, "Quantum deep reinforcement learning for digital twin-enabled 6G networks and semantic communications: Tutorial on adoptions and security," *IEEE Trans. Netw. Sci. Eng.*, vol. 13, pp. 2053–2076, Sep. 2025.

[30] B. E. Arfeto, S. Tariq, U. Khalid, T. Q. Duong, and H. Shin, "GenSC-6G: A prototype testbed for integrated generative AI, quantum, and semantic communication," *IEEE Commun. Mag.*, vol. 63, no. 10, pp. 28–35, Oct. 2025.

[31] L. Hu, L. Yu, and Z. Qin, "Deep learning-based semantic communication system for wireless image transmission," *IEEE Wireless Commun. Lett.*, vol. 14, no. 8, pp. 2391–2395, Aug. 2025.

[32] X. Huang, C. Wang, X. Tian, Z. Li, C. Zhao, and M. Xiao, "A multi-hop semantic communication system enhanced by semantic importance," *IEEE Access*, vol. 13, pp. 140 685–140 693, Sep. 2025.

[33] C. Xing, J. Lv, T. Luo, and Z. Zhang, "Multi-level similarity for efficient compression in graph-based multi-modal semantic communication," *IEEE Commun. Lett.*, vol. 29, no. 9, pp. 2078–2082, Sep. 2025.

[34] M. Tao, J. Fan, J. Luo, and H. Xie, "Coarse-to-fine semantic communication systems for text transmission," *IEEE Trans. Veh. Technol.*, vol. 74, no. 8, pp. 13 267–13 271, Aug. 2025.

[35] Z. Jin, T. Song, W. Jia, W. Zou, and X. Song, "Task-oriented semantic communication with adaptive semantic reconstruction network," *IEEE Internet Things J.*, vol. 12, no. 17, pp. 35 784–35 798, Sep. 2025.

[36] B. Tang, L. Huang, Q. Li, A. Pandharipande, and X. Ge, "Cooperative semantic communication with on-demand semantic forwarding," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 349–363, Jan. 2024.

[37] Q. Fu, H. Xie, Z. Qin, G. Slabaugh, and X. Tao, "Vector quantized semantic communication system," *IEEE Wireless Commun. Lett.*, vol. 12, no. 6, pp. 982–986, Jun. 2023.

[38] M. Kim and D. Ji, "Fully learnable multi-rate quantization for digital semantic communication systems," *IEEE Wireless Commun. Lett.*, vol. 14, no. 9, pp. 2848–2851, Sep. 2025.

[39] W. Xie, T. Zhang, M. Xiong, J. Wang, and L. Yang, "Semantic communication system based on meta-learning framework," *IEEE Commun. Lett.*, vol. 29, no. 7, pp. 1684–1688, Jul. 2025.

[40] Y. Peng, J. Liu, Y. Zhu, R. Zhang, and F. Yuan, "Efficient semantic communication for underwater images guided by physical priors," *IEEE Signal Process. Lett.*, vol. 32, pp. 3072–3076, Jan. 2025.

[41] J. A. Ansere, S. C. Prabhashana, O. A. Dobre, and T. Q. Duong, "Quantum machine learning DDPG for digital twin semantic vehicular networks," in *Proc. IEEE Int. Conf. Machine Learning Commun. Netw. (ICMLCN)*, Barcelona, Spain, May 2025, pp. 1–6.

[42] X. Chen, C. Huang, G. Chen, D. Feng, and P. Xiao, "The communication and computation trade-off in wireless semantic communications," *IEEE Wireless Commun. Lett.*, vol. 14, no. 7, pp. 2259–2263, Jul. 2025.

[43] Z. Sun, S. Ma, and S. Li, "Task-oriented semantic communication with importance-aware rate control," *IEEE Commun. Lett.*, vol. 29, no. 7, pp. 1520–1524, Jul. 2025.

[44] J. A. Ansere, M. Kamal, I. A. Khan, and M. N. Aman, "Dynamic resource optimization for energy-efficient 6G-IoT ecosystems," *Sensors*, vol. 23, no. 10, p. 4711, May 2023.

[45] Y. Y. Sun, J. J. Liu, J. D. Wang, Y. R. Cao, and N. Kato, "When machine learning meets privacy in 6G: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2694–2724, Dec. 2021.

[46] L. V. Nguyen, T. T. Nguyen, O. A. Dobre, and T. Q. Duong, "Leveraging stable diffusion with context-aware prompts for semantic communication," in *Proc.*

IEEE Int. Conf. Mobile Ad-Hoc. Smart Sys. (MASS), SENET Workshop, Seoul, South Korea, Sep. 2024, pp. 1–6.

[47] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018.

[48] J. Fu, "Scene segmentation with dual relation-aware attention network," *IEEE Trans. Neural Netw.*, vol. 32, no. 6, pp. 2547–2560, Jun. 2021.

