A Survey on Semantic Communication Networks: Architecture, Security, and Privacy

Shaolong Guo, Yuntao Wang, *Member, IEEE*, Ning Zhang, *Senior Member, IEEE*, Zhou Su, *Senior Member, IEEE*, Tom H. Luan, *Senior Member, IEEE*, Zhiyi Tian, *Member, IEEE*, and Xuemin Shen, *Fellow, IEEE*

***Abstract*—With the rapid advancement and deployment of intelligent agents and artificial general intelligence (AGI), a fundamental challenge for future networks is enabling efficient communications among agents. Unlike traditional human-centric, data-driven communication networks, the primary goal of agent-based communication is to facilitate coordination among agents. Therefore, task comprehension and collaboration become the key objectives of communications, rather than data synchronization. Semantic communication (SemCom) aims to align information and knowledge among agents to expedite task comprehension. While significant research has been conducted on SemCom for two-agent systems, the development of semantic communication networks (SemComNet) for multi-agent systems remains largely unexplored. In this paper, we provide a comprehensive and up-to-date survey of SemComNet, focusing on their fundamentals, security, and privacy aspects. We introduce a novel three-layer architecture for multi-agent interaction, comprising the control layer, semantic transmission layer, and cognitive sensing layer. We explore working modes and enabling technologies, and present a taxonomy of security and privacy threats, along with state-of-the-art defense mechanisms. Finally, we outline future research directions, paving the way toward intelligent, robust, and energy-efficient SemComNet. This survey represents the first comprehensive analysis of SemComNet, offering detailed insights into the core principles as well as associated security and privacy challenges.**

***Index Terms*—Semantic communication, networks, artificial intelligence, security, privacy, and trust.**

I. Introduction

With the surge of advances in artificial intelligence (AI), our era is evolving from digitization (i.e., connected things) to intellectualization (i.e., connected intelligence). Under this evolution, communication devices are gradually empowered with human-like intelligence and reasoning capabilities [1], [2]. As predicted by Research&Market [1], by 2030, over 125 billion AI devices (known as agents[1]) will connect to the Internet, indicating their pivotal role in future communication applications.

Received 1 May 2024; revised 16 October 2024 and 28 November 2024; accepted 28 November 2024. Date of publication 13 December 2024; date of current version 20 October 2025. This work was supported in part by NSFC under Grant U22A2029, Grant 62302387, and Grant U23A20276; and in part by the Key Research and Development Program of Shaanxi under Grant 2022GXLH-01-25. *(Corresponding author: Zhou Su.)*

Shaolong Guo, Yuntao Wang, Zhou Su, and Tom H. Luan are with the School of Cyber Science and Engineering, Xi'an Jiaotong University, Xi'an 715500, China (e-mail: zhousu@ieee.org).

Ning Zhang is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada.

Zhiyi Tian is with the School of Computer Science, University of Technology Sydney, Sydney, NSW 2006, Australia.

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

Digital Object Identifier 10.1109/COMST.2024.3516819

A. Motivation of Semantic Communication

Unlike traditional Internet users requiring data communications, the goal of agent-based communications is to facilitate the task executions of agents. As a result, the agent-based communication presents three salient features:

- *Computing-oriented communications*, that the communications between agents should be sufficient to meet the computing requirements for task executions,
- *Persistent communications*, that communications among agents are typically persistent until all tasks are finished,
- *Memory-based communications*, that new information transmitted may depend on the previous information as the new task of agents may depend on previous ones.

Due to the features of agent-based communications, the traditional data communication framework may no longer fit, which motivates the new paradigm of semantic communication (SemCom) [3]. SemCom targets computing-oriented communications by first identifying the communication purposes (i.e., computing requirements) of agents; only the desired semantic meanings of data sufficient for computing are transmitted. To extract the semantic meaning of data, SemCom enables agents to exchange valuable and prior information (called the knowledge bases (KBs) as illustrated in Fig. 1), which is continuously updated to fit the persistent and memory-based features of agent-based communications. In this manner, SemCom can significantly reduce redundant data, facilitating collaborative agents to perform complex tasks.

Based on the topology, SemCom can be further divided into two categories:

Paired Semantic Communication: As illustrated in Fig. 1(a), the paired SemCom only considers the semantic-aware communication between two agent nodes. Compared to traditional communication, paired SemCom embraces various benefits, including ultra-high efficiency, enhanced reliability, and strong compatibility.

[1] Beyond the basic communication functions, these agents are equipped with advanced AI algorithms [2].

TABLE I
SUMMARY OF IMPORTANT ABBREVIATIONS IN ALPHABETICAL ORDER

Abbr.	Definition	Abbr.	Definition	Abbr.	Definition
AC	Access Control	AEs	Adversarial Examples	AI	Artificial Intelligence
AuthN	Authentication	Comm.	Communication	CSI	Channel State Information
DL	Deep Learning	DoS	Denial of Service	DP	Differential Privacy
GAI	Generative Artificial Intelligence	GANs	Generative Adversarial Networks	IoT	Internet of Things
IP	Intellectual Property	JSCC	Joint Source-Channel Coding	KBs	Knowledge Bases
KG	Knowledge Graph	ML	Machine Learning	PLA	Physical Layer Authentication
PLS	Physical Layer Security	QKD	Quantum Key Distribution	QoE	Quality-of-Experience
RIS	Reconfigurable Intelligent Surface	RL	Reinforcement Learning	SemCom	Semantic Communication
SemComNet	Semantic Communication Networks	SI	Semantic Information	SNR	Signal-to-Noise

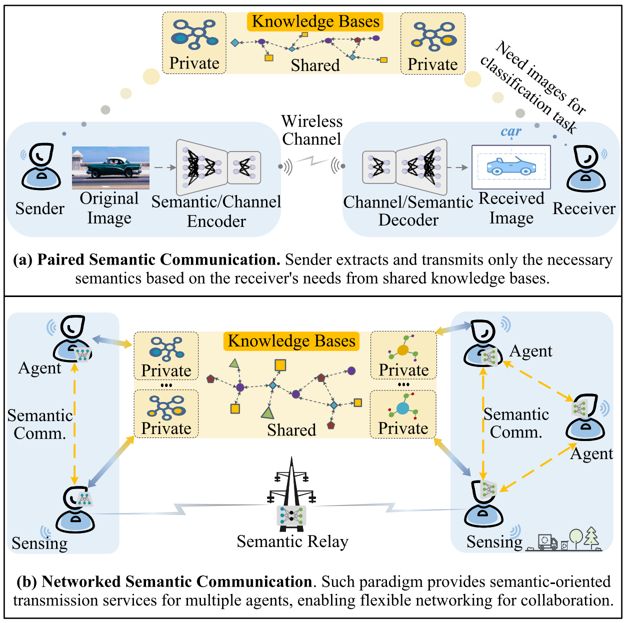


Fig. 1. Illustration of paired semantic communication and networked semantic communication.

First, SemCom exhibits ultra-high transmission efficiency [2], [3], in which pervasive AI and shared KBs allow agents to achieve *understand-before-transmit*. This capability enables SemCom to excel at extracting and transmitting compact semantic information (SI) while filtering out unnecessary data, thereby alleviating the transmission burden [3], [4].

Second, SemCom demonstrates enhanced transmission reliability, even under harsh communication channel conditions (e.g., high bit error rate) [4], [5], [6]. Impaired data from noisy wireless channels can be effectively corrected through semantic-level reasoning, guided by accumulated knowledge [7], [8]. Besides, the reconstruction of SI in SemCom relies on the matched semantic decoder, which enhances data security by making it more challenging for unauthorized entities to interpret or misuse the transmitted data.

Third, SemCom exhibits strong compatibility with existing communication infrastructures. Note that SemCom does not entirely replace traditional communication techniques. Instead, they can complement and reinforce each other. On the one hand, traditional communication technologies, such as reconfigurable intelligent surface (RIS) [9] and orthogonal frequency division multiplexing (OFDM) [10], can be integrated into SemCom to facilitate stable and reliable SI transmission. On the other hand, SemCom can be potentially integrated with existing communication infrastructures to enhance the performance of current systems. Researchers have established a field test network for 6G communication, demonstrating that SemCom could achieve 6G transmission capabilities on existing 4G links [11].

Networked Semantic Communication: The semantic communication networks (SemComNet) represent a network of agent nodes [12], [13] using SemCom. As illustrated in Fig. 1(b), SemComNet could provide semantic-oriented transmission services for multiple agents, enabling efficient and flexible networking [1]. As such, they allow seamless collaboration among agents with shared intents and objectives to execute complex tasks [1]. With SemComNet, a vast array of emerging intelligent applications are supported across key domains, including smart homes, autonomous driving, smart factories, digital twins, virtual reality (VR), Metaverse, and Tactile Internet.

However, extending paired SemCom to a networked environment serving multiple agents presents inherent challenges, including but not limited:

- *Training burden*. SemCom benefits from powerful AI models for semantic processing [14], but training these AI-driven semantic models requires high-performance computing and extensive labeled datasets. In addition, these models are typically task-specific and tailored to pairs of transceivers. As the number of agents in SemComNet increases, the resource demands for training and deploying semantic models grow exponentially, particularly when communicating with different agents and handling various tasks. This scaling challenge can quickly overwhelm the available resources of individual agents.
- *KBs consensus*. In paired SemCom, a critical prerequisite and challenge is establishing consensus on KBs between two agents [15]. This becomes more complex in SemComNet involving multiple agents, as SemComNet should ensure semantic compatibility and continuously maintain multiple KBs among agents. As new knowledge emerges or existing knowledge becomes outdated, the relevant KBs need to be upgraded. It is unrealistic for any single agent to update and coordinate these KBs independently, necessitating coordinated and collaborative efforts.

- *Topology dynamics*. In SemComNet, nodes can dynamically join and depart from the networks. As a result, both the training of semantic models for individuals and their KBs consensus may need to be continuously updated across the entire network to accommodate the network dynamics.

B. Motivation of Securing SemComNet

Due to the new structure of communications, the SemComNet present distinct challenges in security and privacy issues, which should be addressed foremost before its deployment [16], [17]. SemComNet face a variety of security risks and privacy breaches across its procedures:

- Hardware fragility: Compared to traditional networks, intelligent SemComNet heavily rely on computing resources (e.g., GPUs or other AI accelerators) for various AI-driven semantic codecs training. However, hackers may launch sponge examples attack [18] to vanish the acceleration hardware strategies. Such a threat may cause excessive energy consumption and severe performance degradation, and even result in denial of service (DoS) to other agents.
- Communication risk: SemComNet rely on conveying and interpreting concise SI for efficiency improvement. However, semantics ambiguity (i.e., semantic noise) [8], [19], [20] and the fragility nature of semantic models [19], [21] expose huge security risks. For instance, hackers could exploit these vulnerabilities via semantic adversarial attacks to mislead transceivers, as validated in [10], [22]. Specifically, by adding imperceptible noise to transmitted data and injecting modified data into semantic codecs at either transmitter or wireless channel sides [23], hackers could manipulate SI extraction or interpretation drastically. Besides, the subtle modification to transmitted data makes such attacks covert and hard-to-detect [16].
- KBs breach: Through proactive sensing and knowledge discovery, agents can establish and enrich their private KBs [24]. These private KBs typically contain sensitive information (e.g., individual preferences and behaviors), which opens new avenues for privacy-focused hackers [16]. For instance, malicious agents may attempt unauthorized access and retrieval of private information through brute-force attack [25], which poses a serious threat to confidentiality of SemComNet.

Apart from the discussed threats, existing security countermeasures may be ineffective and lack in adaptability due to intrinsic features of SemComNet, such as *heterogeneous components*, *autonomous intelligence*, and *large-scale structure*. Specifically, 1) the integration of heterogeneous components within SemComNet, including various communication modes, data modalities, and agents & KBs types, presents enormous interoperability difficulties [26]. 2) Since SemComNet comprise numerous intelligent agents with high-level autonomy, monitoring and managing their behaviors within the decentralized SemComNet become a challenge. Moreover, agents beyond effective control may emerge as potential security vulnerabilities, leading to exploratory attacks and data breaches. 3) The key prerequisite for reliable SemCom is the frequent querying of KBs to synchronize context and establish consensus between sender and receiver [15], which is challenging in the large-scale and time-varying SemComNet among diverse participants. Besides, attackers can exacerbate this challenge through attacks such as desynchronization of KBs and KBs poisoning attacks [16]. Consequently, it is imminent to develop a secure, trustworthy, and green SemComNet to overcome these challenges.

C. Related Works

The field of SemCom has garnered considerable research attention, giving rise to several surveys that explore its different aspects to date. For instance, Lan et al. [27] review the key components, enabling technologies, and design approaches of wireless SemCom. Shi et al. [13] introduce the classic SemCom and a semantic-aware network model including its architecture and open problems. Luo et al. [2] present recent advancements in DL-based end-to-end SemCom, covering various use cases and future trends. Qin et al. [28] comprehensively survey the principles and challenges of DL-driven SemCom systems for multi-modal data transmission. Uysal et al. [29] provide a view of semantic-aware networked architecture from the data importance aspect. Zhang et al. [12] investigate an AI-native SemCom-empowered network and discuss the prototyping, potential application scenarios, and key challenges. Gündüz et al. [30] provide a holistic review of semantic and task-oriented communication from an information-theoretic perspective. Yang et al. [31] systematically review the fundamentals of SemCom, potential applications, and open issues in 6G communication systems. Lu et al. [32] present a survey-cum-tutorial on AI-empowered SemCom technology from the ecosystem, frameworks, techniques, to application. Trevlakis et al. [33] provide a comprehensive analysis of SemCom in 6G networks which covers semantic knowledge, timeliness, and information theory aspects, along with proposing a networked SemCom architecture. However, as shown in Table II, existing surveys [2], [27], [28], [29], [30], [32] primarily focus on constructing SemCom between paired agents, neglecting the networked paradigm for collaborative agents. Although the works [12], [13], [31], [33] explore the networked SemCom, they fail to cover critical aspects such as the working modes comprehensively, use cases of SemComNet, and the security and privacy considerations. To bridge these research gaps, our paper provides a systematic review of the fundamentals of SemComNet, including an in-depth discussion of the three-tier architecture, working modes, enabling technologies, and practical use cases. Besides, with the growing significance of SemComNet, studying the security and privacy issues becomes essential for real-world deployment, which has been largely overlooked in previous surveys. Our survey offers a taxonomy of security and privacy threats across SemComNet's three layers and explores both existing and potential defense mechanisms. Table II provides a summary of our work's contributions about previous surveys in the field of SemCom.

TABLE II
A COMPARISON OF CONTRIBUTION BETWEEN OUR SURVEY AND RELEVANT SURVEYS

Year	**Refs.**	**Contribution**	**A**	**B**	**C**	**D**	**E**	**F**
2021	[27]	Survey on wireless SemCom including enabling technologies, components, and design approaches.	●	◐	●	○	○	○
2021	[13]	Discussions on the semantic-aware network framework including its architecture and open problems.	●	●	○	◐	○	○
2022	[2]	Discuss recent advancements in workflow, use cases, and key issues of end-to-end SemCom.	●	●	●	○	○	○
2022	[28]	Explore the principles and challenges of DL-driven SemCom for transmitting multi-modal data.	●	●	◐	○	○	○
2022	[29]	Review on semantic-aware communication from a data significance aspect.	●	○	●	○	○	○
2022	[12]	Discuss the key components and architecture of wisdom SemCom and review application scenarios and open issues.	●	◐	●	●	○	○
2023	[30]	A systematic taxonomy in semantic and task-oriented communication from an information-theoretic perspective.	●	●	○	○	○	○
2023	[31]	Comprehensive survey on fundamentals, potential applications, and open issues of SemCom-driven 6G systems.	●	●	●	◐	○	○
2024	[32]	A tutorial-cum-tutorial on AI-driven SemCom from the ecosystem, frameworks, techniques, to application.	●	●	●	○	○	○
2024	[33]	Discuss paired SemCom design, theoretical framework, as well as SemCom networking architectures.	●	●	○	●	○	○
Now	**Ours**	Comprehensive survey of the fundamentals, security, and privacy of SemComNet, discussions on the general architecture, working modes, application scenarios, and security/privacy threats of the SemComNet, review on critical challenges, potentially advanced solutions, and research directions in building future SemComNet.	●	●	●	●	●	●

A: Paired SemCom Design ***B***: Multi-modal Data ***C***: Use Cases ***D***: Networked SemCom Architecture ***E***: Security & Privacy Threats ***F***: Security & Privacy Solutions
● : fully included ◐ : partially included ○ : not included

In this paper, we provide a systematic review of the general architecture, as well as the potential security/privacy threats, state-of-the-art countermeasures, and future trends of SemComNet. To the best of our knowledge, this is the first survey paper in the literature presenting the fundamentals of SemComNet for multi-agent interaction, as well as an in-depth analysis of the security and privacy aspects. The main target of our work is 1) to provide readers with a general understanding of SemComNet, as well as the scope and implications of security threats and challenges in SemComNet, and 2) to highlight effective/potential paths and methods to prevent these threats, thereby fostering the secure SemComNet for various intelligent applications. The contributions of this work are four-fold.

- *Overviews of SemComNet (Section II):* A three-tier SemComNet architecture including the *control layer*, the *semantic transmission layer*, and the *cognitive sensing layer* is first introduced. Then we elaborate on working modes (i.e., *paired*, *clustered*, and *networked*), along with the use cases, enabling technologies, and evaluation metrics.
- *Security Threats and Challenges (Section III):* We present a taxonomy of security and privacy threats in the SemComNet across three layers and investigate the critical challenges to address them.
- *State-of-the-Art Countermeasures (Section IV):* We review the existing defense solutions from both academia and industry, as well as assess their potential for establishing the secure, trustworthy, reliable, and privacy-preserving SemComNet.
- *Future Research Directions (Section V):* We discuss open research opportunities and outline future directions of SemComNet aiming to facilitate new starters conducting research in this field.

Finally, we draw conclusions in Section VI. The organizational structure of this work is depicted in Fig. 2, and Table I summarizes a list of key acronyms.

II. AN OVERVIEW OF SEMANTIC COMMUNICATION NETWORKS

In this section, we first clarify the roadmap of semantic communication (SemCom) and the motivation of the proposed semantic communication networks (SemComNet). Then, we introduce an envisaged architecture for SemComNet, in which working modes, enabling technologies, use cases, and semantic metrics are elaborated.

1) Roadmap of SemCom: Research in SemCom follows three major trends: the shift from single-modal to multi-modal data transmission, the evolution from single-task to multi-task execution, and the progression from paired SemCom to networked SemCom.

- *From Single-modal to Multi-modal Data Transmission.* SemCom research initially concentrated on single-modal data, such as text, images, audio, or video, as shown in Table III. While efficient, single-modal systems have limitations in fully capturing the diversity of communication needs. The shift to multi-modal transmission integrates data from different sources (e.g., combining text with images or audio), which enhances the semantic richness. However, this comes with increased system complexity, particularly in SI extraction, data alignment, and fusion across modalities.
- *From Single-task to Multi-tasks Execution.* SemCom systems have demonstrated exceptional performance in single-task scenarios such as image classification [3] and speech recognition [4] in both efficiency and accuracy. However, as communication scenarios become increasingly complex, the focus of SemCom is shifting towards the simultaneous handling of multiple tasks [15], [34]. For instance, in SemCom-empowered autonomous driving, multiple tasks such as environmental perception and object detection need to happen concurrently with minimal latency. Nevertheless, transitioning to multi-task SemCom introduces several challenges. For instance, task interference may arise when different tasks require varying features, potentially leading to performance degradation [34].
- *From Paired SemCom to Networked SemCom.* In the future, SemCom is evolving from a paired paradigm [2], [3], [34] to a networked paradigm [13]. This transition to SemComNet introduces inherent challenges (e.g., training burden, KBs consensus, and topology dynamics, as discussed in the Introduction), which significantly increase system complexity and

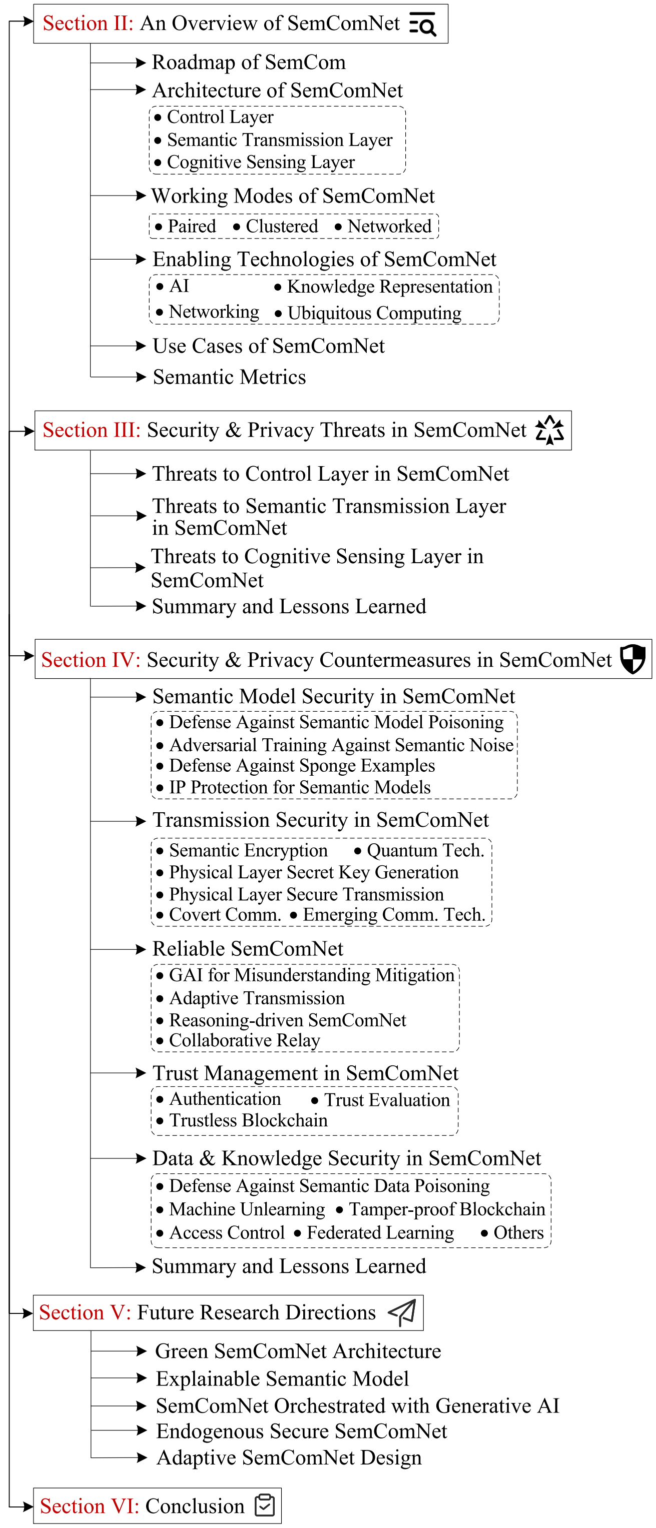


Fig. 2. Organization structure of this paper.

require effective interactions and high-level coordination among multiple agents. For instance, a critical prerequisite for the success of SemComNet is the establishment and continuous updating of shared KBs among these agents [13], which demands collaborative efforts to gather environmental semantics [35]. Despite the complex and long-term challenges of achieving networked SemCom, the potential benefits for multi-agent communication scenarios are substantial.

TABLE III
COMMON AND TYPICAL DATASETS USED BY SEMCOM DIVIDED BY MODALITY TYPE

Data Modality	Typical Datasets
Text	European Parliament [3]; WMT 2018 News [36]
Image	MNIST, CIFAR-10, SVHN, and USPS [37]; Stanford Online Products, CUB-200-2011, Cars1, and In-Shop Clothes [36]
Audio	Librispeech [4]; LJSpeech [38]
Video	CamVid [39]; Vimeo90K, HEVC test dataset, and UVG [40]

In this paper, we extend the concept of paired SemCom to networked SemCom [13], [29] and propose the state-of-the-art SemComNet. As illustrated in Fig. 1, it represents a semantic-oriented networking paradigm that could better support connected multi-agent intelligent interaction [32] via SemCom. In the proposed architecture (refer to Section II-A), by leveraging KBs maintenance, task scheduling, resource allocation, SI delivery, as well as environmental perception and cognitive reasoning, SemComNet effectively enhance the collaborative semantic transmission abilities of agents [32]. As such, SemComNet enable multiple agents to efficiently realize semantic interactions and collectively accomplish complex tasks for 6G and beyond applications [12].

According to the executed tasks, we further classify the SemComNet into *clustered SemCom* and *networked SemCom*. The former allows multiple agents collaboratively to accomplish a common task, while the latter is expected to handle more complex application scenarios that comprise various transmission tasks [34] and to support coordination across these tasks (refer to Section II-B for more details).

A. Architecture of SemComNet

In this paper, we propose a general SemComNet architecture for information sharing comprising multiple agents. Based on previous studies [13], [26], [41], we have identified three layers of functionality, which are shown in Fig. 3 and described as follows:

- *Control layer*: It serves as the management and orchestration component, responsible for managing shared KBs, scheduling tasks, and dynamically allocating resources across the cloud-edge-end architecture within SemComNet.
- *Semantic transmission layer*: This layer ensures efficient interaction among collaborative agents by providing semantic-oriented information delivery services.
- *Cognitive sensing layer*: It upgrades the sensing layer by integrating cognitive ability, which is tasked with environment sensing, agents' intent inference, and structuring accumulated data into contextual knowledge to enrich private KBs.

1) Control Layer: As the top-level management and orchestration component, this layer is mainly responsible for three functions, i.e., *KBs maintenance*, *task scheduling*, and

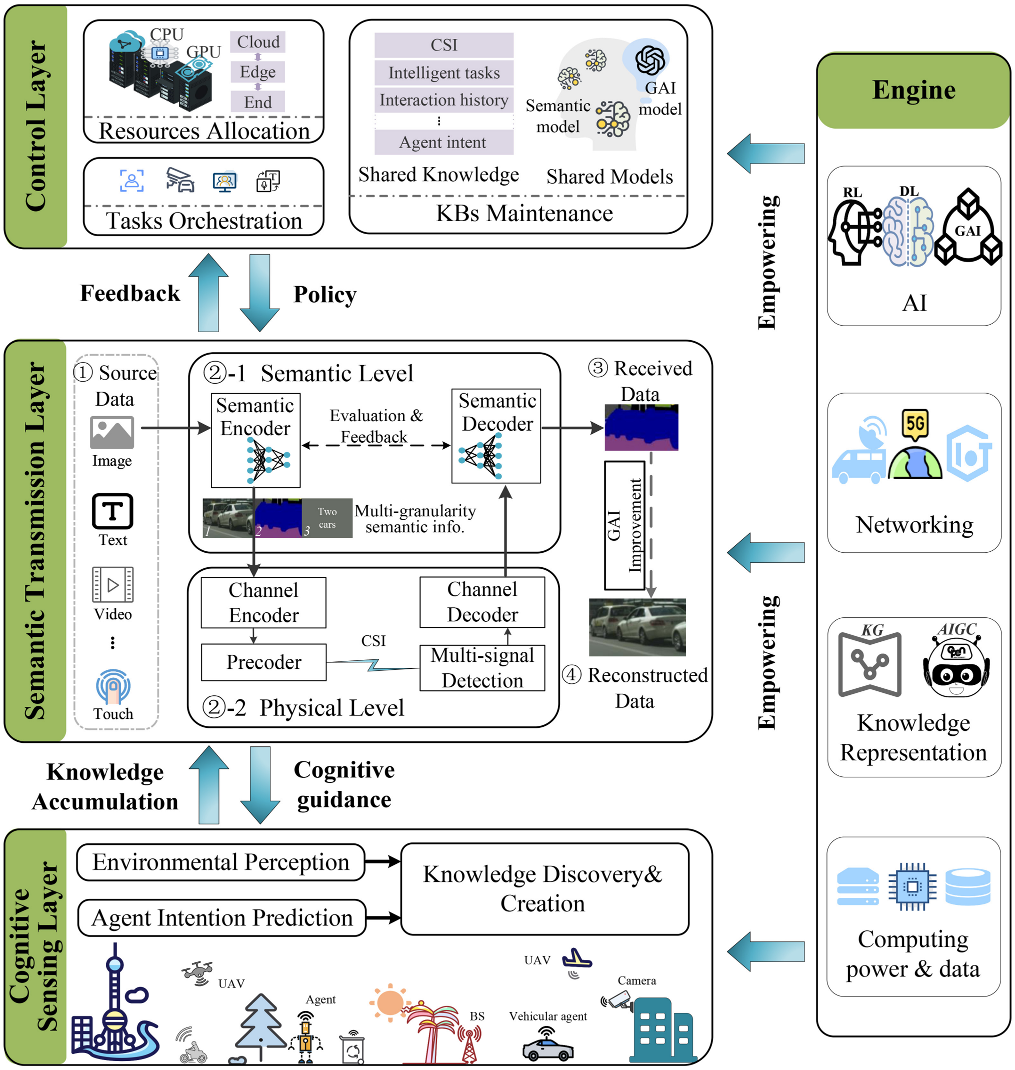


Fig. 3. Architecture of the SemComNet.

overall resource coordination, covering key aspects of network management including data & knowledge, operation, and resources.

a) Maintenance of knowledge bases: This layer manages multiple shared KBs to enhance the semantic understanding and reasoning abilities of agents in SemComNet. To support diverse SemComNet applications, these KBs contain diverse shared semantic models (i.e., semantic codecs) and comprehensive contextual knowledge (e.g., source data, channel state information (CSI), and communication task requirements). On the one hand, agents need to deploy diverse semantic models for efficient SI extraction and reconstruction in various SemComNet applications. However, training these models is time-consuming and costly for agents. In response, this layer provides various shared semantic models, enabling agents to transition quickly from traditional to semantic-oriented service provisioning. Consequently, the shared models alleviate the necessity to train desired models from scratch. On the other hand, this layer offers extensive prior knowledge to assist agents during semantic reasoning and decoding phases, thereby improving the efficiency and reliability (e.g., eliminating semantic ambiguity) of interaction. The knowledge can be represented in various forms including knowledge graph (KG) [20], structured databases, and parameters in GAI models [42], [43], as detailed in Section II-C3.

b) Intelligent communication tasks scheduling: Task scheduling within the control layer is essential for orchestrating network operations, especially in a SemComNet with heterogeneous agents and ever-increasing tasks (e.g., 3D video conferencing and VR transmission). It ensures tasks are executed at optimal times, in the correct sequence, and by the appropriate agents to maximize network performance [13]. On the one hand, it involves aligning task requirements with the capabilities of different agents, optimizing resource utilization, and ensuring timely task assignments. On the other hand, it evaluates and monitors system performance metrics such as response times, quality-of-experience (QoE), and agent feedback, with a focus on ensuring the accuracy, relevance, and contextual coherence of transmitted SI.

c) Adaptive overall resource allocation: To enhance the performance of SemComNet, the resources including communication, and computational power (including storage) need to be carefully allocated to facilitate efficient interaction within large-scale SemComNet. Given that SemComNet require significant computing power to train various semantic codecs and update KBs frequently, the cloud-edge-end hierarchical framework is utilized for on-demand resource utilization [44]. Specifically, the cloud offers massive-scale high-performance computing resources (e.g., powerful CPUs, GPUs, and memory) that are suitable for global KBs maintenance and shared semantic model training. By leveraging edge resources located at the network edge (e.g., base stations and access points), the data transfer latency and privacy leakage can be reduced [45], thereby facilitating real-time processing and access of shared KBs for agents. Pervasive end-agents are equipped with limited computing capabilities, allowing them to handle simpler tasks such as environment sensing & processing, updating privacy-sensitive KBs, and conducting semantic-oriented information delivery.

2) Semantic Transmission Layer: This layer ensures efficient and reliable delivery of SI through two key stages: preparation and semantic-aware transmission. The preparation stage lays the foundation for effective semantic transmission by focusing on both knowledge synchronization and model training. In terms of knowledge, agents align with shared KBs to establish a common understanding. Regarding model training, the objective is to equip agents with semantic models to handle diverse tasks. Current methods for training semantic codecs can be categorized into three types: private learning, federated learning (FL), and centralized learning [46].

During the semantic-aware transmission stage, the process operates through four phases: a) multi-source data fusion, b) SI extraction, c) physical-layer reliable transmission, and d) semantic recovery and enhancement.

a) Multi-source data fusion: This initial phase involves the collection and fusion of multi-modal data sources (e.g., image, point cloud, and video) [20], [36], [47].

b) SI extraction: This phase is a critical operation within this layer which employs a semantic encoder to extract meaningful and desired information from informative multi-modal source data [20], [48] while filtering out redundant and known knowledge for the receiver. Guided by prior knowledge from shared KBs, the semantic encoder can adaptively extract multi-level SI on demand. For instance, for image transmission with various task requirements, high-level SI containing a general summary of an image (e.g., two cars) suffices for classification tasks, especially for resource-limited transmitters. Meanwhile, middle-level SI provides a more specific understanding of objects, regions, or specific structures, making it suitable for scene comprehension tasks. Furthermore, low-level SI involves pixel-level details, significantly enhancing precise scene reconstruction tasks, even with significant knowledge gaps among transceivers. Additionally, the transmitter dynamically adjusts its strategy according to the recipient's feedback and channel conditions, enhancing adaptability to transceiver needs and environmental variations.

c) Physical-layer reliable transmission: To maximize SI transmission efficiency and mitigate interference among multiple agents, this layer employs precoding techniques (e.g., beamforming and spatial multiplexing) to optimize transmitted signals at the source. The precoder leverages spatial diversity to enhance signal quality and enable simultaneous data transmission. Then, by leveraging channel coding, the extracted semantic data streams are reliably delivered through physical channels. Besides, to offer better performance gains, the DL-based joint source-channel coding (JSCC) paradigm [49], [50] can map the source data to channel symbols for enhanced efficiency and flexibility while well-settle cliff-effect by training in an end-to-end manner.

d) Semantic recovery and enhancement: After receiving the transmitted SI, the receiver agent first performs multi-user signal detection to identify and separate signals related to itself. This process involves employing intelligent algorithms such as intelligent radio [2], to estimate the CSI, and to detect and differentiate between the various signals. Then, guided by shared KBs matched by the control layer, the receiver performs semantic decoding and provides feedback to the transmitter. Besides, to conduct semantic reasoning and parsing with auxiliary knowledge and context, the receiver can rectify transmission errors, such as content blurring or partially missing. For instance, injecting blurred SI into GAI models to accomplish content correction and enhancement [51], thereby improving the accuracy and reliability of SemComNet.

3) Cognitive Sensing Layer: This layer represents an upgraded version of the traditional sensing layer [26] by incorporating human-like cognitive processing capabilities into system design [52]. Specifically, unlike the traditional sensing layer, this layer not only has direct interfaces with the physical environment to actively perceive the environment through sensors but also integrates the cognitive abilities of agents. For human beings, cognition involves acquiring understanding and knowledge through thought, experience, and senses [53]. Similarly, for intelligent agents, the cognitive abilities include understanding sensor data, reasoning about agents' intents, and discovering knowledge to enrich their private KBs [52], [53]. The cognitive sensing layer involves three features: *a) perceiving the surrounding environment*, *b) inferring the intents of agents*, and then *c) organizing the accumulated information into knowledge* for the SemComNet [35]. A detailed discussion of these features follows:

a) What do the agents have? The pervasive agents equipped with on-body sensors can actively sense and collect environmental semantics such as light, sound, and CSI from the surroundings [35]. For instance, vehicular agents leverage various sensors (e.g., GPS and cameras) to collect data about other vehicles and traffic conditions. Besides, these intelligent agents within SemComNet possess the capability of semantic understanding [1] to interpret environmental data including identifying specific events or entities. For instance, they could recognize specific patterns in sound or objects in images, thereby understanding events occurring in the environment. Subsequently, this layer processes the raw data into meaningful descriptions regarding environmental and contextual aspects, such as "vehicle traffic is growing".

b) What do the agents want? Identifying the agents' intents is crucial for meeting their different transmission

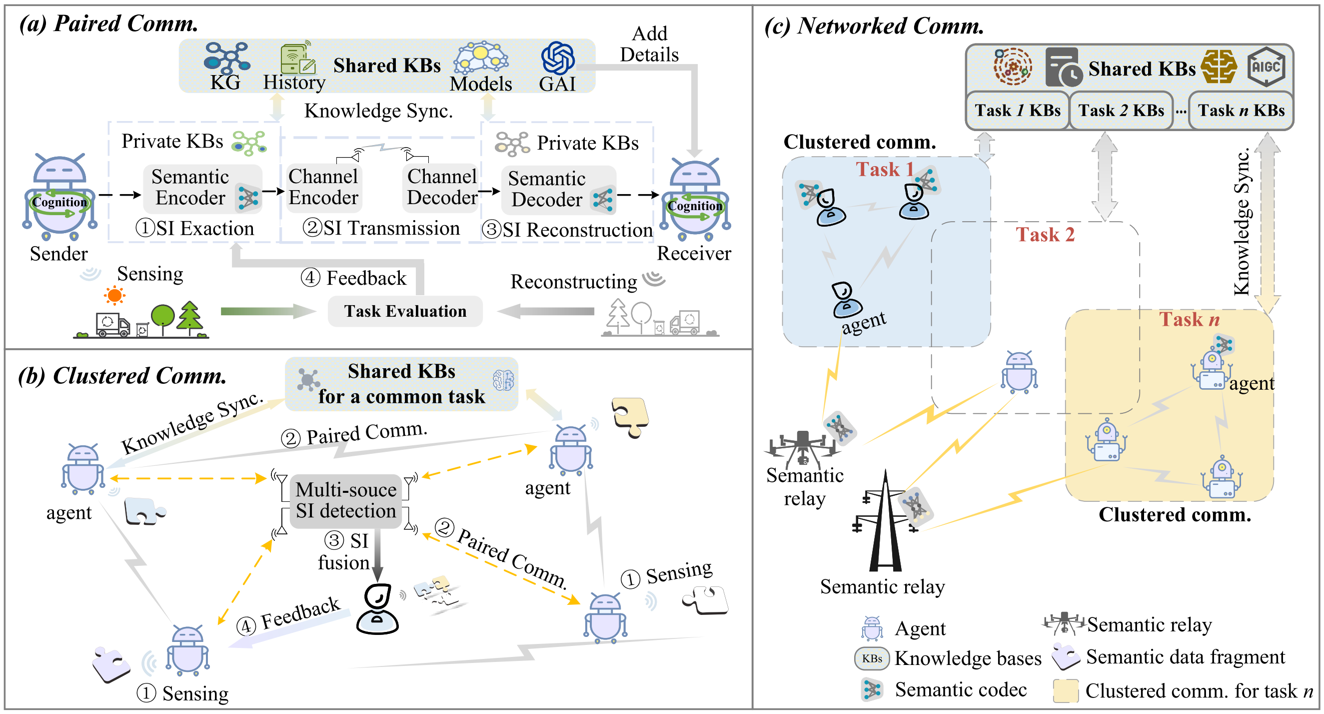


Fig. 4. SemComNet working modes. (a) Paired SemCom: interaction between two agents via SemCom. (b) Clustered SemCom: each agent collaborates with others via paired SemCom within a cluster for a common task. (c) Networked SemCom: multiple clustered agents interact for different tasks, where intra-cluster agents interact via clustered SemCom and inter-cluster communication may assisted by semantic relay.

requirements, thereby improving agents' QoE and unleashing the potential of SemCom [12], [53]. For instance, in the SemComNet-empowered smart transportation scenarios, the intent of vehicular agents for communication may be either "warn" of imminent hazards (e.g., sudden stops ahead), "coordinate" actions among vehicles (for safe intersection crossing), or "optimize" traffic flow through intelligent traffic signal management. Understanding these intents is crucial for effective environment perception and semantic noise reduction [1]. To accurately infer agents' intent, this layer analyzes these agents' behaviors (e.g., habit preferences, bodily movements, emotional states) and interaction histories, employing statistical tools and AI technologies for predictive insights. The AI-based implementation approaches can be categorized into four types [54]: (1) traditional machine learning (ML)-based methods; (2) DL-based methods; (3) RL-based methods; and (4) cognitive model-based methods. Specifically, traditional ML algorithms such as Bayesian networks and hidden Markov models excel in performing pattern recognition from collected data in an explainable manner. When dealing with extensive historical data, DL approaches such as recurrent neural networks and long short-term memory (LSTM) models excel in analyzing action sequence dependencies and inferring underlying intents from extensive data. Furthermore, to adapt to unseen environments, RL techniques empower agents to learn optimal inference strategies through trial and error. Moreover, in scenarios with scarce training data, the cognitive model [53], [55], [56] leverages the theories from cognitive science (e.g., theory of mind [57]), which could predict the agents' decision-making processes and intents. For instance, inspired by human cognitive processes, Rabinowitz et al. [57] propose the machine theory of mind network, which can learn other agents' behaviors to model their mental states including desires, beliefs, and intents from limited data.

c) How do the agents acquire knowledge? After environmental perception and intent inference processes, this layer could effectively transform the derived sensing data and agents' intents into structured forms of knowledge. For instance, KGs and large GAI model-based KBs store knowledge in the form of graph structures and model parameters [42], [51], respectively. In [20], Li et al. illustrate how to construct a cross-modal knowledge graph (CKG) as the KBs, as depicted in Fig. 5. The transformation of raw multi-modal data into structured knowledge includes four phases: 1) data collection and preprocessing to standardize input, 2) multi-modal knowledge extraction to derive entities, relations, and attributes, 3) cross-modal knowledge fusion to integrate data and expand the KBs, and 4) information storage and retrieval using graph databases for efficient knowledge access. As such, it enriches private KBs, enabling agents to enhance their understanding and contextual reasoning within SemComNet. Furthermore, an essential interplay occurs between private KBs of agents and shared KBs of the control layer for continuous knowledge updating and alignment [44]. This interaction ensures the integration of individual knowledge with globally shared knowledge, such as aggregating new environmental semantics [35] to shared KBs and removing obsolete knowledge. Consequently, this process ensures the consistency of multiple KBs in SemComNet and boosts the accuracy of SemCom [13]. However, agents may be reluctant

TABLE IV
A SUMMARY OF THREE WORKING MODES IN THE SEMCOMNET

	Paired Semantic Comm.	Clustered Semantic Comm.	Networked Semantic Comm.
Number of Tasks	1	1	$\geq$ 2 tasks across clusters
Number of Participants	2 agents	> 2 agents form a cluster	Multiple clusters
Communication Modes	End-to-end semantic comm.	Paired semantic comm. within a cluster	Clustered semantic comm. and inter-cluster via semantic translators
Application Scenarios	Collaborative task execution between two agents	Collaborative task execution within a cluster of agents	Scenarios requiring coordination across tasks
Types of KBs	Private and shared KBs between two agents	Private and task-special KBs among clustered agents	Multi-layered comprising public, task-specific, and private KBs

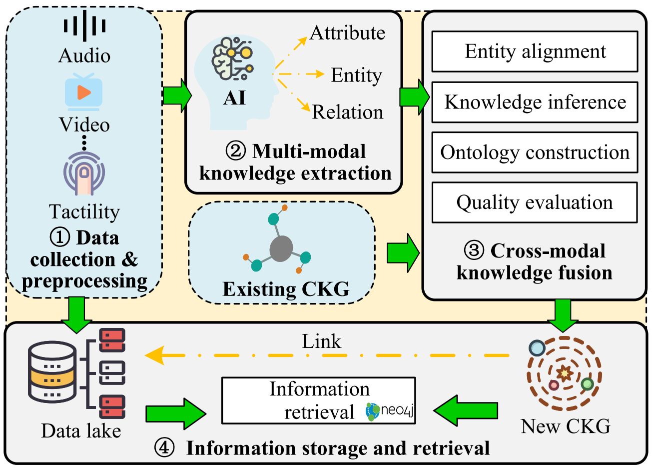


Fig. 5. Illustration of the construction of cross-modal knowledge graph (CKG) in [20], which involves four key steps: 1) data collection and preprocessing, 2) multi-modal knowledge extraction, 3) cross-modal knowledge fusion that combines extracted knowledge and existing CKG, and 4) information storage and retrieval, supported by graph databases (e.g., neo4j).

to share their private KBs due to privacy concerns [13]. FL provides a promising solution [58], allowing agents to collaboratively construct shared KBs without exposing sensitive information, as detailed in Section IV-E5.