[49] A. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[50] E. Bourtsoulatze, D. B. 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, Sep. 2019.

[51] Y. Xiao and M. Krunz, "Distributed optimization for energy-efficient fog computing in the tactile Internet," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 11, pp. 2390–2400, Nov. 2018.

[52] H. Zhang, H. Wang, Y. Li, K. Long, and A. Nallanathan, "DRL-driven dynamic resource allocation for task-oriented semantic communication," *IEEE Trans. Commun.*, vol. 71, no. 7, pp. 3992–4004, Jul. 2023.

[53] C. Bian, Y. Shao, and D. Gündüz, "DeepJSCC-1++: Robust and bandwidth-adaptive wireless image transmission," in *Proc. IEEE Global Commun. Conf.*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 3148–3154.

[54] Y. Xiao, Z. Sun, G. Shi, and D. Niyato, "Imitation learning-based implicit semantic-aware communication networks: Multi-layer representation and collaborative reasoning," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 3, pp. 639–658, Mar. 2023.

[55] S. R. Hasan, M. Z. Chowdhury, M. Sayem, and Y. M. Jang, "Quantum communication systems: Vision, protocols, applications, and challenges," *IEEE Access*, vol. 11, pp. 89 988–90 007, Jul. 2023.

[56] J. A. Ansere, E. Gyamfi, V. Sharma, H. Shin, O. A. Dobre, and T. Q. Duong, "Quantum deep reinforcement learning for dynamic resource allocation in mobile edge computing-based IoT systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 6, pp. 6221–6233, Jun. 2024.

[57] B. M. Mao, Y. C. Kawamoto, and N. Kato, "AI-based joint optimization of QoS and security for 6G energy harvesting Internet of things," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7032–7042, Aug. 2020.

[58] J. Dai, "Nonlinear transform source-channel coding for semantic communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2300–2316, Aug. 2022.

[59] G. Shi, H. Li, D. Gao, M. Yang, and Y. Dong, "An image adaptive rate mechanism in semantic communication for image endogenous semantics," *IEEE Trans. Veh. Technol.*, vol. 73, no. 9, pp. 13 425–13 439, Sep. 2024.

[60] V.-T. Hoang, V.-L. Nguyen, R.-G. Chang, P.-C. Lin, R.-H. Hwang, and T. Q. Duong, "Adversarial attacks against shared knowledge interpretation in semantic communications," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 2, pp. 1024–1040, Apr. 2025.

[61] A. Kirillov *et al.*, "Segment anything," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Paris, France, Oct. 2023, pp. 4015–4026.

[62] Y. Xiao and M. Krunz, "AdaptiveFog: A modelling and optimization framework for fog computing in intelligent transportation systems," *IEEE Trans. Mobile Comput.*, vol. 21, no. 12, pp. 4187–4200, Dec. 2021.

[63] ——, "Dynamic network slicing for scalable fog computing systems with energy harvesting," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 12, pp. 2640–2654, Sep. 2018.

[64] Y. Xiao, G. Shi, Y. Li, W. Saad, and H. V. Poor, "Toward self-learning edge intelligence in 6G," *IEEE Commun. Mag.*, vol. 58, no. 12, pp. 34–40, Dec. 2020.

[65] Y. Xiao, M. Hirzallah, and M. Krunz, "Distributed resource allocation for network slicing over licensed and unlicensed bands," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 10, pp. 2260–2274, Oct. 2018.

[66] Y. Xiao, Y. Li, G. Shi, and H. V. Poor, "Reasoning on the air: An implicit semantic communication architecture," in *Proc. IEEE ICC Workshop*, Seoul, South Korea, May 2022, pp. 1–7.

[67] Z. Wang, H. Xu, J. Liu, H. Huang, C. Qiao, and Y. Zhao, "Resource-efficient federated learning with hierarchical aggregation in edge computing," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Vancouver, BC, Canada, May 2021, pp. 1–10.

[68] B. Guler, A. Yener, and A. Swami, "The semantic communication game," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 4, pp. 787–802, Sep. 2018.

[69] W. Gong, H. Tong, S. Wang, Z. Yang, X. He, and C. Yin, "Adaptive bitrate video semantic communication over wireless networks," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Hangzhou, China, Nov. 2023, pp. 122–127.

[70] F. X. Tang, Y. C. Kawamoto, N. Kato, and J. J. Liu, "Future intelligent and secure vehicular network toward 6G: Machine-learning approaches," *Proc. IEEE*, vol. 108, no. 2, pp. 292–307, Feb. 2020.

[71] X. Shi *et al.*, "BSSNet: A real-time semantic segmentation network for road scenes inspired from autoencoder," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 5, pp. 3424–3438, May 2024.