B. Working Modes of SemComNet

As shown in Fig. 4, the SemComNet have three working modes according to the communication modes: (1) *paired SemCom* where SI is directly transmitted between two agents rather than raw data streams, (2) *clustered SemCom* enabling collaborative interactions and the exchange of SI among agents within clusters for the same task, and (3) *networked SemCom* facilitating connections for different clustered agents engaged in diverse tasks, either through direct intra-cluster connections or indirectly assisted by *semantic translators*.[2] In Table IV, we summarize the main differences between the above three working modes in SemComNet.

1) Paired Semantic Communication: As illustrated in Fig. 4 (a), this mode allows two agents to interact and aims to deliver the key meaning behind the source data, instead of transmitting the raw bit streams. Before the interaction, both communicating entities undergo the preparation stage. Here, agents engage in synchronizing background knowledge and aligning their shared KBs to establish a common ground and context for communication. To alleviate the computation burden of agents, lightweight semantic codecs can transfer from large, powerful models within shared KBs via knowledge distillation.[3]

The transmission involves three principal phases, i.e., SI exaction, SI transmission, and SI reconstruction. *i) SI exaction phase.* In this phase, the transmitter uses the semantic encoder to extract compact SI from original multi-modal data streams, while simultaneously filtering out irrelevant information. The work [31] categorizes semantic encoder designs into four main approaches: DL-based, RL-based, KB-assisted, and semantic-native approaches. Among these, various DL-based solutions [2], [3], [34] are identified as the mainstream direction currently. *ii) SI transmission phase.* After extraction, the transmitting agent transmit processed SI to the receiver. The channel codecs are responsible for ensuring error-free transmission of SI, capable of combating noise, interference, and other challenges related to the physical layer. During channel coding, mechanisms such as error correction, compression, and encryption can be employed to ensure the integrity, efficiency, and security of the SI. *iii) SI reconstruction phase.* Upon receiving the transmitted SI, the receiver utilizes a semantic decoder to interpret and reconstruct the SI into a format relevant to its tasks. The semantic decoder design should be tailored to specific downstream tasks, which can be classified as pragmatic task execution (e.g., image classification) and observable information reconstruction (e.g., video transmission) [37]. During reconstruction, the receiver

[2]Semantic translators (a.k.a semantic relays [59], [60], [61], [62]) act as intelligent relays that facilitate the semantic meaning forwarding between the transceivers. Their role is similar to that of human translators who bridge interactions between individuals from different cultural backgrounds. For instance, when the source and destination have mismatched KBs or incompatible semantic codecs, the semantic translator decodes the SI using its shared knowledge with the source and re-encodes it based on its shared knowledge with the destination.

[3]Knowledge distillation (KD) is an ML model compression technique that transfers knowledge from a large, complex teacher model to a smaller student model. In SemComNet, KD allows agents to create lightweight yet effective semantic codecs from larger models (e.g., pre-trained semantic models), benefiting resource-constrained agents with limited computational power.

interacts with KBs to ensure the SI aligns with the contextual background.

Finally, a quality assessment is conducted to verify the reconstructed information maintains reliability and suitability at the semantic level. Based on the assessment and CSI, the transmission strategy is adaptively adjusted to mitigate semantic noise and channel-related disturbances.

2) Clustered Semantic Communication: This mode involves multiple agents to collaboratively accomplish a common task, as depicted in Fig. 4 (b). Agents are organized into clusters based on common interests in tasks such as cooperative navigation and real-time object detection [63]. Within each cluster, sharing domain-specific knowledge and task-specific semantic models enables efficient collaboration and SI exchange. Achieving clustered SemCom requires the following four steps. *i) Perception and information collection.* At first, agents utilize their sensors to actively collect data from the surrounding environment [35]. Then, each agent independently cognitive sensing processes this data locally and extracts relevant semantic fragments[4] that contribute to the shared task, such as collaborative object perception. *ii) Semantic fragment sharing via paired SemCom.* Within the cluster, agents share the locally extracted semantic fragments with other agents via paired SemCom. *iii) SI fusion.* Upon receiving SI from other agents, the recipient applies multi-signal detection and multi-modal data fusion [48]. These processes aim to reconstruct a more comprehensive and holistic semantic representation. *iv) Re-transmission for missing information.* During the semantic reconstruction phase, the receiver identifies missing or incomplete information via verification. Subsequently, feedback or requests are sent to the respective agents, prompting them to re-transmit the necessary information fragments [44]. This process aims to achieve a complete semantic understanding and fulfill the task requirements.

3) Networked Semantic Communication: This paradigm is expected to accommodate complex application scenarios comprising various tasks [36] and to support coordination across tasks [13]. In this paradigm, connected agents are organized into different clusters based on their involved tasks. Within each task, agents efficiently exchange SI via clustered SemCom. Meanwhile, for inter-cluster communication, agents belonging to different clusters can be assisted or coordinated by *semantic translators*, a.k.a. semantic relay nodes [59], [60], [61], [62]. These translators, as depicted in Fig. 4 (c), could be acted by powerful unmanned aerial vehicles and base stations, which possess cross-domain knowledge and diverse semantic models. Their role extends beyond basic SI forwarding, emphasizing semantic translation to guarantee accurate interpretation and exchange of SI at the semantic level [59]. Consequently, semantic translators not only mitigate knowledge background disparities among transceivers across clusters, but also alleviate agents from the necessity of training multiple semantic models across tasks. As such, it boosts connectivity and efficiency for networked agents.

The networked SemCom mode relies on multi-layered KBs [44], including public, task-specific, and private KBs. Specifically, public KBs provide common-sense knowledge and general models accessible to networked agents. Within each cluster, agents share task-specific KBs, which concentrate on domain expertise and task-related knowledge (e.g., diseases, symptoms, and treatments in medical diagnosis tasks). Additionally, private KBs owned by individual agents mainly contain personalized and sensitive knowledge accumulated by each agent (e.g., preferences, interests, and communication context). These multi-layered KBs provide essential context and knowledge for SI extraction, delivery, and reconstruction for agents.

C. Enabling Technologies of SemComNet

1) AI: Serving as the foundation of SemComNet, AI techniques including DL, RL, and GAI significantly enhance the ability of SI extraction and overall efficiency in dynamic communication conditions and diverse task demands. Specifically, DL models such as LSTM and Transformer [2], [3] allow semantic models to robustly abstract key features and capture long-range dependencies within data. Given the frequent need for model updates in SemComNet, which can lead to service interruptions and significant time consumption, transfer learning [3] facilitates the reuse of existing models and knowledge for new tasks, accelerating training and improving data efficiency. Moreover, to mitigate "catastrophic forgetting" [64] in transfer learning, continual learning becomes essential in SemComNet which helps semantic models adapt to new tasks without forgetting learned knowledge. To address the issue of scarce training samples in the SemComNet environment, GAI models such as generative adversarial networks (GANs) and diffusion models prove beneficial [45], [51]. These models could create extensive, high-quality, and personalized data samples resembling real-world scenarios for training diverse semantic models. Besides, RL empowers agents with autonomous implicit semantic reasoning [44] and real-time decision-making in SemComNet tasks such as autonomous driving and drone navigation.

2) Networking Technology: In SemComNet, networking technology such as IoT, Bluetooth Low Energy (BLE), Long Range (LoRa), as well as 5G and beyond (B5G) plays a crucial role in enhancing the performance and efficiency, achieving reliable network connections, and real-time SI transmission between agents. The increasing affordability and advancing intelligence of IoT devices [65] have made it feasible to deploy numerous intelligent agents. Simultaneously, the abundant deployment of sensors in the IoT provides SemComNet with rich perceptual data, facilitating knowledge accumulation and SI comprehension. Moreover, BLE offers short-range and low-latency wireless networking, ideal for intra-cluster communication, enabling seamless data exchange with minimal energy consumption [66]. LoRa enhances SemComNet

[4]Semantic fragment (a.k.a. semantic seb [12]) refers to the smallest unit of meaningful information (i.e., SI) in a SemCom system. Generally, a semantic seb consists of a feature or set of features extracted from the source data, which vary depending on the method used to extract them and the specific communication needs or tasks. Moreover, the same content's semantic sebs may be interpreted inconsistently across agents with different backgrounds. Note that, the definition and standardization of semantic sebs remain evolving, necessitating further research.

TABLE V
ENABLING TECHNOLOGIES AND THEIR ROLES IN SEMCOMNET

Enabling Tech.	Role	Covered Tech.
AI	Semantic-aware data processing and decision making	ML, DL, RL, and GAI
Networking Technology	Ensure efficient SI transmission and reliable connectivity for all agents	5G and beyond, LoRa, BLE, and IoT
Knowledge Representation	Manage and represent knowledge for efficient retrieval and KBs construction	KGs, ontologies, and scene graphs
Ubiquitous Computing	Provide on-demand and sufficient computing resources for SemComNet	Mobile, edge, and cloud computing

by enabling long-range, reliable communication for inter-cluster agents spread across wide areas, even in challenging environments. Besides, B5G [12] support high-speed data transmission, increased network capacity, and global communication coverage, including remote areas and oceans, fulfilling the real-time response and efficient data transfer requirements of SemComNet.

3) Knowledge Representation Technology: Acting as "memory" of SemComNet, the knowledge representation such as knowledge graph (KG) [20], [31], [67], ontologies [33], [68], and scene graphs [69] can transform contextual information and experience into a comprehensible format for agents. They also empower SemComNet to manage, search, and reason over vast amounts of knowledge. For instance, KGs can be integrated into SemComNet to organize and aggregate vast unstructured data into a structured knowledge format from diverse domains [67], presented in graphical form. By formally defining the concepts and their relationships, ontologies [33] offer a structured representation that aids in identifying entities (e.g., people, places, and events) within data. Besides, Zhang et al. [69] use scene graphs to capture and store knowledge about the objects and their relationships in the original image.

4) Ubiquitous Computing: SemComNet require significant computing power to sustain AI model training and storage, as well as multiple KBs management and synchronization. In response, ubiquitous computing [23] establishes an environment for SemComNet where computing resources are seamlessly and invisibly embedded in various agents (e.g., wearable devices and sensors), ensuring widespread availability. To enhance agents' QoE, the cloud-edge-end network architecture offers on-demand access to computational and storage resources. The cloud tier offers powerful computing resources for training shared semantic models, while the edge tier facilitates rapid computation processing for nearby agents. Additionally, through collaborative maintenance of multiple pre-cached KBs across the cloud-edge-end infrastructure, agents can efficiently access required knowledge, improving semantic delivery performance.

5) Summary: As shown in Table V, AI, networking, knowledge representation, and ubiquitous computing technologies form the fundamental pillars of SemComNet. AI enhances semantic understanding and intelligent optimization, networking technology provides reliable and scalable transmission, knowledge representation structures and manages knowledge, and ubiquitous computing supplies the necessary computational resources. These four technologies collectively enable intelligent and semantic-aware interaction among agents within SemComNet.

D. Use Cases of SemComNet

With the assistance of SemComNet, a vast array of emerging intelligent applications across numerous key domains are enabled [31], [70], such as autonomous driving, holographic-type communication, and the Metaverse.

1) Autonomous Driving: Autonomous driving depends on real-time data exchange among vehicles, roadside infrastructure, and cloud servers to make safe driving decisions. Equipped with sensors such as cameras and LiDAR, autonomous vehicles continuously generate vast amounts of data to monitor their surroundings and provide critical inputs for decision-making. However, conventional communication systems often struggle to meet the ultra-reliable, low-latency, and high-throughput transmission demands required for fast decision-making and stable coordination. SemComNet represent the novel paradigm for intelligent collaboration in autonomous driving by integrating user intent and semantics into the communication process. This enhances data exchange and decision-making capabilities among vehicles, roadside infrastructure, and other entities, enabling efficient vehicle-road collaboration while meeting the strict demands of perception accuracy and latency in smart transportation systems.

In a SemComNet-enabled autonomous driving scenario, each vehicle maintains private KBs with sensitive data (e.g., historical routes and location coordinates) while sharing semantic KBs with an edge server. These shared KBs include background knowledge (e.g., city maps) and task-specific semantic models. For tasks requiring rapid inference, vehicles reconstruct SI from multi-modal data and shared KBs in real-time to make local decisions (e.g., acceleration, braking). By accessing the edge server's knowledge, vehicles expand their perception range and make accurate decisions with minimal computation. In more complex scenarios, vehicular agents transmit SI to the edge server and other nearby vehicles, which collaborate to make final decisions. This collaborative process enhances task performance by leveraging the complementary nature of SI and task correlations. For instance, Feng et al. [71] propose a multi-user SemCom system for edge intelligence-enabled autonomous driving. They employ the JSCC method and fusion module for various users and data modalities to reduce data redundancy and optimize task coordination, thus improving transmission efficiency.

2) Holographic-Type Communication: Holographic-type communication [72] leverages holographic display systems to capture individuals and their surroundings remotely and transmit them over networks. At the receiver's end, laser projections create real-time 3D holograms, enabling interaction with virtual projections. This technology provides an immersive, multi-modal experience via naked-eye holography. However, current holography requires substantial data volumes for synthesizing images and 3D data. For example, multi-angle holographic videos demand terabit-per-second (Tbps) transmission rates and sub-1 ms latency [72] to achieve full immersion, which remains challenging.

Instead of transmitting all raw data, SemComNet focus on delivering only critical SI, reducing the data volume for transmission. For instance, in holographic video conferencing, this might involve transmitting key features such as facial expressions instead of full volumetric data. In [73], Cheng et al. propose a semantic-driven framework, named SemHolo, for enhancing holographic telepresence in 6G networks. By prioritizing task-relevant aspects (e.g., key gestures and facial expressions) over full volumetric data, SemHolo enables the efficient delivery of 3D content. A proof-of-concept implementation demonstrates the feasibility of real-time interactive telepresence while ensuring visual quality and reconstruction fidelity. Besides, the multi-modal data transmission capabilities of SemComNet could link different modalities through a common semantic framework. This allows for expressing the same semantic meaning across various modalities, enabling simultaneous recovery of multi-modal information from a single semantic context. By focusing on semantic representations, SemComNet address the bandwidth challenges of holographic communication while improving efficiency.

3) Metaverse: The Metaverse, envisioned as a next-generation Internet, offers a fully immersive and high-fidelity virtual universe where users interact in computer-generated environments [74]. For instance, Metaverse concerts provide large online audiences with a 360-degree panoramic view and an interactive experience, allowing digital avatars to engage with both the venue and other participants. However, large-scale multi-user activities in the Metaverse demand significant resources from both backend servers and user devices (e.g., VR headsets). Specifically, for the Metaverse servers, rendering detailed virtual environments with numerous users requires immense computing and transmission resources. As for Metaverse users, they face the challenge of constantly receiving and rendering potentially thousands of other participants' positions and actions.

In practice, not all content needs to be transmitted or rendered to maintain a seamless and immersive experience [75]. SemComNet address this by transmitting only semantically meaningful data, reducing storage and transmission loads. For instance, rather than transmitting full 3D models of a user's gestures, SemComNet transmits only essential information, such as key movements that trigger interactions. This reduces data load, accelerates processing, and improves system performance [31]. To ensure trustworthy SemCom in the Metaverse, Chen et al. [76] propose a secure multi-user SemCom system based on intelligent radio and GANs, while FL balances privacy and communication accuracy.

E. Semantic Metrics

Unlike traditional communication systems that use metrics such as *symbol-error rate* (SER) or *bit-error rate* (BER), semantic metrics in SemComNet focus on the significance and meaning of information [33]. These metrics can be classified into three categories [70], i.e., *significance-oriented*, *meaning-oriented*, and *combined* metrics. Specifically, significance-oriented metrics evaluate the value and relevance of information based on its timing, timeliness, and impact on subsequent tasks [31], [33]. These metrics help prioritize information transmission in resource-constrained scenarios. For example, the *age of information* (AoI) [77] and *value of information* (VoI) [29] are two representative metrics used for significance evaluation in SemComNet [33]. AoI quantifies the timeliness of information at the destination by measuring the time elapsed since the latest data was generated at the source, while VoI considers both the timeliness and quality of the transmitted information [33]. Besides, the meaning-oriented metrics assess the similarity between the transmitted and received SI, prioritizing semantic fidelity over pixel or bit-level accuracy. Lastly, combined metrics (e.g., semantic QoE [78], [79]) integrate aspects of semantic fidelity, communication costs, and user experience, providing a more holistic evaluation.

Besides, the above semantic metrics can significantly influence the security and privacy aspects of SemComNet. Take significance-oriented metrics as an example, they can be used to prioritize time-sensitive information and detect timing anomalies that may indicate malicious activities. For instance, an attacker may exploit vulnerabilities (e.g., delaying or replaying outdated or irrelevant data) to confuse or mislead the receiving agent. By monitoring metrics such as AoI or VoI, SemComNet can identify and discard data that is excessively old or delayed beyond acceptable thresholds, thereby mitigating the risk of timing-based attacks.

III. Security and Privacy Threats in SemComNet

Despite the promising prospects of SemComNet, it faces significant security and privacy issues, which have not been widely discussed. As critical systems, SemComNet's security and privacy concerns must encompass the confidentiality, integrity, and availability (CIA) of resources (including data, shared knowledge, and semantic models). Specifically, confidentiality issues primarily involve unauthorized access to sensitive resources, as well as potential privacy and intellectual property (IP) infringement. Integrity is compromised by attacks that alter or corrupt data/knowledge (e.g., launch false data injection, poisoning, and adversarial attacks) which compromise its accuracy and semantic meaning. Availability focuses on threats disrupting services, such as DoS and malware attacks, affecting SemComNet's operational functionality for legitimate users.

Besides, to effectively and comprehensively identify these threats, we conduct a layer-by-layer analysis. As shown in Fig. 6, we categorize a broad scope of security/privacy threats in SemComNet (from Sections III-A–III-C) from its three layers of functionality, i.e., control, semantic transmission, and cognitive sensing layers. Fig. 7 provides three illustrative examples of attacks within these layers.

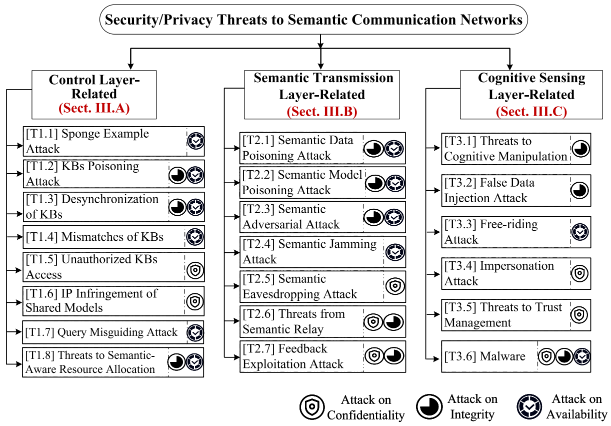


Fig. 6. The taxonomy of security/privacy threats to SemComNet from the three functional layers (i.e., control layer, semantic transmission layer, and cognitive sensing layer).

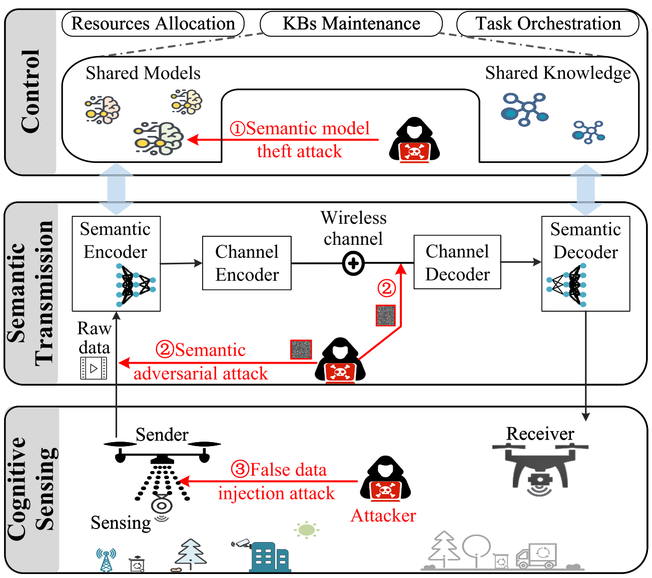


Fig. 7. An illustrative example of semantic model theft, semantic adversarial, false data injection attacks in the SemComNet.

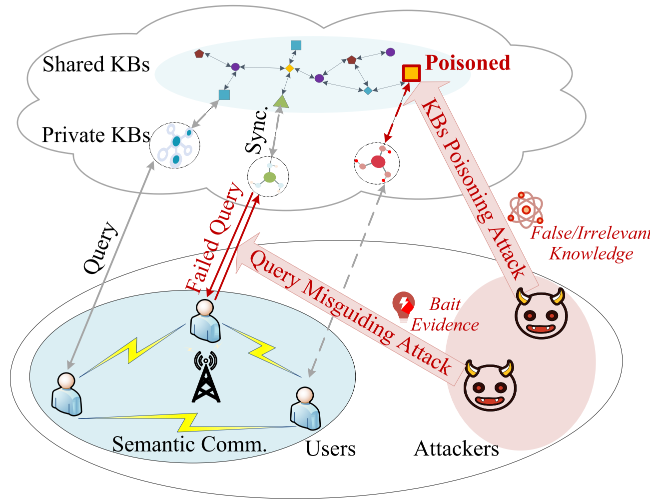


Fig. 8. An illustrative example of KBs poisoning attack and query misguiding attack in the SemComNet.

A. Threats to Control Layer in SemComNet

As outlined in Section II-A1, the control layer in SemComNet is tasked with top-level management and orchestration for tasks, resources, and shared KBs. However, adversaries could launch specialized attacks such as DoS (e.g., by injecting sponge examples) which aims at depleting resources and rendering SemComNet unavailable. Besides, compared with traditional communication networks, a distinctive feature of SemComNet is its maintenance of multiple KBs to provide plenty of shared knowledge and semantic models for collaborative agents. However, this feature incurs vulnerabilities such as KBs poisoning, unauthorized KBs access, and theft of shared semantic models. Below, we enumerate the typical threats related to the control layer in SemComNet.

- *Sponge Examples Attack [T1.1].* In traditional networks, DoS attack [18], [80] disrupt normal operations by overwhelming systems with excessive traffic, preventing legitimate users from accessing network services. In SemComNet, a variant known as sponge examples attack emerges [18], [81], which targets energy consumption and latency in DL components. The threat is severe due to SemComNet's reliance on computational resources (e.g., powerful CPUs and GPUs) for DL-driven semantic models training. Hackers can inject crafted sponge examples into semantic codecs, nullifying hardware acceleration and causing excessive energy consumption and response delays. In [18], Shumailov et al. introduce two methods for generating sponge examples: gradient-based and genetic-based. The former is a glass-box approach requiring access to model parameters, while the latter is a closed-box technique that optimizes samples based on energy or latency metrics by simply querying the model. To mitigate such threats, one solution is worst-case performance analysis [18], which establishes processing time and energy thresholds using natural examples. Inputs exceeding these thresholds are rejected, ensuring system robustness. A more comprehensive solution is discussed in Section IV-A3.
- *Knowledge Bases Poisoning Attack [T1.2].* Unlike conventional communication networks, SemComNet relies heavily on shared KBs to aid semantic understanding and reasoning [32]. However, this reliance on KBs introduces new attacks that are relatively easy to exploit but challenging to detect. Specifically, malicious entities may manipulate the storage nodes of KBs (e.g., cloud and edge servers) via unauthorized access [23] to influence their cached knowledge [16]. As shown in Fig. 8, one specific attack involves KBs poisoning [23], where attackers inject false, harmful, or misleading knowledge into KBs, thereby deceiving receivers and deteriorating the overall performance of SemComNet. To mitigate this, access control (AC) and blockchain technologies, as discussed in Section IV-E3, offer secure KBs management, ensuring tamper-proof records and preventing unauthorized modifications.
- *Desynchronization of Knowledge Bases [T1.3].* The primary target of this attack is to undermine both the

integrity and availability of KBs. Attackers may induce inconsistency or desynchronization of KBs by disrupting network connections, manipulating update frequencies, or tampering with updated versions. This desynchronization within SemComNet allows adversaries to secretly delay, modify, or even destroy KBs without detection [23]. As a result, inconsistent semantic understandings arise among agents, compromising the overall robustness and effectiveness of SemComNet. For instance, malicious entities might intentionally introduce conflicting or outdated knowledge, thereby desynchronizing multiple KBs and interfering with the accurate extraction of SI. Currently, this type of attack and its corresponding defenses remain unexplored in the field of the SemCom domain, warranting further research.

- *Mismatches of Knowledge Bases [T1.4].* This threat may affect the availability of KBs matching mechanism, leading to semantic noise between transceivers [2]. Before interaction, participating agents in SemComNet should synchronize and align their prior knowledge [3], [15] to maintain consistent and up-to-date contextual understanding. However, in realistic scenarios, agents may be reluctant to share their sensitive knowledge with others due to privacy concerns and substantial communication burdens, resulting in KBs mismatches between transmitter and receiver [24]. These mismatches may incur semantic-level misunderstandings (e.g., semantic noise) at the receiver [2]. Besides, the dynamic nature of the environment necessitates continual updates in KBs, further exacerbating the mismatches or disparities between them. To effectively mitigate this risk, techniques such as semantic relay (detailed in Section IV-C4) and resilient semantic understanding schemes (discussed in Sections IV-C1 and IV-C3) may be beneficial.
- *Unauthorized Knowledge Bases Access [T1.5].* Unauthorized access [80] in communication systems occurs when entities gain access to systems or data without proper authorization, potentially leading to threats such as data theft, system manipulation, or service disruption. In SemComNet, such a threat expands the attack surface due to the existence of multiple privacy-sensitive KBs [24]. For instance, regionally shared KBs contain relatively personal and sensitive knowledge that is only accessible to agents within the coverage of edge servers [44]. Curious agents may attempt unauthorized access or retrieval of data from KBs via various approaches such as brute-force attacks [25] and impersonation attacks [80], which pose a severe risk of privacy breaches. Moreover, attackers can tamper with the KBs with poisoned data or backdoors [82] after gaining unauthorized access, resulting in deteriorated communication performance and loss of confidentiality. More defense details can refer to Sections IV-E3 and IV-E2.
- *IP Infringement of Shared Semantic Models [T1.6].* SemComNet accelerate the shift of transmission services from traditional to semantic-oriented paradigms by providing diverse shared semantic models for authorized agents. As such, the computational burden of these agents may be significantly reduced, mitigating the need to train semantic codecs from scratch. However, significant IP threats may arise in SemComNet due to the replicability characteristic of these shared semantic models, which allow authorized entities to resell these valuable models for illegal profit without being detected [83]. Furthermore, during the model duplication and distribution phases, these semantic models are susceptible to risks such as theft (as depicted in the upper part of Fig. 7), counterfeiting, and unauthorized imitation [84]. Considering the expensive and time-consuming nature of their training process, the above risks may pose significant economic and confidentiality losses for SemComNet. To effectively mitigate IP infringement, techniques such as watermarking and blockchain can be beneficial, as discussed in Section IV-A4.
- *Query Misguiding Attack [T1.7].* In SemComNet, agents frequently query the KBs [25] to synchronize relevant knowledge. However, adversaries may manipulate KBs and alter the query behavior [25] through various methods such as injecting malicious SQL code, installing malware, or phishing. For instance, adversaries can impede agents from generating informative queries by introducing misleading evidence, referred to as *bait evidence* [85], illustrated as Fig. 8. For instance, when an agent queries KBs regarding the first president of the USA, an expected response should provide information about George Washington. Yet, adversaries might craft misleading queries, such as "Was George Washington the first president of the United States?" to mislead the KBs and obstruct the retrieval of accurate knowledge. Currently, research on defending against such threats in SemComNet is still lacking and requires further study.
- *Threats to Semantic-Aware Resource Allocation [T1.8].* SemComNet offer a semantic-aware and spectrum-efficient paradigm that enables users to transmit semantic-rich information with minimum spectrum resources [69], [79]. Especially in scenarios with limited resources (e.g., spectrum and power), allocation strategy is decided based on semantic importance. It means prioritizing the transmission of semantic-rich information using limited spectrum resources [86]. However, hackers may execute the feature importance-aware attack [87] by injecting transferable adversarial examples (AEs) into semantic codecs. These examples not only divert the attention of codecs to areas with poor semantic content but also lead to the failure of scarce spectrum allocation (e.g., resources are occupied by semantic-poor information). As a result, the integrity of semantic codecs and the availability of services may be compromised. At present, defense strategies against such threats in SemComNet remain unexplored and require further investigation.

B. Threats to Semantic Transmission Layer in SemComNet

This layer provides semantic transmission services tailored to the needs of agents and the semantics of sent data, enabling efficient and intelligent communication akin to human interactions [44]. However, SemCom, often referred to as "AI +

Communication" [2], [28], [32], inherit dual vulnerabilities. On the one hand, sensitive SI and computing tasks are exposed to broadcast channels, offering attackers the chance to launch attacks such as semantic jamming and semantic eavesdropping. On the other hand, advanced AI-empowered semantic codecs are susceptible to manipulation through semantic adversarial or poisoning attacks. Such threats can alter and forge transmitted meanings, introducing semantic noise [2] and disrupting communication services in SemComNet. Below, we identify the typical attacks related to the semantic transmission process.

- *Semantic Data Poisoning Attack [T2.1].* In SemComNet, adversaries can launch semantic poisoning attacks to undermine the integrity and availability of semantic models, reducing transmission accuracy [19]. These attacks fall into two categories: data poisoning and model poisoning, which depend on whether the attacker contaminates the transceivers' training datasets or directly alters semantic models [25], [88]. In data poisoning attacks, attackers inject or modify a subset of the training data to degrade the performance of semantic codecs or induce unexpected behaviors. Such attacks are further divided into dirty-label poisoning and clean-label poisoning. Dirty-label poisoning alters both the data and corresponding labels, making it easier to detect due to the inconsistency between the data and labels. In contrast, clean-label poisoning is subtler as only the data is manipulated. Common defenses against data poisoning attacks include data preprocessing [89], robust training [88], and poison filtering mechanisms [90] (as detailed in Section IV-E1).
- *Semantic Model Poisoning Attack [T2.2].* Different from data poisoning attacks that require access to training data, semantic model poisoning attacks directly target the model's parameters or training process. These attacks typically occur when attackers gain access to semantic codecs' learning and update procedures, which are out of the transceivers' control. Such attacks are particularly severe in collaborative or distributed training settings, where multiple agents jointly train semantic models using local data and necessitate frequent model exchanges. In such scenarios, attackers may control some clients and manipulate their updates during training, or implate backdoors that trigger specific behaviors under certain conditions [82]. Typical defensive strategies against model poisoning include model sanitization in centralized learning settings [91], as well as robust aggregation and anomalous clients detection [92], [93] in collaborative training settings (as detailed in Section IV-A1).
- *Semantic Adversarial Attack [T2.3].* Existing works [94] show that deep neural networks (DNNs) are susceptible to AEs, i.e., subtle perturbations added to the inputs of a DL model, causing it to make incorrect decisions. Since SemComNet relies on advanced DL models, these vulnerabilities can be inherited, and the over-the-air transmission of SI further exacerbates the risk. Specifically, as shown in the middle part of Fig. 7, the semantic adversarial attack can target both the sender and receiver sides. At the sender, attackers introduce subtle perturbations to distort semantic meaning and impede accurate SI extraction. For instance, Hu et al. [95] leverage iterative fast gradient sign method (FGSM) [94] to generate sample-dependent semantic noise that misleads the transmitter, which may dramatically alter the semantics of data. At the receiver (potentially via the wireless channel), subtle perturbations to semantic decoder inputs may fail semantic decoding and misunderstanding between transceivers. To mitigate such attack, solutions within ML domain provide valuable insights, such as adversarial training [8], [10], [95], defensive distillation [94], detection of AEs [94], and weight perturbation [95] (as detailed in Section IV-A2).
- *Semantic Jamming Attack [T2.4].* The conventional jamming attack [96] disrupts communication by emitting interference or high-power signals [97], causing bit-level interruptions or degraded communication quality. In contrast, semantic jamming attack targets the integrity of the transmitted data's semantics [98], [99], degrading the consistency and quality of semantic content and preventing accurate interpretation at the receiver [19]. For instance, Tang et al. [99] propose a semantic jammer that generates jamming data streams by randomly sampling from a Gaussian distribution. However, their scheme [99] assumes the jammer shares the same network structure as the sender, which may be impractical for real-world scenarios. Besides, to counter such attack, Tang et al. [99] introduce a GAN-inspired framework, where a semantic jammer (generator) and a robust receiver (discriminator) optimize strategies through adversarial gaming, thereby improving the receiver's resilience to semantic jamming. Apart from directly countering jamming signals, techniques such as spread spectrum (detailed in Section IV-B6) [100] and covert communication (detailed in Section IV-B4) [19], [101], [102] can conceal the existence of SI transmission, making it difficult for jammers to detect and target SI.
- *Semantic Eavesdropping Attack [T2.5].* Eavesdropping, a common threat in traditional communication, involves intercepting transmitted data to expose sensitive information [96]. In SemComNet, a semantic eavesdropping attack involves intercepting transmitted signals to infer their semantic meanings [21], [103]. To evaluate the threat, Du et al. [19] present the semantic secrecy outage probability (SSOP), which quantifies the success rate of intercepting and decoding SI. Defending against semantic eavesdropping is easier than traditional eavesdropping, as accurately decoding intercepted SI is highly challenging without access to paired semantic codecs and shared knowledge. However, two significant threats remain. Firstly, the rapid advancement of GAI models enables attackers to use these models as general semantic decoders [104], capable of semantic interpretation and intent understanding [42]. Secondly, during semantic codecs training, gradient sharing of models can expose sensitive data. As demonstrated in [105], observing gradients allows attackers to reconstruct

fine-grained training data. To mitigate these risks, techniques such as semantic data encryption (detailed in Section IV-B1), covert communication (discussed in Section IV-B4) [19], [101], [102], physical layer security technologies (discussed in Sections IV-B2, IV-B3, and IV-B6), and quantum technologies (detailed in Section IV-B5) could be employed for secure semantic transmission.

- *Threats from Semantic Relay [T2.6].* In wireless communication systems, relays extend communication range and improve data transmission reliability through cooperative strategies [96] such as amplify and forward (AF). However, untrusted relays introduce security risks, including eavesdropping, data tampering, or selectively dropping transmissions. In SemComNet, these threats persist and may lead to more severe consequences. Semantic relay nodes not only forward data but also translate semantic signals at the semantic level [60], [61], especially in scenarios where transceivers lack matched knowledge [59], [62]. Typically, these relay nodes store rich background knowledge [59], including privacy-sensitive information about transceivers. On the one hand, curious relays may steal this private information, resulting in more severe privacy breaches. On the other hand, malicious relays may exploit their advantageous position to inject viruses [106] into SemComNet or manipulate the relayed SI via adding semantic noise [22], potentially causing erroneous interpretations of semantics. However, current defense strategies against such threats in SemComNet remain unexplored and warrant further investigation.
- *Feedback Exploitation Attack [T2.7].* In SemComNet, the transmitter adjusts its transmission strategy according to feedback from the receiver and physical channel (e.g., CSI). If the receiver struggles to interpret the content, the sender transmits more detailed, low-level SI to aid understanding [107]. However, two security issues arise during this feedback process, i.e., feedback leakage and feedback tampering. The former occurs when malicious entities analyze feedback to infer sensitive information about agents, such as their preferences and behavior patterns, potentially constructing detailed agents' profiles and threatening confidentiality [16]. Meanwhile, the latter refers to malicious feedback modification, leading to inappropriate adjustments in transmission rates. That potentially causes communication failures and affects the availability of SemComNet. Currently, the defense against such attack in the SemComNet domain remains an open issue, warranting further research.

C. Threats to Cognitive Sensing Layer in SemComNet

As described in Section II-A3, the cognitive sensing layer in SemComNet is distinguished for its advanced cognitive capabilities, including environment perception, agent intention inference, knowledge discovery, and private KBs establishment. However, these capabilities also introduce unique vulnerabilities during the data & knowledge perception, processing, and sharing phases. Various threats exist within the cognitive sensing layer, including but not limited to false perception data injection, free-riding, and cognitive manipulation via "information bombs". Additionally, risks such as unauthorized private KBs access and private KBs poisoning attacks, mirror the vulnerabilities found in the control layer. We list typical threats to the cognitive sensing layer in SemComNet as below.

- *Threats to Cognitive Manipulation [T3.1].* These threats in SemComNet involve deliberate interference with agents' cognitive processes to influence their judgment and reasoning. Cognition manipulation can stem from adversarial information dissemination, such as spreading rumors, or deceptive content (e.g., Deepfake images generated by GANs [150]). Such tactics compromise decision-making and the integrity of SemComNet. Additionally, the easy-to-use GAI models enable malicious agents to create biased "information bombs" [81] via malicious prompts, disrupting public cognition and shaping virtual opinion leaders. Defenses against these threats include adversarial information detection techniques, categorized as spatial-based, frequency-based, and data-driven methods [150]. Spatial-based techniques enhance local forgery detection by focusing on specific spatial regions, frequency-based methods analyze differences in the frequency domain, and data-driven approaches improve detection generalization by training with diverse fake data [81].
- *False Data Injection Attack [T3.2].* To enrich private KBs and improve environmental understanding, agents in SemComNet autonomously collect environmental semantics via on-body sensors [35]. However, as illustrated in the bottom part of Fig. 7, attackers might attempt to mislead SemComNet by injecting false data. For instance, GPS spoofing attack [153] can override legitimate signals with counterfeit ones, causing GPS receivers to misinterpret geographical information. This deception can introduce flawed environmental semantics into the KBs, leading to incorrect decisions and undermining SemComNet's integrity. To effectively mitigate such threat, techniques such as robust authentication and provenance tracking may be beneficial, as discussed in Sects. IV-D1 and IV-D2.
- *Free-riding Attack [T3.3].* In multi-agent SemComNet, collaboration is essential for task completion (e.g., collaborative perception) and ensuring smooth system operations. However, selfish agents may avoid contributions while reaping benefits or accessing its resources [23]. For instance, selfish agents may deliberately provide low-quality [132] or semantic-poor sensing data to save their energy and resources. If widespread, such free-riding behavior can deplete system resources, reduce service availability, and degrade the QoE of legitimate agents. Mitigation strategies include trust evaluation (detailed in Section IV-D3) and effective AC (discussed in Section IV-E3).

- *Impersonation Attack [T3.4].* Adversaries can exploit authentication vulnerabilities in SemComNet to impersonate legitimate agents, compromising system confidentiality [80]. This could be achieved by stealing the credentials through password cracking or phishing [151]. Once successful, attackers can misuse compromised identities for malicious activities, such as submitting false sensory data or violating privacy. The authentication mechanism (discussed in Section IV-D1) with information-theoretic security guarantees can help prevent such impersonation threats.
- *Threats to Trust Management [T3.5].* In the dynamic environment of SemComNet, where autonomous agents frequently change their behavior (e.g., uploading sensory data), managing trust is crucial to mitigate interaction risks [145]. However, establishing an effective trust management system is challenging due to the need for real-time monitoring of agent behaviors and reputations [135], which may compromise confidentiality. Furthermore, agent interactions often involve limited or incomplete data sharing, increasing the complexity of determining the reputation of other agents and establishing trust relationships. To address these issues, trust-free architectures (discussed in Section IV-D2) and efficient trust evaluation mechanisms (covered in Section IV-D3) offer alternative solutions to enhance trust management in SemComNet.
- *Malware Attack [T3.6].* This attack compromises service availability, confidentiality, and integrity of data/knowledge in SemComNet. Specifically, malware can swiftly spread between agents, causing deviations in their behavior patterns and roles. For instance, malicious code or instructions [81] injected into a benign agent can turn it into an aggressive entity. In addition, certain malware families such as ransomware, spyware, and worms may steal private data, erase sensory data, or delete accumulated knowledge [151]. To resist these threats, malware detection, defense, and prevention are crucial [152]. Malware detection helps identify threats before damage occurs [151], while defense and prevention focus on mitigating and reversing malware activities before and after infection [106]. In practice, a multilayered approach combining these strategies can enhance resilience against complex malware threats in SemComNet.

D. Summary and Lessons Learned

SemComNet inherit several security/privacy threats from traditional networks, with varying impacts. These can be classified into: i) enhanced impact of existing attacks. Traditional attacks, such as threats from relay nodes, become more severe in SemComNet, leading to increased risks of privacy breaches and semantic manipulation. ii) Reduced impact of existing attacks. Some existing attacks are easier to defend in SemComNet. For instance, semantic eavesdropping is harder to execute, as decoding intercepted SI without proper codecs and knowledge is highly challenges for attackers. iii) Variant of existing attacks. Traditional attacks such as DoS evolve into sponge example attacks, exploiting SemComNet's reliance on computational resources to cause excessive energy consumption and delays. Additionally, some brand-new threats arise from the cutting-edge technologies underpinning SemComNet, such as its reliance on multiple KBs and AI techniques. These introduce risks such as KB poisoning, unauthorized access, and desynchronization attacks.