[72] J. Ji, R. Shi, S. Li, P. Chen, and Q. Miao, "Encoder-decoder with cascaded CRFs for semantic segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 5, pp. 1926–1938, May 2021.

[73] Y. Zhou, T. Do, H. Zheng, N. Cheung, and L. Fang, "Computation and memory efficient image segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 1, pp. 46–61, Jan. 2018.

[74] B. Rekabdar, C. Mousas, and B. Gupta, "Generative adversarial network with policy gradient for text summarization," in *Proc. IEEE 13th Int. Conf. Semantic Comput. (ICSC)*, New York, NY, USA, Jan. 2019, pp. 204–207.

[75] H. Pan, Y. Hong, W. Sun, and Y. Jia, "Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 3, pp. 3448–3460, Mar.

2023.
[76] G. J. Brostow, J. Fauqueur, and R. Cipolla, "Semantic object classes in video: A high-definition ground truth database," *Pattern Recognit. Lett.*, vol. 30, no. 2, pp. 88–97, Jan. 2009.
[77] L. Xiao, N. Wang, and G. Yang, "A reading comprehension style question answering model based on attention mechanism," in *Proc. IEEE 29th Int. Conf. Appl.-Specific Syst., Archit. Processors (ASAP)*, Milan, Italy, Jul. 2018, pp. 1–4.
[78] K. Chen, T. Zhao, M. Yang, L. Liu, A. Tamura, R. Wang, M. Utiyama, and E. Sumita, "A neural approach to source dependence based context model for statistical machine translation," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 1, pp. 266–280, Feb. 2018.
[79] J. Nie, C. Zheng, C. Wang, Z. Zuo, X. Lv, S. Yu, and Z. Wei, "Scale–relation joint decoupling network for remote sensing image semantic segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–12, Nov. 2022.
[80] C. Zheng, Y. Jiang, X. Lv, J. Nie, X. Liang, and Z. Wei, "SSDT: Scale-separation semantic decoupled transformer for semantic segmentation of remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, no. 10, pp. 9037–9052, Jan. 2024.
[81] Y. Xiao, X. Zhang, Y. Li, G. Shi, and T. Basar, "Rate-distortion theory for strategic semantic communication," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Mumbai, India, Nov. 2022, pp. 279–284.
[82] G. Shi, Y. Xiao, Y. Li, and X. Xie, "From semantic communication to semantic-aware networking: Model, architecture, and open problems," *IEEE Commun. Mag.*, vol. 59, no. 8, pp. 44–50, Aug. 2021.
[83] F. Wang, X. Xu, Y. Chen, and X. Li, "Fuzzy semantics for arbitrary-shaped scene text detection," *IEEE Trans. Image Process.*, vol. 32, pp. 1–12, Aug. 2023.
[84] Y. Xu, Y. Wang, W. Zhou, Y. Wang, Z. Yang, and X. Bai, "TextField: Learning a deep direction field for irregular scene text detection," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5566–5579, Nov. 2019.
[85] C. Liang, Y. Sun, D. Niyato, and M. A. Imran, "Knowledge graph fusion based semantic communication framework," *IEEE Trans. Mobile Comput.*, vol. 24, no. 11, pp. 11 416–11 429, Nov. 2025.
[86] T. Han, Q. Yang, Z. Shi, S. He, and Z. Zhang, "Semantic-preserved communication system for highly efficient speech transmission," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 245–259, Jan. 2023.
[87] E. Gyamfi, J. A. Ansere, M. Kamal, S. Mir, and K. A. Bonsu, "Adaptive federated blockchain-powered security for MEC-assisted industrial IoT ecosystems," in *Proc. 19th Int. Conf. Emerging Technol. (ICET)*, Topi, Pakistan, Nov. 2024, pp. 1–6.
[88] C. Liang, H. Du, Y. Sun, D. Niyato, J. Kang, D. Zhao, and M. A. Imran, "Generative AI-driven semantic communication networks: Architecture, technologies, and applications," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 1, pp. 27–47, Feb. 2025.
[89] L. Xia, Y. Sun, D. Niyato, X. Li, and M. A. Imran, "Joint user association and bandwidth allocation in semantic communication networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 2, pp. 2699–2711, Feb. 2024.
[90] Q. Zhou, L. Wang, G. Gao, B. Kang, W. Ou, and H. Lu, "Boundary-guided lightweight semantic segmentation with multi-scale semantic context," *IEEE Trans. Multimedia*, vol. 26, pp. 7887–7900, Mar. 2024.
[91] A. Shumin, L. Qingmin, L. Zongqing, and J. Xue, "Efficient semantic segmentation via self-attention and self-distillation," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 15 256–15 266, Sep. 2022.
[92] L. Qingxuan, S. Xin, C. Changrui, D. Junyu, and H. Zhou, "Parallel complement network for real-time semantic segmentation of road scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 5, pp. 4432–4444, May 2022.
[93] J. A. Ansere, E. Gyamfi, Y. Li, H. Shin, O. A. Dobre, T. Hoang, and T. Q. Duong, "Optimal computation resource allocation in energy-efficient edge IoT systems with deep reinforcement learning," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 4, pp. 2130–2142, Dec. 2023.
[94] L. V. Nguyen, T. T. Nguyen, O. A. Dobre, and T. Q. Duong, "Digital semantic communication with neural image compression," in *Proc. IEEE INFOCOM 2025 Workshops*, London, UK, May 2025, pp. 1–2.
[95] A. Masaracchia, V.-L. Nguyen, D. B. da Costa, E. Ak, B. Canberk, V. Sharma, and T. Q. Duong, "Toward 6G-enabled URLLCs: Digital twin, open RAN, and semantic communications," *IEEE Commun. Standards Mag.*, vol. 9, no. 1, pp. 13–20, Mar. 2025.
[96] B. Chen, C. Gong, and J. Yang, "Importance-aware semantic segmentation for autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 137–148, Jan. 2019.
[97] L. Li, B. Qian, J. Lian, W. Zheng, and Y. Zhou, "Traffic scene segmentation based on RGB-D image and deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1664–1669, May 2018.
[98] Y. Kang, K. Yamaguchi, T. Naito, and Y. Ninomiya, "Multiband image segmentation and object recognition for understanding road scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1423–1433, Dec. 2011.
[99] V. Miclea and S. Nedevschi, "Real-time semantic segmentation-based stereo reconstruction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1514–1524, Apr. 2020.
[100] J. H. Anajemba, C. Iwendi, I. Razzak, J. A. Ansere, and I. M. Okpalaoguchi, "A counter-eavesdropping technique for optimized privacy of wireless industrial IoT communications," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 6445–6454, Sep. 2022.
[101] K. Yang, X. Hu, L. M. Bergasa, E. Romera, and K. Wang, "PASS: Panoramic annular semantic segmentation," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 10, pp. 4171–4185, Oct. 2020.
[102] E. Romera, J. M. Alvarez, L. M. Bergasa, and R. Arroyo, "ERFNet: Efficient residual factorized ConvNet for real-time semantic segmentation," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 1, pp. 263–272, Jan. 2018.
[103] T. Wu, S. Tang, R. Zhang, J. Cao, and Y. Zhang, "CGNet: A light-weight context guided network for semantic segmentation," *IEEE Trans. Image Process.*, vol. 30, pp. 1169–1179, Dec. 2021.