Besides, cross-layer attacks present significant challenges in the multi-layered architecture of SemComNet. Such attacks exploit interactions and dependencies among layers, complicating detection and defense. Attackers can exploit the dependencies between layers, amplifying the effects of their attacks. For instance, cognitive manipulation in the sensing layer may contaminate agents' private KBs, which can then spread to shared KBs in the control layer. The contaminated shared KBs can adversely affect the semantic transmission layer, leading to incorrect semantic interpretations and reduced communication performance. The interconnected nature of SemComNet amplifies potential damage, highlighting the need for robust, multi-layered defense mechanisms to detect and mitigate threats at every level.

In summary, we have categorized the security and privacy threats of SemComNet (i.e., from Sections III-A–III-C) across their three functional layers, i.e., control layer, semantic transmission layer, and cognitive sensing layer. Besides, as depicted in Fig. 9, we have reviewed existing/potential defense approaches for the above security/privacy issues within SemComNet from five perspectives: semantic model security, semantic transmission security, reliability, trust, and data/knowledge security.

IV. Security and Privacy Countermeasures in SemComNet

In this section, we provide an in-depth and up-to-date discussion of security and privacy defenses tailored for SemComNet across five crucial aspects, i.e., semantic model security, semantic transmission security, reliable SemComNet, trust management in SemComNet, and data & knowledge security in detail.

A. Semantic Model Security

Advanced AI techniques (e.g., DL and RL) empower semantic codecs to efficiently extract and interpret SI. As the performance of SemComNet is heavily dependent on the capabilities of these AI-driven semantic codecs, ensuring their security becomes crucial [2], [31]. In the following, we discuss countermeasures to protect semantic models from poisoning attacks, AEs, and sponge examples, as well as strategies to safeguard the IP of these models.

1) Defense Strategies Against Semantic Model Poisoning: Semantic model poisoning attacks directly manipulate the parameters of semantic codecs, rather than contaminating training data. In centralized learning, defenses against model poisoning focus on model sanitization, which detects and mitigates malicious parameter modifications. Techniques such as fine-pruning, which removes dormant or suspicious neurons and retrains the model on clean data, can restore normal

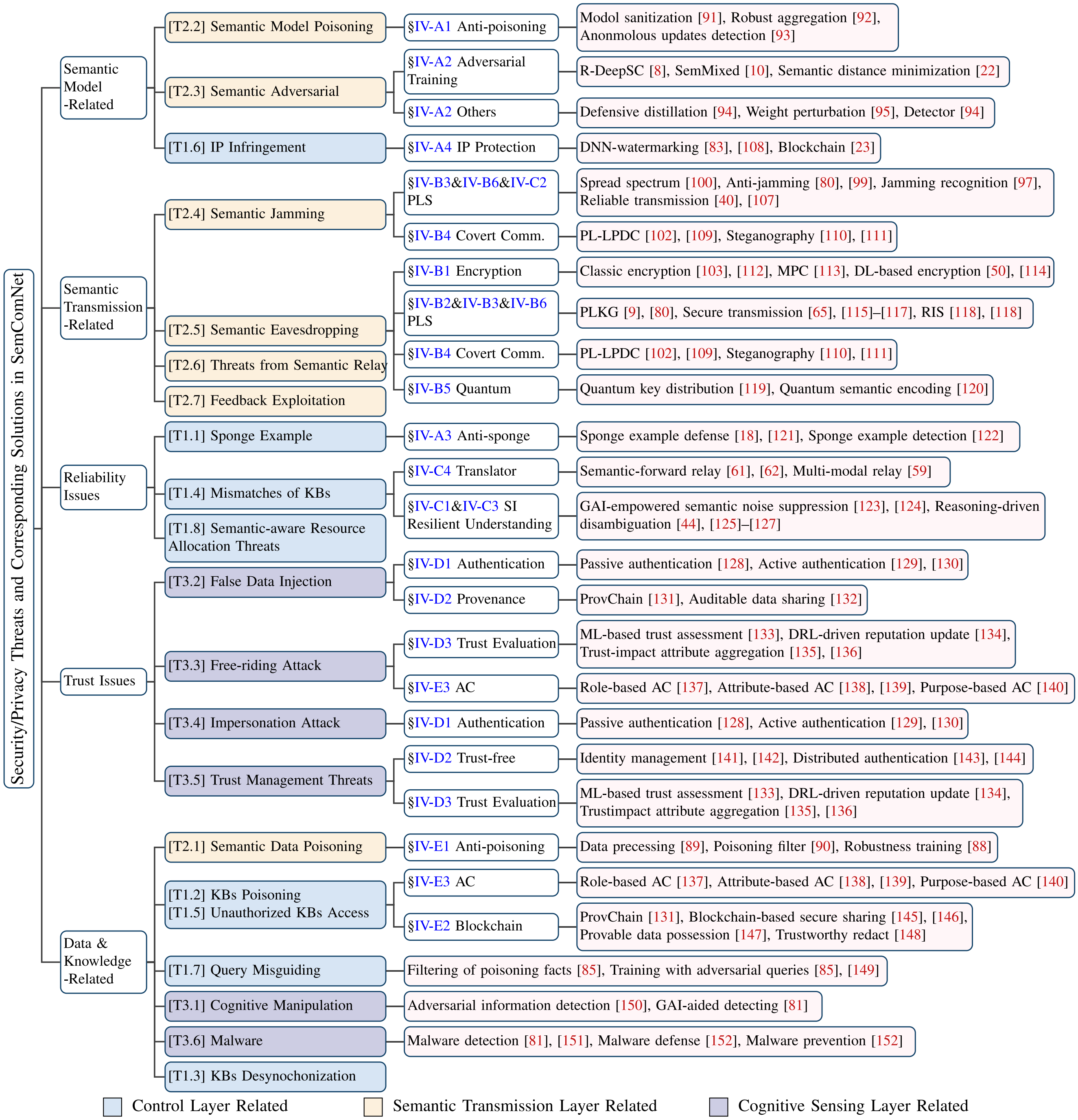


Fig. 9. The taxonomy of security/privacy threats to SemComNet from five aspects (i.e., model, transmission, reliability, trust, and data & knowledge) and corresponding existing/potential defenses in the SemComNet.

model function. For instance, Wang et al. [91] present a system for detecting and mitigating DNNs backdoor attacks by exploiting the model's sensitivity to input perturbations. However, this method faces scalability issues for large models and assumes the availability of clean training data. In distributed learning, robust aggregation methods such as secure aggregation protocols [92] help prevent malicious updates. For instance, Krum as a Byzantine-resilient aggregation rule, filters out abnormal model weights by sorting and selecting participants' updates, thus enhancing the robustness of the global model [92]. However, Krum's complexity increases with more participants due to pairwise distance calculations. An alternative approach involves detecting malicious clients. For instance, Zhao et al. [93] propose a client-side cross-validation scheme where each update is evaluated using other clients' local data. This method adjusts update weights based on the evaluations during aggregation.

2) Adversarial Training for Semantic Noise Resistant SemComNet: SemComNet are susceptible to semantic adversarial attack due to the vulnerability of DNN-based codecs and the broadcast of wireless medium [10], [19], [22]. These attacks may induce semantic noise that will gradually distort

TABLE VI
SUMMARY OF KEY LITERATURE ON SEMANTIC MODEL SECURITY IN SEMANTIC COMMUNICATION NETWORKS

Ref.	Security Threat	⋆ Purpose • Advantages ◦ Limitations † Evaluation Metrics	Utilized Technology
[91]	Semantic backdoor attack	⋆ Detect whether a model is backdoored exploit activation statistics • Robust and general tools for detecting and mitigating backdoor attack ◦ Unscalable to large-scale model and strong assumption (e.g., clean training dataset) † Attack success rate, classification accuracy, and false positive/negative rate	Backdoor detection, optimization
[93]	Semantic poisoning attack	⋆ Leverage clean data through cross-validation for detecting poisoning updates • High defense strength and robustness without compromising model accuracy ◦ Limited by validation data size and susceptible to malicious clients † Model accuracy, computation cost, and probability of evading the detection	Model validation
[8]	Semantic adversarial attack	⋆ Enhance the resilience of semantic codecs against adversarial examples • High robustness, semantic fidelity, and compatibility with existing training pipelines ◦ Increased computational cost, difficulty in converging, and low defense generalization † BLEU score and BERT score	Adversarial training
[108]	Semantic model stealing, surrogate model attacks	⋆ Trace illegal usage and copyright verification of AI models • High watermark capacity and strong generalization ability ◦ Cannot resist ambiguity attack and lack resilience against pre-processing techniques † Peak SNR, SSIM, normalized correlation, and success rate of watermark extracted	Model watermarking

the desired meaning conveyed in SI. Current defense schemes span diverse categories [94], i.e., adversarial training, defensive distillation, AEs detector, and methods involving weight concealment and interference [95]. Among these approaches, adversarial training stands out as a simple yet highly effective method. It has been extensively studied and proven to enhance the resilience of semantic models to AEs [8], [10], [19]. This approach involves incorporating AEs into the training data and continually generating new AEs during each iteration of training process. For instance, Peng et al. [8] employ adversarial training, specifically the fast gradient method to identify perturbations that significantly disrupt the system and subsequently train the system to withstand these AEs. In [8], the *Bilingual Evaluation Understudy* (BLEU) score is utilized to assess the semantic quality of received data, and the *Bidirectional Encoder Representations from Transformers* (BERT) score measures the semantic similarity between text sentences. The results, indicated by high BLEU and BERT scores, demonstrate the proposed scheme's effectiveness in mitigating semantic distortions caused by adversarial noise.

However, this training strategy might increase computational costs and may not generalize well to AEs from other adversaries because it trains the model on narrowly crafted AEs, which increases the risk of overfitting. Besides, the work [8] does not consider the interference caused by the massive connections between agents, which can introduce various adversarial perturbations in wireless channels. To simulate this, Nan et al. [10] introduce *SemAdv* to create semantic-oriented perturbations for physical-layer adversarial attacks during SI transmission. To defend this, they propose an adversarial training approach *SemMixed* to enhance the resilience of SemComNet against various physical adversarial perturbations. The experiments not only use standard evaluation metrics but also introduce a new security metric called *misleading rate*, which evaluates the model's tendency to mislead predictions toward specific image categories. Simulation results highlight that the proposed defense strategy significantly reduces the attack success rate.

3) Safeguarding SemComNet Against Sponge Examples: To mitigate the threat posed by sponge examples, a straightforward solution is to adopt worst-case performance analysis, as proposed by Shumailov et al. in [18]. By establishing processing time and energy thresholds from model profiling with natural examples, inputs surpassing these limits will be rejected, thereby helping to ensure system availability. However, this solution has limitations. Specifically, setting optimal thresholds requires careful calibration tailored to the hardware platform and model performance. If thresholds are not correctly set, the defense could yield suboptimal results. Additionally, attackers could manipulate sponge examples to stay within these thresholds while still consuming excessive resources and causing delays. With large volumes of sponge examples, the system's resources may be overwhelmed, potentially leading to a DoS situation. Thus, relying solely on threshold-based defenses may be insufficient. To bolster the robustness of SemComNet, supplementary defense mechanisms from both defense [121] and detection [122] could be considered. For instance, Wong et al. [121] propose training models with AEs generated using techniques such as FGSM. This approach improves semantic models' resilience to perturbations and reduces their susceptibility to manipulation by sponge examples.

When direct defense strategies are not feasible, detection becomes critical. For instance, Chen et al. [122] introduce a novel detection scheme by analyzing the sequence of queries made to the victim model. Unlike traditional defenses that focus on identifying malicious inputs, they leverage the temporal nature of attacks and track the similarity between successive queries. By calculating the k-nearest-neighbor distance between a new query and stored examples, the system can detect suspicious patterns indicative of ongoing attacks. Simulation results show its effectiveness even under closed-box attack scenarios, where attackers may use query blinding techniques to obscure the sequence. Additionally, this detection method is compatible with existing defenses against zero-query attacks [122], offering a comprehensive solution for detecting adversarial examples.

4) IP Protection for Securing Model in SemComNet: Powerful semantic codecs are crucial for intelligent communication tasks in SemComNet, but training process is often expensive and time-consuming. To protect the IP of these models, watermarks are vital [83], [108]. In [83], Zhang et al. extend digital watermarking from multimedia to DNNs and propose three DNN-compatible watermark generation algorithms to confirm model ownership via detecting preset patterns. However, the work [83] fails to resist surrogate model attack[5]. In response, Zhang et al. [108] introduce a task-agnostic method that embeds a spatially invisible watermark within the networks' outputs, maintaining high extraction success rates even against surrogate models. Apart from watermark-based IP protection, blockchain technology offers a trustworthy and traceable mechanism for semantic codecs and knowledge sharing within SemComNet [23], providing an additional layer of trustworthiness and transparency.

B. Semantic Transmission Security

Existing and potential solutions to ensure transmission security in SemComNet include cryptography, physical layer security (PLS), covert communication, quantum technology, and other emerging communication techniques, as detailed below.

1) Semantic Data Encryption for Secure Transmission: Encryption schemes play a crucial role in safeguarding the privacy, security, and integrity of semantic data, thereby enhancing the overall reliability of SemComNet, which has received increasing attention [103], [113], [156]. In [103], Chen et al. propose a cryptography method using random permutation and substitution to counter model inversion eavesdropping attack [21], where attackers intercept and reconstruct the original message. Experiments show this approach is effective in both glass-box and closed-box scenarios. Similarly, Chen et al. [156] employ RSA and AES algorithms in SemCom to prevent eavesdroppers from accessing sensitive information, reducing eavesdropping accuracy (i.e., eavesdropping ACC) to approximately 10% on CIFAR10 and ImageNet datasets.

However, this approach [103] relies heavily on key management and distribution, posing risks associated with key exposure. To address these challenges, advanced cryptographic techniques such as secure multi-party computation (MPC) [113] and homomorphic encryption (HE) [113] eliminate the need for key sharing. MPC enables agents in SemComNet to perform collaborative computations (e.g., addition, comparisons, and searches) without revealing raw data. To promote its integration into ML, Knott et al. [113] propose a user-friendly framework named CRYPTEN by offering MPC services through familiar ML abstractions. Meanwhile, HE enhances data security by allowing computations directly on encrypted data, ensuring privacy during transmission and processing without requiring decryption.

However, the above solution [113] may impose excessive communication and computing burdens on agents. Moreover, traditional encryption [103] such as permutation and substitution disrupt intrinsic SI correlations, reducing reconstruction quality. To solve this, Xu et al. [50] propose a DL-based joint encryption and source-channel coding (DJESCC) for SemComNet. This approach secures visual content by transforming images into visually protected forms while preserving reconstruction quality. However, encrypting and decrypting entire information may result in reconstruction errors. In response, Sun et al. [155] introduce a privacy-preserving JSCC scheme named DPJSCC, as shown in Fig. 10. It separates private and public information into distinct subcodewords using disentangled IB. The private subcodewords are encrypted with a password-based algorithm, while public ones are directly transmitted. DPJSCC is effective even against advanced eavesdroppers, achieving eavesdropping accuracy near random guessing while maintaining strong reconstruction quality. Experiments show DPJSCC reduces eavesdropping accuracy by up to 18% compared to DeepJSCC [157] and adversarial JSCC [158].

2) Physical-Layer Lightweight Secret Key Generation for SemComNet: Physical layer key generation (PLKG) emerges as a quantum-resistant and key transmission-free solution, offering a promising avenue for securing SemComNet. By leveraging the unique properties of the transmission medium, such as wireless channel fading, PLKG enables secret key generation directly between communicating parties. For instance, Zhao et al. [9] present a PLKG scheme using a reconfigurable intelligent surface (RIS) [159] to enhance the key generation rate by exploiting the randomness of semantic drifts between the transmitter and receiver. However, the idealized assumption of spatial constraints (e.g., eavesdroppers being half a wavelength away) in [9] may not hold in real-world scenarios, where eavesdroppers could occupy various positions, potentially impacting system vulnerability differently. Additionally, PLKG can integrate seamlessly with hybrid cryptographic schemes, as demonstrated in [80], combining traditional cryptographic strengths with PLKG's unique advantages. This integration bolsters security and efficiency, making PLKG an advanced solution for safeguarding SemComNet.

3) Physical Layer Secure Transmission for SemComNet: Since its inception [160], physical layer secure transmission technology has attracted substantial attention [96]. This technology seamlessly integrates the aspects of data transmission and encryption, offering a promising approach for securing SemComNet. Currently, there are several key technical approaches [96]: beamforming [116], power allocation [19], cooperative interference [117], and artificial noise (AN) injection [115], [161]. These methods intentionally amplify the gap between the legitimate and eavesdropping communication channels. This allows authorized recipients to decode confidential SI successfully, while semantic signals at eavesdroppers are deliberately scrambled and irreversibly corrupted.

[5] A surrogate model attack targets the IP of DL models by creating a substitute model that mimics the target model's behavior. In a closed-box scenario, the attacker does not have access to the target model's internal structure or parameters but can observe its input-output pairs. By feeding input samples into the target model to obtain predictions, the attacker constructs a training dataset and trains a surrogate model to approximate the target model. After several iterations, the attacker successfully creates a model that behaves similarly to the original, thus stealing its IP.

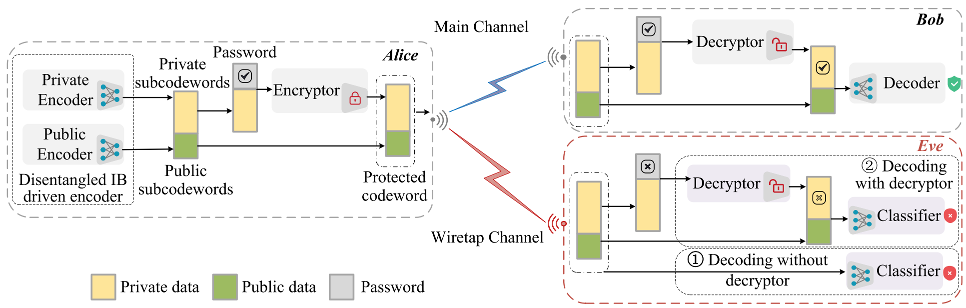


Fig. 10. Illustration of the DPJSCC scheme [155], which depicts the disentangled information bottleneck (IB) guided privacy-protective JSCC framework for image transmission.

(i) Beamforming. Implementing the beamforming technique in SemComNet makes the SI transmission in specific directions which ensures only intended receivers can capture it. For instance, Lin et al. [116] propose a frequency diverse array beamforming approach that optimizes frequency offsets and transmits beamforming to maximize secrecy rates, particularly in close legitimate user and eavesdropper scenarios. *(ii) Power allocation.* A well-designed power allocation scheme enhances the security of the transmission significantly [19]. For instance, Yin et al. [65] introduce a multi-domain resource multiplexing scheme leveraging co-channel interference among IoT nodes. Their alternating optimization with successive convex approximation methods enhances security of physical layer. *(iii) Cooperative interference.* To effectively disrupt unauthorized channels, cooperative relay, and friendly jamming techniques can be introduced. In [117], Li et al. introduce two innovative cooperative interference alignment (IA) schemes, the former adjusts the spatial signature of one interference and strength of all interference, ensuring orthogonality with the desired transmission, and the other modifies all interfering signal strength, preserving orthogonality. *(iv) AN injection.* As traditional IA methods are prone to secret signal cancellation, Hu et al. [115] propose an AN-assisted IA scheme that minimizes secrecy outage probability through optimized power allocation. Although these techniques are effective at the physical layer, achieving holistic security across SemComNet requires exploring cross-layer optimization, such as optimizing and integrating PLS schemes with other layers (e.g., the control layer) within SemComNet.

4) Covert Communication for Secure Transmission: In SemComNet, as shown in Fig. 11, semantic data encryption ensures content security through advanced encryption [50], [114], [155], but it incurs higher computational complexity and key management overhead. PLS protects the content by ensuring the message remains inaccessible to eavesdroppers, even if communication is detected. Key performance indicators for PLS include *secrecy outage probability* (SOP) and *average secrecy capacity* (ASC) [19], [96]. In contrast, as shown in Fig. 11, covert communication focuses on undetectable transmissions. It conceals the transmission itself, preventing detection by a warden, and is measured by metrics such as *covert communication capacity* [162], *detection error probability* (DEP), and *covert rate* [19], [101], [102].

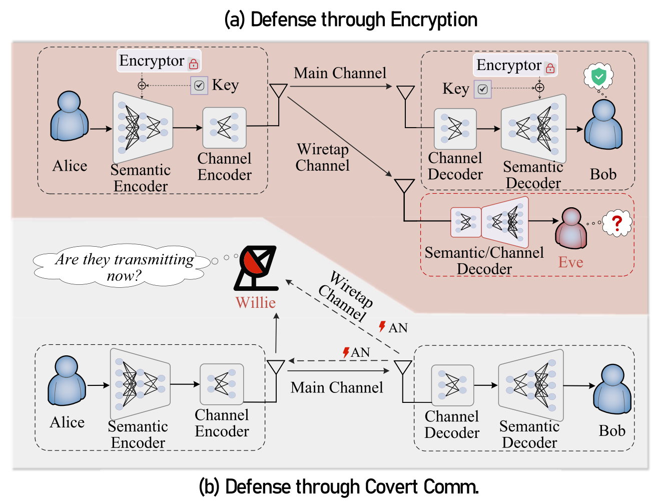


Fig. 11. An illustrative comparison of defense mechanisms through semantic encryption [114] and covert communication [102].

One paradigm of covert communication is physical-layer low probability of detection communication (PL-LPDC). For text semantic transmission in wireless networks, as shown in Fig. 11, Hu et al. [102] propose a covert communication scheme using a full-duplex receiver to generate AN, which interferes with Willie's detection. By jointly optimizing transmit power, AN power, and the number of semantic symbols per word, they maximize semantic spectral efficiency while ensuring minimum semantic similarity and maintaining covertness. Detection performance is evaluated using the false alarm rate and missed detection rate. Their analysis shows that a fixed AN power can enhance covert transmission. For image transmission, Wang et al. [109] present a PL-LPDC framework, where friendly jammers deploy jamming signals to defend against eavesdroppers. A multi-agent policy gradient algorithm is proposed to improve system performance, allowing devices and jammers to identify vulnerable devices and optimize transmission parameters in SemComNet. However, the lack of complexity and scalability analysis, along with the intricate training procedures, may limit its practical deployment in SemComNet.

Unlike PL-LPDC, which focuses on concealment and low detection probability, physical-layer steganography [110] aims to hide and encrypt the content of communication, offering

TABLE VII
SUMMARY OF KEY LITERATURE ON SEMANTIC TRANSMISSION SECURITY IN SEMANTIC COMMUNICATION NETWORKS

Ref.	Security Threat	⋆ Purpose ● Advantages ○ Limitations † Evaluation Metrics	Utilized Technology
[114]	Semantic eavesdropping attack	⋆ Encrypts SI to protect against unauthorized access during transmission ● Ensures the confidentiality and integrity of transmitted SI ○ Low scalability, complicated key management, and high processing delay † BLEU score and convergence speed	Symmetric encryption, adversarial training
[9]	Semantic eavesdropping, spoofing attacks	⋆ Generate secret key by leveraging semantic drifts and update under RIS assistance ● High-security guarantee, fast key generation, and enhanced channel randomness ○ Idealized spatial constraint and simplified channel model † Key generation rate, P-value, and randomness pass rate	Physical layer key generation, RIS
[115]	Passive eavesdropping, interference attacks	⋆ Artificial noise-assisted interference alignment for mitigating eavesdropping ● Practical assumption, enhanced security, and mitigated interference ○ Lack cross-layer optimization and may impact transmission efficiency † Secrecy outage probability and power allocation ratio	Interference alignment, artificial noise
[109]	Semantic eavesdropping, jamming attacks	⋆ Covert and secure SI transmission assisted with friendly jammer ● Low probability of detection and suitable for highly confidential scenarios ○ Lack complexity and scalability analysis, as well as complex training procedures † Average accuracy of answering the questions	Covert comm., multi-agent RL
[120]	Semantic poisoning, adversarial attacks	⋆ Quantum SemCom for reliable and secure interaction ● Strong security guarantees and future-proof against quantum attacks ○ Challenging in fragile/expensive photonic quantum resources and scalability issues † Distance of semantic Hilbert space, semantic decoding error, and fidelity	Quantum embedding, quantum ML
[154]	Unauthorized access, eavesdropping attack	⋆ RIS-assisted secure and inverse semantic-aware wireless sensing ● Efficiency in transmission, enhanced security, cross-layer compatibility ○ Hardware cost and implementation complexity, and productive attenuation	RIS

a higher level of confidentiality. It embeds secret SI within the physical properties of the communication channel (e.g., signal power, phase, or frequency), making it imperceptible to unauthorized users. For instance, Yamaguchi et al. [111] propose a steganography security approach that conceals the secret signal by embedding it into the cover data, making detection difficult without prior knowledge. Future research efforts are required in designing intelligent covert communication in SemComNet, integrating adaptive optimization, theoretical completeness, and lightweight implementation to accommodate diverse environmental conditions.

5) Quantum Technology for Secure-Enhanced SemComNet: Quantum technology, based on the principles of quantum physics, can significantly enhance security within SemComNet [120], [163]. Quantum key distribution (QKD) plays a pivotal role in establishing secure key exchanges in SemComNet, providing a tamper-evident framework for SemComNet. Unlike classical cryptographic methods, QKD can detect eavesdropping attempts. Any interference is rapidly identified by the communicating parties, alerting them to potential intruders and safeguarding the confidentiality of semantic data.

In [119], Kaewpuang et al. focus on QKD-aided secure SI transmission within SemComNet. Furthermore, they address resource allocation challenges in QKD deployment for SI transmission by proposing a two-stage stochastic optimization model that optimizes QKD resource deployment, with Shapley values ensuring fair cost allocation among cooperative QKD service providers. Apart from QKD, quantum semantic encoding can further protect the privacy of semantic content, offering enhanced security and efficient data transmission. For instance, Khalid et al. [120] explore quantum SemCom employing quantum embedding and quantum ML to encode data into quantum states, which are securely teleported using quantum principles. However, such quantum solutions face challenges in resource optimization and scalability, primarily due to the reliance on fragile and expensive photonic quantum resources.

6) Emerging Communication Technology Assisting Secure SemComNet: RIS, an emerging wireless technology [159], can enhance the security and efficiency of SI transmission in SemComNet. For instance, Wang et al. [118] propose using a simultaneous transmitting and reflecting RIS to prevent eavesdropping during SI transmission. By optimizing transmission and reflection coefficients, they enhance legitimate semantic signal transmission and create interference for eavesdroppers. In [154], Du et al. present an inverse semantic-aware wireless sensing framework for SemComNet, which uses the RIS amplitude response matrix for secure data encryption and semantic hash sampling for efficient self-supervised decoding. Experimental results show a 95% reduction in data volume without affecting sensing tasks, offering a resource-saving solution for the secure SemComNet.

Furthermore, the above passive RIS schemes [118], [154] face challenges, such as productive attenuation where reflection link fading is proportional to distances. To tackle this issue, Li et al. [164] introduce active RIS to reconfigure the propagation within the wireless environment, while utilizing the on-off control scheme for active RIS phase shifts. Apart from RIS, spread spectrum techniques such as frequency hopping spread spectrum and direct sequence spread spectrum can also be utilized in SemComNet [100]. As such, the semantic signal is spread across a wide spectrum of frequencies, or pseudo-random sequences are employed. This helps mitigate the disruptive effects of interference on the channel, ensuring a robust and secure communication environment.

TABLE VIII
SUMMARY OF KEY LITERATURE ON ENSURING RELIABLE SEMCOMNET

Ref.	Security Threat	⋆ Purpose • Advantages ◦ Limitations † Evaluation Metrics	Utilized Technology
[123]	Semantic noise, distortion	⋆ Leverage GAI model to enhance perceptual quality and semantic reliability • Improved communication efficiency and reduced misunderstanding errors ◦ High hallucination risk and heavy computational burden on agents † Peak SNR, MS-SSIM, and LPIPS	StyleGAN
[165]	Semantic noise, eavesdrop and interference attack	⋆ Covert SI transmission and accurate SI decoding via multi-modal GAI models • Strong secure guarantee and accurate content generation capability ◦ Complex training procedures and costly resource consumption † SSIM, detection error probability, and bit error probability	GAI, multi-modal prompts, covert comm.
[107]	Physical and semantic interference	⋆ Adaptive bit rate control mechanism for SI transmission under harsh conditions • High efficiency, strong fault tolerance, and cost-saving ◦ Lack of large-scale and real-world test † BLEU score	Semantic HARQ, adaptive transmission
[44]	Semantic interference, and privacy exposure of KBs	⋆ Leverage RL for reasoning implicit SI while protecting privacy • Automatic reasoning and robust semantic error correction ◦ Low sample efficiency and lack generalization to unseen data † Semantic reasoning accuracy, average symbol recovery accuracy, and symbol error rate	Inverse RL
[61]	Mismatch of KBs between transceiver	⋆ Assist in forwarding and interpreting SI while minimizing semantic noise • Extended communication range and reliable translation capability ◦ High computational complexity and long transmission delay † BLEU score and sentence similarity	Semantic relay

C. Reliable SemComNet

In SemComNet, reliability stands as a vital property facilitating semantic-oriented service provisioning and effective agent interaction across diverse environments. Specifically, ensuring a reliable SemComNet involves considerations spanning two dimensions, i.e., resilient semantic interpretation and reliable SI transmission. The former dimension guarantees accurate understanding and reconstruction of semantic content, thereby enhancing tolerance to semantic ambiguity. On the other hand, the latter dimension focuses on maintaining stable SI transmission performance even in harsh conditions, such as extra-low SNR, strong perturbation [166], and long transmission distances. We delve into discussions regarding enhancing SemComNet's reliability through various approaches, as below.

1) GAI for Misunderstanding Mitigation in SemComNet: GAI models such as ChatGPT can enhance the resilience of semantic interpretation (e.g., suppressing semantic noise) within SemComNet [97]. Specifically, these models aid in reducing transmission traffic and latency [51]. They also excel in reconstructing semantic-consistent details from SI at the destination side, even when encountering challenges such as semantic noise and mismatches in KBs between the transmitter and receiver. For instance, to enhance the perceptual quality of reconstructed data, Erdemir et al. [123] introduce two innovative GAI-based JSCC frameworks: InverseJSCC and GenerativeJSCC. Unlike traditional approaches focusing solely on distortion metrics, the schemes proposed in [123] optimize a combination of *mean squared error* (MSE) and *learned perceptual image patch similarity* (LPIPS) losses to ensure semantic fidelity. Simulations show that the proposed scheme exceeds DeepJSCC not only in distortion metrics such as *peak signal-to-noise-ratio* (PSNR) and *multiscale structural similarity index measure* (MS-SSIM), but also in perceptual quality measured by LPIPS, even with KBs mismatches. To achieve superior perceptual quality in limited bandwidth and low SNR scenarios, Chen et al. [124] treat image recovery as an inverse problem. They employ invertible neural networks with the diffusion model to aid the reconstruction process. However, the above works [123], [124] neglect the aspect of privacy protection during transmission in GAI-driven SemComNet. To address this, as shown in Fig. 12, Han et al. [167] leverage the StyleGAN inversion method to extract disentangled SI from the original image. Meanwhile, they employ privacy filters to replace private SI with natural features guided by KBs, thereby safeguarding sensitive SI. The results have proven successful in enhancing communication efficiency and reducing misunderstanding errors.

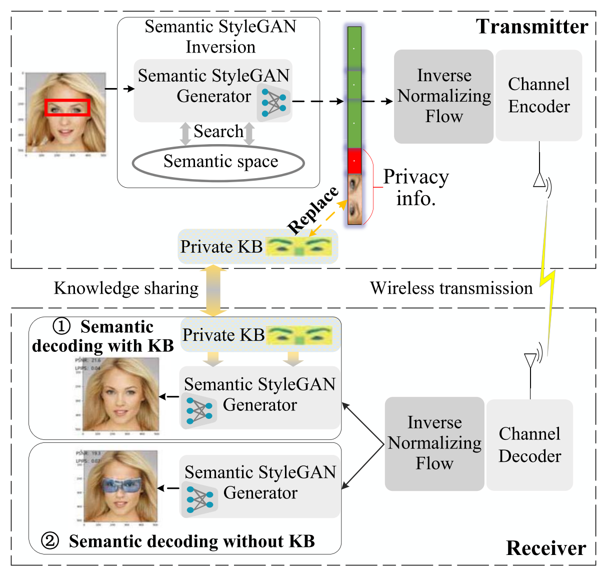


Fig. 12. Illustration of a reliable and privacy-preserving SemCom system for image transmission employing GAI and privacy filters in [167]. This system accurately reconstructs semantically consistent details from SI at the destination, effectively mitigating semantic noise risks. Besides, it utilizes a privacy filter and KBs to replace privacy information with corresponding natural features, ensuring privacy protection.

TABLE IX
SUMMARY OF KEY LITERATURE ON TRUST MANAGEMENT IN SEMANTIC COMMUNICATION NETWORKS

Ref.	Security Threat	⋆ Purpose • Advantages ◦ Limitations † Evaluation Metrics	Utilized Technology
[130]	Impersonation attack, eavesdropping attack	⋆ Leverages unique physical layer properties for secure authentication • Information-theoretic guarantee, lightweight processing speed, and high compatibility ◦ Challenging in imperfect channel conditions and large-scale environments † Secure authentication efficiency	Physical layer authentication
[141]	False data injection, impersonation attack, malware attack	⋆ Provide robust privacy-preserving identity management for industrial SemComNet • Guarantee unforgeability, traceability, revocability, and public verifiability ◦ Without consideration of privacy-preserving and processing speed † Computation cost, communication cost, and gas cost	Cryptographic tool, blockchain
[132]	Free-riding attack, misuse of shared knowledge/data	⋆ Trust-free environment for fine-grained data authorization and traceable audit • Low computation overhead in shared knowledge/data audit ◦ Lack large-scale and real-world performance test	Smart contract, trusted computing
[134]	False feedback attack, internal attack	⋆ Dynamic reputation mechanism for limiting wrong feedback from malicious agents • Enhanced system reliability and adaptability to agent behavior ◦ Overheads grow exponentially as the number of agents rises † Reputation and average reward	Dempster-Shafer theory, DRL
[136]	Semantic-poor data quality	⋆ Reputation evaluation model to filter out trusted agents with semantic-rich data • Strong reliability and practicality via real-world validation ◦ Difficulty in obtaining trust indicators † Reputation and QoS score	Experience-Reputation trust evaluation

However, a notable challenge arises from the inherent instability generation capabilities of GAI models, such as the issue of *hallucinaion* [165], [168]. This instability poses challenges, particularly in applications requiring accurate information transfer (e.g., medical imaging diagnosis). To address this issue, Du et al. [165] propose a GAI-aided approach that employs multi-modal prompts for accurate content decoding. By leveraging generative diffusion models and covert communication, the approach facilitates the transmission of multi-modal prompts and ensures precise image regeneration under energy constraints. Nonetheless, the integration of large-scale GAI models into SemComNet presents challenges, demanding substantial computing power for training. This strains the deployment of GAI models in resource-limited agents within SemComNet. In summary, achieving trustworthy and resource-efficient integration of GAI into SemComNet necessitates further investigation.

2) Reliable Adaptive Transmission in SemComNet: In dynamic network conditions and deteriorating channel quality, ensuring reliable SI transmission is crucial. Adaptive transmission offers a promising solution to enhance SemComNet reliability [31], [32]. For instance, to improve noise resilience in ultra-low SNR scenarios, Zhou et al. [107] propose a multi-bit length selection strategy with a policy network to dynamically adjust coding rates. They also introduce progressive semantic hybrid automatic repeat request (HARQ) schemes, incorporating incremental knowledge to reduce semantic errors. Results, evaluated using the BLEU metric, show improved efficiency, robustness, and cost-effectiveness. However, the absence of extensive, real-world testing may limit the validation of these findings. Moreover, Wang et al. [40] explore an adaptive bitrate transmission strategy for video data. Using nonlinear transformations and conditional coding, they extract SI from video frames and allocate channel bandwidth dynamically for optimal performance under challenging wireless conditions.

3) Reasoning-Driven Reliable SemComNet: Agents endowed with reasoning capabilities significantly enhance the reliability of SemComNet by mitigating semantic ambiguity and enhancing communication reliability. For instance, Jiang et al. [125] introduced a KG-driven SemComNet which could significantly enhance reliability against semantic noise. However, the above work [125] does not consider the issue of typing errors in large-scale factual KGs within SemComNet. These errors, especially in entity-type pairs, pose a significant obstacle to reliable knowledge extraction and utilization in SemComNet. To address this challenge, Yao and Barbosa [126] propose an innovative active typing error detection algorithm that effectively incorporates both gold and noisy labels. Furthermore, they emphasize the semi-supervised noise models as a feasible solution, offering the possibility to enhance the utilization of varied information for error detection in SemComNet.

RL excels at reasoning implicit SI from transmitted data, which could seamlessly adapt to the dynamic conditions within SemComNet. For instance, an implicit semantic-aware communication architecture is proposed in [44], focusing on both explicit and implicit message semantics reasoning. The proposed generative imitation-based reasoning mechanism could guide destination users to automatically interpret implicit semantics without requiring access to the source data, significantly safeguarding user privacy and reducing semantic noise. However, this method has drawbacks, such as low sample efficiency and limited generalization to unseen data. In response, Thomas and Saad [127] integrate a signaling game and Neuro-Symbolic (NeSy) AI [169] into semantic transmission. The game-theoretical approach creates a compositional language, aiding in generalization and semantic awareness. The NeSy AI integrates experiential learning (neural component) with knowledge-based reasoning (symbolic component), empowering the SemComNet to acquire intricate signaling strategies with limited training samples and minimal data transmission. Currently, GAI models [43], [168], [170] serve as powerful tools with semantic understanding and reasoning capabilities. Leveraging these models could significantly enhance reasoning accuracy in SemComNet [42], [51], [170].

4) Collaborative Relay for Reliable SemComNet: In traditional communication systems, collaborative relays ensure reliable information delivery, particularly in long-distance and low-SNR scenarios [59]. Similarly, semantic relay nodes in SemComNet not only assist in reliable transmission but also help translate and interpret SI from source to destination [62]. For instance, Luo et al. [61] introduce the intelligent relay-assisted SemCom with amplify-and-forward (AF) and decode-and-forward (DF) modes. This system assists in forwarding and interpreting SI while minimizing semantic noise. However, it faces challenges with high computational complexity and long transmission delay.

D. Trust Management in SemComNet

Establishing and managing trust among collaborative agents in SemComNet is crucial for building a trustworthy environment. In response, this section will introduce three approaches for creating a trustworthy SemComNet, i.e., intrinsic trust via physical-layer authentication mechanism, trustless SemComNet through blockchain, and measurable trust evaluation.

1) Authentication Mechanism in SemComNet: Authentication serves as a fundamental component of trust management by providing a preliminary verification of agents' identities, forming the basis for trust establishment and subsequent evaluations. Conventional cryptography-based authentication methods face scalability and complexity challenges in large-scale SemComNet. In contrast, physical layer authentication (PLA) is gaining considerable attention due to its information-theoretic security guarantee, lightweight processing speed, and high compatibility [171]. This approach is increasingly recognized as a viable implementation in heterogeneous and decentralized SemComNet.

PLA schemes can be categorized as passive and active [171]. Passive PLA relies on physical-layer features for transmitter authentication without modifying the source message. For instance, Gao et al. [128] introduce *EsaNet*, a DL-based passive authentication network that extracts a wireless channel fingerprint from environmental semantics (i.e., CSI) to distinguish legitimate users. In contrast, active PLA modifies the source message by embedding a physical-layer tag generated from a secret key, enhancing information-theoretic security according to Shannon's secrecy analysis. In [129], a generalized model for achieving data confidentiality and active wireless PLA is proposed. It highlights the role of channel uncertainty and various design dimensions, such as time, frequency, and space, in enhancing security for SemComNet. To evaluate the transmission efficiency of active PLA schemes, Tan et al. [130] introduces a metric named *secure authentication efficiency* (SAE). By tuning three key parameters that impact SAE (i.e., the probability of message transmission, the probability of message outage, and the probability of secure authentication), they provide a systematic optimization framework and analyze feasibility constraints and optimal solutions.

However, the aforementioned works [129], [130] assume perfect channel conditions, which is unrealistic in real-world scenarios. To address this, Perazzone et al. [172] explore the use of AN in fingerprint embedding for wireless security, focusing on scenarios with imperfect CSI. They examine the impact of AN leakage on security and compare detectors for imperfect CSI. The findings reveal that while AN improves security, its effectiveness diminishes under poor channel conditions.

2) Blockchain for Trustless SemComNet: Blockchain integration in SemComNet is gaining attention for its transparency, decentralization, and tamper-proof features [173]. By enabling decentralized identity management [141], [142], cross-domain authentication [143], [144], and smart contract-driven authorization [132], blockchain provides a robust trust management solution. It eliminates the need for a central certification authority, facilitating decentralized, self-sovereign identity management, where agents can independently publish and query their identities.

For instance, Bao et al. [141] propose a robust privacy-preserving identity management system for industrial IoT, utilizing blockchain and cryptographic tools to ensure unforgeability, traceability, revocability, and verifiability. However, in the open and resource-limited SemComNet scenarios, the need for lightweight, privacy-preserving identity management is crucial, which is neglected in [141]. To tackle this, Xu et al. [142] introduce a blockchain-based system for mobile SemComNet, empowering users with self-sovereign identities (SSIs). Blockchain records legitimate user SSIs and public keys on the blockchain for decentralized authentication, while Chameleon hash ensures efficient revocation of unauthorized users to reduce overhead.

In the fast-paced SemComNet, cross-domain authentication is crucial for rapid identity verification across diverse agents. Chen et al. [144] propose a privacy-preserving cross-domain authentication solution for public key infrastructure. It ensures cross-domain compatibility and rapid response via multiple Merkle hash trees. For industrial networks, Shen et al. [143] employ a consortium blockchain and identity-based signatures for establishing trust among different domains. For authorization, smart contracts running on the blockchain enable automatic authorization policies in SemComNet. For instance, Wang et al. [132] apply smart contracts for fine-grained data access and traceable usage management, with off-chain execution to reduce computational overhead. However, real-world performance in large-scale SemComNet with heterogeneous components remains underexplored.