[104] X. Geng, L. Li, L. Jiao, X. Liu, F. Liu, and S. Yang, "Knowledge-aware geometric contourlet semantic learning for hyperspectral image classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 35, no. 1, pp. 698–712, Jan. 2025.

[105] L. Zhang and L. Zhang, "Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 270–294, Jun. 2022.

[106] S. K. Roy, S. Manna, T. Song, and L. Bruzzone, "Attention-based adaptive spectral–spatial kernel ResNet for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 9, pp. 7831–7843, Sep. 2021.

[107] Y. Su, L. Gao, M. Jiang, A. Plaza, X. Sun, and B. Zhang, "NSCKL: Normalized spectral clustering with kernel-based learning for semisupervised hyperspectral image classification," *IEEE Trans. Cybern.*, vol. 53, no. 10, pp. 6649–6662, Oct. 2023.

[108] H. Wu and S. Prasad, "Semi-supervised deep learning using pseudo labels for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1259–1270, Mar. 2018.

[109] Z. Zhong, J. Li, D. A. Clausi, and A. Wong, "Generative adversarial networks and conditional random fields for hyperspectral image classification," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3318–3329, Jul. 2020.

[110] S. Teng, C. Ning, W. Zhang, N. Wu, and Y. Zeng, "Fast asymmetric and discrete cross-modal hashing with semantic consistency," *IEEE Trans. Comput. Social Syst.*, vol. 10, no. 2, pp. 577–589, Apr. 2023.

[111] D. Zhang and X. Wu, "Scalable discrete matrix factorization and semantic autoencoder for cross-media retrieval," *IEEE Trans. Cybern.*, vol. 52, no. 7, pp. 5947–5960, Jul. 2022.

[112] J. A. Ansere, E. Gyamfi, M. Kamal, M. M. Khan, and K. A. Bonsu, "Quantum-inspired multi-agent computation offloading in edge intelligence-aided IoT networks," in *Proc. 19th Int. Conf. Emerging Technol. (ICET)*, Topi, Pakistan, Nov. 2024, pp. 1–5.

[113] S. Tariq, U. Khalid, B. E. Arfeto, T. Q. Duong, and H. Shin, "Integrating sustainable big AI: Quantum anonymous semantic broadcast," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 86–99, Jun. 2024.