3) Trust Evaluation in SemComNet: Trust evaluation is vital for ensuring security and reliability during cooperation within large-scale SemComNet. Current methods predominantly focus on using Dempster-Shafer theory, fuzzy logic, and Bayes' theorem to analyze and combine trust-impact attributes (such as direct/indirect, subjective/objective, local/global, and historical/present) during evaluation [135]. For instance, Parhizkar et al. [174] propose a trust evaluation mechanism based on both direct trust (from historical transactions) and

TABLE X
SUMMARY OF KEY LITERATURE ON DATA & KNOWLEDGE SECURITY IN SEMANTIC COMMUNICATION NETWORKS

Ref.	Security Threat	⋆ Purpose • Advantages ∘ Limitations † Evaluation Metrics	Utilized Technology
[145]	Desynchronization of KBs, data tampering	⋆ Safeguard data/knowledge integrity and confidentiality during sharing • High agents' QoE, enhanced efficiency, and strengthened security ∘ Potential scalability and latency issues in large-scale deployments † Detection ratio of malicious full nodes and ratio of stakes of legitimate nodes	Coalition-matching game, blockchain
[140]	Unauthorized KBs access, poisoning attack	⋆ Flexible access control policies for big data/knowledge sharing and analysis • Fine-grained, purpose-aware, and low-performance overhead ∘ Complexity of configuration and difficulty in implementing cross-domain access † Set sum of squared error and query processing efficiency	Purpose-based AC
[176]	Privacy exposure from erased data	⋆ Endow trained data/knowledge with deletion capacity in efficient ways • Reduce storage and computational overhead without access to the training data ∘ Lack information-theoretic guarantee and adaptability to non-convex loss † Computational and storage complexity	DP, machine unlearning
[177]	Privacy exposure from raw data/KBs	⋆ Federated semantic codecs training with IB for balancing accuracy and privacy • High rate-distortion performance and convergence in harsh channel conditions ∘ Limited applicability to non-IID data and high communication overhead † Recall accuracy, latency, and error rate	FL, IB
[46]	Gradient leakage attack	⋆ Offer customized and controllable data utilization for agents • High personalization, adaptability, and customization ∘ Degradation of model accuracy † Peak SNR	DP

indirect trust (from social interactions). However, the above work [174] approach relies on subjective influence and simple attribute combinations, leading to practical challenges, such as cold start and sparse history data [175].

The integration of DL methods [175] addresses the challenges of trust evaluation and establishes an automated framework for SemComNet. For instance, Jayasinghe et al. [133] propose an ML-based trust assessment model for IoT services. Utilizing unsupervised learning techniques and support vector machines, this model accurately classifies trust features and calculates trust values. Besides, deep RL (DRL) which combines DL's perceptual capabilities and RL's decision-making abilities, may be well applied in trust or reputation models within SemComNet. In the context of misbehavior detection, Gyawali et al. [134] use DRL to dynamically update vehicle reputation based on Dempster-Shafer theory, improving system resilience against internal attacks. However, as the number of agents increases, communication and computation overheads may become prohibitive. The above works [133], [134], [174] mainly focus on entity-based trust models (i.e., relying on agents' credibility), while neglecting the authenticity of SI. To fill this gap, Truong et al. [136] use virtual interactions and quality assessments to calculate trust indicators (experience and reputation). However, acquiring these indicators may be challenging, and a typically sparse trust matrix may compromise the accuracy of trust assessments in SemComNet.

E. Data & Knowledge Security in SemComNet

SemComNet operates as a data-knowledge dual-driven paradigm. On the one hand, it relies on training semantic models with large-scale data to provide semantic-oriented transmission services. On the other hand, the accumulated knowledge in KBs improves the agents' abilities in understanding and reasoning, thereby facilitating the extraction and reconstruction of SI. Within this paradigm, information security becomes a crucial prerequisite for the development and prosperity of the SemComNet. Next, we delve into discussions regarding data & knowledge security in SemComNet, focusing on aspects such as anti-poisoning, tamper-proofing, access control, and privacy preservation.

1) Defense Strategies Against Semantic Data Poisoning: The semantic data poisoning attacker manipulates the training data used by semantic codecs to degrade their performance. Several effective defense strategies have been proposed, such as data preprocessing [89], poison filtering [90], and robustness training [88]. Data preprocessing techniques, such as data augmentation and transformation, can strengthen the model against such attacks. For instance, Borgnia et al. [89] demonstrate that strong data augmentations (e.g., Mixup and CutMix), significantly improve resistance to poisoning attacks without trading off performance. Apart from data preprocessing, the poison-filter-based defenses [88] focus on detecting and removing corrupted data before training semantic models. Methods such as outlier detection, ensemble-based filtering, validation-based filtering, and k-nearest neighbors (k-NN) filtering [88] can identify and eliminate anomalous data. For instance, Li et al. [90] propose a defense mechanism based on knowledge distillation to counteract data poisoning. By training an auxiliary model on a validation set and distilling its knowledge to create pseudo-labels, they sanitize the dataset and enhance the model's robustness. Additionally, robust training techniques [88] such as robust regression enhance the semantic models' resistance by optimizing the training process to mitigate the impact of malicious data.

2) Blockchain for Tamper-Proof SemComNet: The distributed nature and immutability of blockchain make the SemComNet under the premise of ensuring data availability [131] and sharing [145]. It provides SemComNet with integrity [147], confidentiality [146] of semantic data &

knowledge while maintaining it can be redactable [148]. Specifically, Liang et al. in [131] present a blockchain-based data provenance system named ProvChain to ensure semantic data source security from collection, storage, and verification in three stages. For secure decentralized knowledge sharing, Wang et al. [145] introduce a blockchain-based framework to reduce malicious activities effectively. That holds promise for supporting secure knowledge sharing in SemComNet. However, scalability and latency concerns in extensive SemComNet deployments require careful consideration.

To ensure remote knowledge integrity in cloud storage services, Wang et al. [147] introduce a blockchain-based private provable data possession scheme, to ensure remote data integrity in cloud storage. Compared with existing cryptography schemes, they offer enhanced security, efficiency, and practicality while ensuring agent anonymity. As for confidentiality of data and knowledge, Fotiou et al. [146] introduce a decentralized security approach for content distribution utilizing blockchain in a fully distributed manner, ensuring security without relying on central authorities. Lastly, blockchain technology plays a crucial role in ensuring trustworthy data deletion. For instance, Ateniese et al. [148] propose a framework for redacting and compressing the content of blocks in blockchain-based systems. Experiment results show the overhead imposed by having a mutable blockchain is negligible.

3) Data & Knowledge Access Control in SemComNet: Various access control (AC) policies, including role-based [137], attribute-based [138], [139], purpose-based [140], can be employed to safeguard knowledge and semantic data in SemComNet, tailored to specific requirements. For instance, Sultan et al. [137] introduce a cryptographic role-based encryption for AC, where only authorized users decrypt data. It includes efficient user revocation and outsourced decryption, reducing computational load. Theoretical analysis proves its security against chosen-plaintext attacks, making it ideal for practical SemComNet applications. For fine-grained AC in collaborative scenarios, Xue et al. [138] employs attribute-based encryption to facilitate collaborative access based on owner-defined policies, which may be applied to SemComNet to ensure the clustered and networked agents collaboration while preventing unauthorized collusion attempts.

However, the above role and attribute-based AC models usually fall short in allowing agents to perform statistical analysis on sensitive semantic data and knowledge without direct access. To tackle this challenge, on the one hand, Xu et al. [139] propose a privacy-preserving, revocable attribute-based AC using a linear secret sharing scheme and an extended path oblivious random access memory protocol. This ensures privacy and allows for fine-grained access control, such as write access and policy updates. On the other hand, the purpose-aware AC model allows knowledge owners to define intended usage purposes, ensuring privacy-preserving access control. For instance, Xue et al. [140] propose an automatic purpose-aware AC model that distinguishes data processing purposes and enforces access using two mechanisms for structured and unstructured data. While experiments show efficiency in large-scale platforms, the model's complexity and challenges in cross-domain access hinder its broader adoption in SemComNet. Further research is needed to develop context-aware AC that adapts permissions based on specific communication environments, enhancing SemComNet's adaptability and intelligence.

4) Machine Unlearning for Privacy Preservation in SemComNet: In SemComNet, agents may need to erase their communication history and private knowledge for privacy and security reasons. Traditional methods, such as retraining DNN-based codecs from scratch, are resource-intensive and disrupt communication services. The emerging concept of machine unlearning [178] offers an efficient solution for data removal and model adaptation. It provides a lightweight approach to safeguard agents' right to be forgotten while maintaining seamless communication, effectively balancing privacy and efficiency.

In [179], Golatkar et al. propose a method called scrubbing to selectively remove any information about a dataset from the network's weights without requiring access to the original data or retraining. While effective, its implementation in SemComNet is challenging due to its reliance on simple and well-structured models. To simplify unlearning, Sekhari et al. [176] present an approximate unlearning algorithm that reduces computational and storage complexities. By applying disturbance updates during training to subtract the influence of specific samples, an unlearned model was produced. However, the work [176] has practical limitations, including reliance on convex loss functions and the lack of an information-theoretic guarantee for unlearning.

The above centralized unlearning methods [176], [179] require access to all training data, while the federated unlearning, which does not, is more practical in the distributed training setting of SemComNet. It aims to erase a client's contributions from the trained models and can be categorized into server-side and client-side approaches [180]. Server-side unlearning is more efficient as it allows clients' contributions to be erased without their participation. For instance, Wu et al. [181] propose a server-based approach that subtracts a client's accumulated updates from the global model and uses knowledge distillation to restore its performance. In contrast, client-side federated unlearning involves clients erasing the influence of specific data samples. Liu et al. [182] propose a Newton-type retraining algorithm that removes the influence of specific samples, using the Fisher information matrix to minimize retraining costs. Their approach achieves efficient data erasure while maintaining model accuracy.

5) FL for Privacy Preservation in SemComNet: Centralized learning in SemComNet requires aggregating raw data from users, which leads to traffic congestion and privacy concerns [36], [46], [177]. FL offers a more privacy-preserving and efficient alternative by keeping user data on devices and reducing the training burden through shared efforts. Researchers have explored integrating FL with SemComNet for tasks such as training semantic codecs and constructing KBs [46], [177], [183]. For semantic codecs training, Zhao et al. [46] propose an online inference and offline FL framework that balances privacy with data utilization by combining privacy-protected model training and personalized deployment. For shared KB training, Wei et al. [177] introduce the federated semantic learning (FedSem) framework. This approach allows agents to train local KBs and transmit learned knowledge (as model parameters) to the server. The IB mechanism is also employed to limit the amount of shared knowledge, ensuring privacy protection and improving rate-distortion and convergence performance.

6) Potential Privacy Protection Approaches for SemComNet: Differential privacy (DP), another widely used technique in SemComNet, safeguards semantic data and knowledge by adding noise. For instance, Min et al. [184] apply the DP mechanism to randomize semantic locations, using RL to adaptively select the perturbation policies based on location sensitivity and attack history. Experimental results show effectiveness in balancing privacy and QoE loss. Additionally, to resist the gradient leakage attacks where attackers can infer private training data from shared semantic model parameters or gradients, Zhao et al. [46] employ edge-side model aggregation with DP. This approach allows agents to add random noise to parameter sharing based on their diverse privacy needs, ensuring personalized privacy protection. However, adding noise may markedly degrade the accuracy of semantic codecs, presenting a trade-off between model accuracy and privacy protection that needs careful consideration. In response, Liu et al. [185] propose APB-DP, an adaptive privacy budget-based DP scheme that balances model performance and privacy in collaborative training. Simulations show that APB-DP reduces the *privacy leakage rate* (PRL) by 13% and *performance loss rate* (PLoss) by 71% compared to the standard FL scheme. Additionally, techniques such as knowledge distillation [76], secure MPC [113], and trusted execution environment [186] also offer potential solutions safeguarding privacy in SemComNet.

F. Summary and Lessons Learned

This section has discussed advanced security/privacy countermeasures for SemComNet, encompassing semantic model security, transmission security, robustness, trust mechanisms, and data & knowledge security (i.e., from Sections IV-A–IV-E). The insights from the literature review highlight potential pathways to building a secure, privacy-preserving, robust, and trustworthy SemComNet. The summary and key lessons learned from this section are listed as follows.

- *Semantic model security*. Semantic models, foundational to SemComNet, leverage AI techniques for semantic understanding tasks. However, they face threats like AEs, poisoning examples, sponge examples, and model theft. In Section IV-A, we have explored key strategies to mitigate these risks. Specifically, adversarial training, defensive distillation, and AE detectors improve model robustness against AEs, while detection and adversarial training methods mitigate sponge examples. For poisoning threats, anomaly detection, data sanitization, and robust aggregation prove effective. Besides, watermarking and blockchain-based techniques are promising for IP protection. Table VI summarizes these defenses for semantic model-related threats.
- *Semantic transmission security*. The broadcast nature of communication channels allows any nearby user to intercept transmitted SI, creating opportunities for adversaries to launch attacks such as semantic eavesdropping and jamming [19]. For secure SI transmission in SemComNet, we have learned that existing cryptographic schemes, PLS techniques (including secret key generation and secure transmission approaches), covert communication, quantum technology, and other emerging communication technologies (e.g., RIS and spread spectrum techniques) can offer some insights for protecting SI transmission. Table VII compares existing/potential countermeasures for addressing semantic transmission threats.
- *Reliable SemComNet*. Reliability in SemComNet is essential to tolerate perturbations during transmission and maintain resilience in SI interpretation (e.g., reducing semantic noise). We have learned that AI techniques including ML, GAI, and RL could enhance the resilience of semantic interpretation while optimizing bandwidth usage. Besides, adaptive transmission strategies could offer some insights for enhancing reliability. Moreover, incorporating reasoning abilities and collaborative relays within SemComNet further enhances robustness by improving communication reliability and ensuring precise SI interpretation. Table VIII summarizes existing/potential countermeasures for addressing reliability risks in SemComNet.
- *Trust management in SemComNet*. To build a trustworthy SemComNet, trust solutions can be characterized as intrinsic, trust-free, and quantifiable approaches. Specifically, PLA mechanisms establish intrinsic trust by verifying agents' identities through inherent properties (e.g., RF fingerprints), forming the foundation for trust evaluation. Moreover, blockchain creates a trustless environment by eliminating reliance on central authorities, making it ideal for large-scale identity management. Besides, measurable trust evaluation mechanisms assess the reliability of agents or semantic content. Table IX summarizes existing and potential countermeasures for addressing trust management challenges in SemComNet.
- *Data & knowledge security in SemComNet*. For data & knowledge security in SemComNet, the distributed nature and expanded attack surface intensify existing threats. In response, blockchain emerges as a promising solution, ensuring availability, integrity, and confidentiality throughout their life cycle (e.g., collection, storage, sharing, and destruction). Moreover, various AC policies (e.g., role-based, attribute-based, and purpose-based) allow for tailored restrictions on agents' access to critical information. Lastly, for privacy protection, advanced approaches such as machine unlearning, FL, and DP offer effective solutions. However, further advancements tailored to SemComNet's unique characteristics are needed. A comparison of existing/potential defenses tailored for

data/knowledge security and privacy issues is presented in Table X.

V. Future Research Directions

Recent SemComNet research has shown significant advancements. However, several critical issues remain unexplored. This section explores key challenges in SemComNet research and outlines potential future directions.

A. Green SemComNet Architecture

In large-scale SemComNet, frequent knowledge/model sharing significantly increases resource and energy consumption. To address this, SemComNet should adopt green and eco-friendly designs to minimize resource waste and prolong agents' battery life. On the one hand, various communication technologies (e.g., ultra-massive MIMO and RIS) are increasingly deployed to meet the growing demand for high data throughput [12]. A potential research direction is ensuring seamless compatibility between SemComNet and existing communication infrastructure. It ensures a smooth transition to a more intelligent and green paradigm without wasting current infrastructure investments. On the other hand, the growing complexity of communication tasks imposes a significant burden on resource-constrained agents. For instance, the reliability of data-driven codecs requires rich high-quality datasets for training and sufficient background knowledge for support. It is necessary to explore eco-friendly strategies that efficiently construct self-learning [187] and self-updating datasets & KBs, reducing the maintenance cost. Additionally, from an economic perspective, exploring the utilization of low-power hardware, such as neuromorphic chips [188] and quantum photonic chips [189], offers a promising avenue for energy-efficient and secure task execution.

B. Explainable Semantic Model

The mathematical representation of semantics is inherently challenging, but the powerful semantic extraction capabilities of DL have made SemCom a reality [3], [28]. However, the opaque, closed-box nature of DL-based semantic codecs results in a lack of explainability, raising concerns about reliability and social acceptance, particularly in safety-sensitive domains such as transportation and healthcare. Explainable AI (XAI) offers a promising avenue by making decision-making processes transparent and traceable. In SemComNet, XAI can be applied across its three layers, such as improving resource allocation interpretation in the control layer, clarifying SI extraction and reconstruction in the transmission layer, and enhancing knowledge acquisition clarity in the cognitive sensing layer. Additionally, integrating domain-specific knowledge to enhance contextual explainability remains an open challenge. For instance, KGs offer structured and interpretable representations of knowledge [190], enabling contextual understanding and the incorporation of domain expertise. However, adapting KGs to SemComNet and ensuring their dynamic updates in response to evolving contexts pose critical challenges.

C. SemComNet Orchestrated With Generative AI

The integration of GAI models (e.g., DALL-E and GPT-4) with SemComNet holds transformative potential, enabling applications such as KBs creation and content refinement. These models can generate semantically consistent, context-aware content to enrich KBs and enhance semantic reasoning [51]. However, their incorporation into SemComNet faces challenges in trustworthiness, sustainability, and personalization, which require further research. Firstly, GAI models are prone to *hallucination* that generates outputs that appear plausible but are nonsensical or adversarial, thereby compromising service reliability [168]. Secondly, the resource-intensive and time-consuming nature of GAI models, both for training and inference, pose challenges for resource-constrained agents [51]. Techniques such as quantization and model compression can enhance efficiency in large-scale SemComNet deployments while maintaining performance [45]. Thirdly, while GAI models offer general-purpose solutions, adapting them to personalized scenarios through fine-tuning remains an open research area requiring further attention.

D. Endogenous Secure SemComNet

As security threats in SemComNet become more complex and sophisticated, traditional patch-like protection solutions are increasingly ineffective. There is an urgent need to develop a SemComNet framework with endogenous security [191], where security mechanisms are integrated by design into the system architecture from the outset. This proactive approach, which incorporates self-protection, self-evolution, rapid response, and autoimmunity capabilities, is aimed at adapting to dynamic environments and resisting both known and unknown threats. For instance, physical-layer steganography technology [110] can encrypt and conceal SI by exploiting the unique properties of the physical channel, providing quantum-resistant security for data transmission within SemComNet. However, developing a comprehensive and systematically integrated endogenous security framework for SemComNet remains a significant challenge. Moreover, given the collaborative nature of SemComNet, where multiple agents work together to achieve common objectives, further research is needed to design privacy-preserving and endogenous security mechanisms for collaborative decision-making and conflict resolution. Such mechanisms would ensure secure cooperation among agents, especially in knowledge sharing and SI transmission, while safeguarding confidentiality and integrity in multi-agent interactions.

E. Adaptive SemComNet Design

As a future direction, adaptive SemComNet should address two critical challenges to support flexible collaboration and diverse demands in practical scenarios. Firstly, achieving effective SemCom among multiple users in dynamic and complex networks requires overcoming challenges such as knowledge synchronization and adaptation to evolving network topologies. Synchronizing shared KBs is crucial for maintaining consistent semantic understanding, yet it becomes challenging in networked SemCom with changing topologies or intermittent connections [13]. Despite its significance, the implementation of adaptive networked SemCom remains underexplored and warrants further investigation. Secondly, efficiently handling multiple semantic tasks simultaneously is another pressing issue [34]. The SI extracted by different DL-based codecs for semantic tasks (e.g., behavior identification and anomaly detection) is often incompatible and non-interchangeable [15], leading to inefficiencies when facing diverse tasks. A promising solution lies in developing unified semantic models capable of supporting multiple tasks. Such models could leverage advanced techniques such as continual learning, meta-learning, and multi-task learning to dynamically adapt to various tasks without requiring extensive retraining or storing multiple models.

VI. Conclusion

In this work, a thorough survey of SemComNet on fundamental concepts, security/privacy concerns, and countermeasures aspects has been presented. Firstly, we have introduced a novel three-layered architecture of SemComNet, consisting of the control layer, semantic transmission layer, and cognitive sensing layer. Afterward, we have discussed three working modes (i.e., paired, clustered, and networked) of SemComNet, along with the supporting technologies, use cases, and evaluation metrics. Next, our survey has revealed critical security and privacy threats of SemComNet from the three functional layers, which have not been comprehensively investigated in existing research. Then, to build a secure and robust SemComNet, the security and privacy countermeasures have been reviewed and examined from both academic and industrial perspectives, and the key challenges have been discussed to build tailored defenses in SemComNet. Finally, we have outlined the future research directions for SemComNet. We hope this work serves as a valuable guideline for security and privacy in SemComNet and encourages further research in this emerging area.

References

[1] P. Zhang, H. Yang, Z. Feng, Y. Cui, J. Dai, and X. Qin, "Toward intelligent and efficient 6G networks: JCSC enabled on-purpose machine communications," *IEEE Wireless Commun.*, vol. 30, no. 1, pp. 150–157, Feb. 2023.

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

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

[4] 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.

[5] Y. Wang, W. Chen, T. H. Luan, Z. Su, Q. Xu, and R. Li, "Task offloading for post-disaster rescue in unmanned aerial vehicles networks," *IEEE/ACM Trans. Netw.*, vol. 30, no. 4, pp. 1525–1539, Aug. 2022.

[6] H. Zeng, Z. Su, Q. Xu, R. Li, Y. Wang, and M. Dai, "USV fleet-assisted collaborative computation offloading for smart maritime services: An energy-efficient design," *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 14718–14733, Oct. 2024.

[7] K. Huang, Q. Lan, Z. Liu, and L. Yang, "Semantic data sourcing for 6G edge intelligence," *IEEE Commun. Mag.*, vol. 61, no. 12, pp. 70–76, Dec. 2023.

[8] X. Peng et al., "A robust deep learning enabled semantic communication system for text," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2022, pp. 2704–2709.

[9] R. Zhao, Q. Qin, N. Xu, G. Nan, Q. Cui, and X. Tao, "SemKey: Boosting secret key generation for RIS-assisted semantic communication systems," in *Proc. IEEE Veh. Technol. Conf. (VTC-Fall)*, 2022, pp. 1–5.

[10] G. Nan, Z. Li, J. Zhai, Q. Cui, G. Chen, and X. Du, "Physical-layer adversarial robustness for deep learning-based semantic communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2592–2608, Aug. 2023.

[11] C. T. Blog. "Chinese engineers field test a 6G network with semantic communications on 4G infrastructure." 2024. Accessed: Sep. 22, 2024. [Online]. Available: https://techblog.comsoc.org/2024/07/15/chinese-engineers-field-test-a-6g-network-with-semantic-communications-on-4g-infrastructure/

[12] P. Zhang et al., "Toward wisdom-evolutionary and primitive-concise 6G: A new paradigm of semantic communication networks," *Engineering*, vol. 8, pp. 60–73, Jan. 2022.

[13] 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.

[14] Z. Qin, L. Liang, Z. Wang, S. Jin, X. Tao, and W. Tong, "AI empowered wireless communications: From bits to semantics," *Proc. IEEE*, vol. 112, no. 7, pp. 621–652, Jul. 2024.

[15] Z. Tian, H. Vo, C. Zhang, G. Min, and S. Yu, "An asynchronous multi-task semantic communication method," *IEEE Netw.*, vol. 38, no. 4, pp. 275–283, Jul. 2024.

[16] M. Shen, J. Wang, H. Du, D. Niyato, X. Tang, and J. Kang, "Secure semantic communications: Challenges, approaches, and opportunities," *IEEE Netw.*, vol. 38, no. 4, pp. 197–206, Jul. 2024.

[17] H. Peng, Z. Zhang, Y. Liu, Z. Su, T. H. Luan, and N. Cheng, "Semantic communication in non-terrestrial networks: A future-ready paradigm," *IEEE Netw.*, vol. 38, no. 4, pp. 119–127, Jul. 2024.

[18] I. Shumailov, Y. Zhao, D. Bates, N. Papernot, R. Mullins, and R. Anderson, "Sponge examples: Energy-latency attacks on neural networks," in *Proc. IEEE Eur. Symp. Security Privacy (EuroS&P)*, 2021, pp. 212–231.

[19] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, and M. Guizani, "Rethinking wireless communication security in semantic Internet of Things," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 36–43, Jun. 2023.

[20] A. Li, X. Wei, D. Wu, and L. Zhou, "Cross-modal semantic communications," *IEEE Wireless Commun.*, vol. 29, no. 6, pp. 1–8, Dec. 2022.

[21] Y. Wang, S. Guo, Y. Deng, H. Zhang, and Y. Fang, "Privacy-preserving task-oriented semantic communications against model inversion attacks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 10150–10165, Aug. 2024.

[22] J. Kang, J. He, H. Du, Z. Xiong, Z. Yang, and X. Huang, "Adversarial attacks and defenses for semantic communication in vehicular metaverses," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 48–55, Aug. 2023.

[23] Y. Wang, Z. Su, S. Guo, M. Dai, T. H. Luan, and Y. Liu, "A survey on digital twins: Architecture, enabling technologies, security and privacy, and future prospects," *IEEE Internet Things J.*, vol. 10, no. 17, pp. 14965–14987, Sep. 2023.

[24] S. Cheng, X. Zhang, Y. Sun, Q. Cui, and X. Tao, "Knowledge discrepancy oriented privacy preserving for semantic communication," *IEEE Trans. Veh. Technol.*, vol. 73, no. 8, pp. 11637–11646, Aug. 2024.

[25] C. Li, L. Zeng, X. Huang, X. Miao, and S. Wang, "Secure semantic communication model for black-box attack challenge under metaverse," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 56–62, Aug. 2023.

[26] T. Qiu, N. Chen, K. Li, M. Atiquzzaman, and W. Zhao, "How can heterogeneous Internet of Things build our future: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2011–2027, 3rd Quart., 2018.

[27] Q. Lan et al., "What is semantic communication? A view on conveying meaning in the era of machine intelligence," *J. Commun. Inf. Netw.*, vol. 6, no. 4, pp. 336–371, 2021.

[28] Z. Qin, X. Tao, J. Lu, W. Tong, and G. Y. Li, "Semantic communications: Principles and challenges," 2022, *arXiv:2201.01389*.

[29] E. Uysal, O. Kaya, A. Ephremides, J. Gross, M. Codreanu, and P. Popovski, "Semantic communications in networked systems: A data significance perspective," *IEEE Netw.*, vol. 36, no. 4, pp. 233–240, Jul./Aug. 2022.

[30] D. Gündüz, Z. Qin, I. E. Aguerri, H. S. Dhillon, Z. Yang, and A. Yener, "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5–41, Jan. 2023.

[31] W. Yang, H. Du, Z. Q. Liew, W. Y. B. Lim, Z. Xiong, and D. Niyato, "Semantic communications for future Internet: Fundamentals, applications, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 213–250, 1st Quart., 2023.

[32] Z. Lu, R. Li, K. Lu, X. Chen, E. Hossain, and Z. Zhao, "Semantics-empowered communications: A tutorial-cum-survey," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 1, pp. 41–79, 1st Quart., 2024.

[33] S. E. Trevlakis, N. Pappas, and A.-A. A. Boulogeorgos, "Toward natively intelligent semantic communications and networking," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 1486–1503, 2024.

[34] G. Zhang, Q. Hu, Z. Qin, Y. Cai, G. Yu, and X. Tao, "A unified multi-task semantic communication system for multimodal data," *IEEE Trans. Commun.*, vol. 72, no. 7, pp. 4101–4116, Jul. 2024.

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

[36] H. Xie, Z. Qin, X. Tao, and K. B. Letaief, "Task-oriented multi-user semantic communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2584–2597, Sep. 2022.

[37] 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.

[38] 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.

[39] 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)*, 2023, pp. 122–127.

[40] S. Wang, J. Dai, Z. Liang, K. Niu, Z. Si, and C. Dong, "Wireless deep video semantic transmission," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 214–229, Jan. 2023.

[41] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[42] M. Xu, H. Du, D. Niyato, J. Kang, Z. Xiong, and S. Mao, "Unleashing the power of edge-cloud generative AI in mobile networks: A survey of AIGC services," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 2, pp. 1127–1170, 2nd Quart., 2024.

[43] B. Min et al., "Recent advances in natural language processing via large pre-trained language models: A survey," *ACM Comput. Surveys*, vol. 56, no. 2, pp. 1–40, 2023.

[44] 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.

[45] Y. Pan, Z. Su, Y. Wang, S. Guo, H. Liu, and R. Li, "Cloud-edge collaborative large model services: Challenges and solutions," *IEEE Netw.*, early access, Aug. 13, 2024, doi: 10.1109/MNET.2024.3442880.

[46] L. Zhao, D. Wu, and L. Zhou, "Data utilization versus privacy protection in semantic communications," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 44–50, Jun. 2023.

[47] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, "A survey on machine learning for data fusion," *Inf. Fusion*, vol. 57, pp. 115–129, May 2020.

[48] X. Luo, R. Gao, H.-H. Chen, S. Chen, Q. Guo, and P. N. Suganthan, "Multi-modal and multi-user semantic communications for channel-level information fusion," *IEEE Wireless Commun.*, vol. 31, no. 2, pp. 117–125, Apr. 2024.

[49] 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.

[50] J. Xu, B. Ai, W. Chen, N. Wang, and M. Rodrigues, "Deep joint source-channel coding for image transmission with visual protection," *IEEE Trans. Cogn. Commun. Netw.*, vol. 9, no. 6, pp. 1399–1411, Dec. 2023.

[51] L. Xia, Y. Sun, C. Liang, L. Zhang, M. A. Imran, and D. Niyato, "Generative AI for semantic communication: Architecture, challenges, and outlook," 2023, *arXiv:2308.15483*.

[52] Q. Wu, G. Ding, Y. Xu, S. Feng, Z. Du, and J. Wang, "Cognitive Internet of Things: A new paradigm beyond connection," *IEEE Internet Things J.*, vol. 1, no. 2, pp. 129–143, Apr. 2014.

[53] H. Seo, J. Park, M. Bennis, and M. Debbah, "Semantics-native communication via contextual reasoning," *IEEE Trans. Cogn. Commun. Netw.*, vol. 9, no. 3, pp. 604–617, Jun. 2023.

[54] Q. Zhao, G. Li, J. Cai, M. Zhou, and L. Feng, "A tutorial on Internet of Behaviors: Concept, architecture, technology, applications, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 1227–1260, 2nd Quart., 2023.

[55] D. D. Bourgin, J. C. Peterson, D. Reichman, S. J. Russell, and T. L. Griffiths, "Cognitive model priors for predicting human decisions," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2019, pp. 5133–5141.

[56] Q. Wu, T. Ruan, F. Zhou, Y. Huang, F. Xu, and S. Zhao, "A unified cognitive learning framework for adapting to dynamic environments and tasks," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 208–216, Dec. 2021.

[57] N. Rabinowitz, F. Perbet, F. Song, C. Zhang, S. M. A. Eslami, and M. Botvinick, "Machine theory of mind," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2018, pp. 4218–4227.

[58] Y. Wang, Z. Su, Y. Pan, T. H. Luan, R. Li, and S. Yu, "Social-aware clustered federated learning with customized privacy preservation," *IEEE/ACM Trans. Netw.*, vol. 32, no. 5, pp. 3654–3668, Oct. 2024.

[59] J. Guo, H. Chen, B. Song, Y. Chi, C. Yuen, and F. R. Yu, "Distributed task-oriented communication networks with multimodal semantic relay and edge intelligence," *IEEE Commun. Mag.*, vol. 62, no. 6, pp. 82–89, Jun. 2024.

[60] 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, 2024.

[61] X. Luo, B. Yin, Z. Chen, B. Xia, and J. Wang, "Autoencoder-based semantic communication systems with relay channels," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC WKSHPS)*, 2022, pp. 711–716.

[62] W. Lin, Y. Yan, L. Li, Z. Han, and T. Matsumoto, "Semantic-forward relaying: A novel framework toward 6G cooperative communications," *IEEE Commun. Lett.*, vol. 28, no. 3, pp. 518–522, Mar. 2024.

[63] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Comput. Commun.*, vol. 30, no. 14, pp. 2826–2841, 2007.

[64] M. De Lange et al., "A continual learning survey: Defying forgetting in classification tasks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3366–3385, Jul. 2022.

[65] Z. Yin, N. Cheng, Y. Hui, W. Wang, L. Zhao, and K. Aldubaikhy, "Multi-domain resource multiplexing based secure transmission for satellite-assisted IoT: AO-SCA approach," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 7319–7330, Nov. 2023.

[66] L. Leonardi, L. L. Bello, and G. Patti, "LoRa support for long-range real-time inter-cluster communications over Bluetooth low energy industrial networks," *Comput. Commun.*, vol. 192, pp. 57–65, Aug. 2022.

[67] B. Wang, R. Li, J. Zhu, Z. Zhao, and H. Zhang, "Knowledge enhanced semantic communication receiver," *IEEE Commun. Lett.*, vol. 27, no. 7, pp. 1794–1798, Jul. 2023.

[68] J. Qiu, Z. Tian, C. Du, Q. Zuo, S. Su, and B. Fang, "A survey on access control in the age of Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 4682–4696, Jun. 2020.

[69] W. Zhang, Y. Wang, M. Chen, T. Luo, and D. Niyato, "Optimization of image transmission in cooperative semantic communication networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 861–873, Feb. 2024.

[70] P. Zhang, W. Xu, Y. Liu, X. Qin, K. Niu, and S. Cui, "Intellicise wireless networks from semantic communications: A survey, research issues, and challenges," *IEEE Commun. Surveys Tuts.*, early access, Aug. 13, 2024, doi: 10.1109/COMST.2024.3443193.

[71] Y. Feng, H. Shen, Z. Shan, Q. Yang, and X. Shi, "Semantic communication for edge intelligence enabled autonomous driving system," *IEEE Netw.*, early access, Sep. 26, 2024, doi: 10.1109/MNET.2024.3468328.

[72] I. F. Akyildiz and H. Guo, "Holographic-type communication: A new challenge for the next decade," *ITU J. Future Evol. Technol.*, vol. 3, no. 2, pp. 421–442, 2022.

[73] R. Cheng, K. Liu, N. Wu, and B. Han, "Enriching telepresence with semantic-driven holographic communication," in *Proc. 22nd ACM Workshop Hot Topics Netw.*, 2023, pp. 147–156.

[74] Z. Wang, D. Liu, Y. Sun, X. Pang, P. Sun, and F. Lin, "A survey on IoT-enabled home automation systems: Attacks and defenses," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 4, pp. 2292–2328, 4th Quart., 2022.

[75] B.-Y. Hsu, C.-Y. Shen, H. S. Yuan, W.-C. Lee, and D.-N. Yang, "Social-aware group display configuration in VR conference," in *Proc. AAAI Conf. Artif. Intell.*, vol. 38, 2024, pp. 8517–8525.

[76] J. Chen, J. Wang, C. Jiang, Y. Ren, and L. Hanzo, "Trustworthy semantic communications for the metaverse relying on federated learning," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 18–25, Aug. 2023.

[77] P. Popovski et al., "A perspective on time toward wireless 6G," *Proc. IEEE*, vol. 110, no. 8, pp. 1116–1146, Aug. 2022.

[78] S. Kadam and D. I. Kim, "Semantic communication-empowered vehicle count prediction for traffic management," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2024, pp. 1–6, doi: 10.1109/WCNC57260.2024.10571211.

[79] H. Zhang, H. Wang, Y. Li, K. Long, and V. C. M. Leung, "Toward intelligent resource allocation on task-oriented semantic communication," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 70–77, Jun. 2023.

[80] M. Mitev, A. Chorti, H. V. Poor, and G. P. Fettweis, "What physical layer security can do for 6G security," *IEEE Open J. Veh. Technol.*, vol. 4, pp. 375–388, 2023.

[81] Y. Yao, J. Duan, K. Xu, Y. Cai, Z. Sun, and Y. Zhang, "A survey on large language model (LLM) security and privacy: The good, the bad, and the ugly," *High-Confidence Comput.*, vol. 4, no. 2, 2024, Art. no. 100211.

[82] Z. Tian, L. Cui, J. Liang, and S. Yu, "A comprehensive survey on poisoning attacks and countermeasures in machine learning," *ACM Comput. Surveys*, vol. 55, no. 8, pp. 1–35, 2022.

[83] J. Zhang et al., "Protecting intellectual property of deep neural networks with watermarking," in *Proc. Asia Conf. Comput. Commun. Security (AsiaCCS)*, 2018, pp. 159–172.

[84] H. Nie, S. Lu, J. Wu, and J. Zhu, "Deep model intellectual property protection with compression-resistant model watermarking," *IEEE Trans. Artif. Intell.*, vol. 5, no. 7, pp. 3362–3373, Jul. 2024.

[85] Z. Xi et al., "On the security risks of knowledge graph reasoning," in *Proc. 32nd USENIX Security Symp. (USENIX Security)*, 2023, pp. 3259–3276.

[86] 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.

[87] Z. Wang, H. Guo, Z. Zhang, W. Liu, Z. Qin, and K. Ren, "Feature importance-aware transferable adversarial attacks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 7619–7628.

[88] Z. Wang, J. Ma, X. Wang, J. Hu, Z. Qin, and K. Ren, "Threats to training: A survey of poisoning attacks and defenses on machine learning systems," *ACM Comput. Surveys*, vol. 55, no. 7, pp. 1–36, 2022.

[89] E. Borgnia et al., "Strong data augmentation sanitizes poisoning and backdoor attacks without an accuracy tradeoff," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2021, pp. 3855–3859.

[90] Y. Li, J. Yang, Y. Song, L. Cao, J. Luo, and L.-J. Li, "Learning from noisy labels with distillation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 1928–1936.

[91] B. Wang et al., "Neural cleanse: Identifying and mitigating backdoor attacks in neural networks," in *Proc. IEEE Symp. Security Privacy (SP)*, 2019, pp. 707–723.

[92] P. Blanchard, E. M. E. Mhamdi, R. Guerraoui, and J. Stainer, "Machine learning with adversaries: Byzantine tolerant gradient descent," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 30, 2017, pp. 119–129.

[93] L. Zhao, S. Hu, Q. Wang, J. Jiang, C. Shen, and X. Luo, "Shielding collaborative learning: Mitigating poisoning attacks through client-side detection," *IEEE Trans. Depend. Secure Comput.*, vol. 18, no. 5, pp. 2029–2041, Sep./Oct. 2021.

[94] J. Zhang and C. Li, "Adversarial examples: Opportunities and challenges," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 7, pp. 2578–2593, Jul. 2020.

[95] Q. Hu, G. Zhang, Z. Qin, Y. Cai, G. Yu, and G. Y. Li, "Robust semantic communications with masked VQ-VAE enabled codebook," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 8707–8722, Dec. 2023.

[96] Y. Liu, H.-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, 1st Quart., 2017.

[97] C. Zhao, H. Du, D. Niyato, J. Kang, Z. Xiong, and D. I. Kim, "Generative AI for secure physical layer communications: A survey," *IEEE Trans. Cogn. Commun. Netw.*, early access, Aug. 5, 2024, doi: 10.1109/TCCN.2024.3438379.

[98] Z. Yang, M. Chen, G. Li, Y. Yang, and Z. Zhang, "Secure semantic communications: Fundamentals and challenges," *IEEE Netw.*, vol. 38, no. 6, pp. 513–520, Nov. 2024.

[99] R. Tang, D. Gao, M. Yang, T. Guo, H. Wu, and G. Shi, "GAN-inspired intelligent jamming and anti-jamming strategy for semantic communication systems," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC WKSHPS)*, 2023, pp. 1623–1628.

[100] K. Li, B. P. L. Lau, X. Yuan, W. Ni, M. Guizani, and C. Yuen, "Toward ubiquitous semantic metaverse: Challenges, approaches, and opportunities," *IEEE Internet Things J.*, vol. 10, no. 24, pp. 21855–21872, Dec. 2023.

[101] X. Chen, J. An, Z. Xiong, C. Xing, N. Zhao, and F. R. Yu, "Covert communications: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 1173–1198, 2nd Quart., 2023.

[102] J. Hu, L. Ye, Y. Chen, X. Zhang, J. Wang, and Z. Chen, "Covert communications for text semantic with finite blocklength," *IEEE Wireless Commun. Lett.*, vol. 13, no. 10, pp. 2842–2846, Oct. 2024.

[103] Y. Chen, Q. Yang, Z. Shi, and J. Chen, "The model inversion eavesdropping attack in semantic communication systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2023, pp. 5171–5177.

[104] H. Xie, Z. Qin, X. Tao, and Z. Han, "Towards intelligent communications: Large model empowered semantic communications," 2024, *arXiv:2402.13073*.

[105] L. Zhu, Z. Liu, and S. Han, "Deep leakage from gradients," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2019, pp. 14747–14756.

[106] M. G. Gaber, M. Ahmed, and H. Janicke, "Malware detection with artificial intelligence: A systematic literature review," *ACM Comput. Surveys*, vol. 56, no. 6, pp. 1–33, 2024.

[107] Q. Zhou, R. Li, Z. Zhao, Y. Xiao, and H. Zhang, "Adaptive bit rate control in semantic communication with incremental knowledge-based HARQ," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1076–1089, 2022.

[108] J. Zhang et al., "Deep model intellectual property protection via deep watermarking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 8, pp. 4005–4020, Aug. 2022.

[109] Y. Wang, Y. Hu, H. Du, T. Luo, and D. Niyato, "Multi-agent reinforcement learning for covert semantic communications over wireless networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2023, pp. 1–5.

[110] N. Xie, J. Zhang, and Q. Zhang, "Security provided by the physical layer in wireless communications," *IEEE Netw.*, vol. 37, no. 5, pp. 42–48, Sep. 2023.

[111] R. Yamaguchi, H. Ochiai, and J. Shikata, "A physical-layer security based on wireless steganography through OFDM and DFT-precoded OFDM signals," in *Proc. IEEE Veh. Technol. Conf. (VTC-Spring)*, 2020, pp. 1–5.

[112] Q. Qin et al., "Securing semantic communications with physical-layer semantic encryption and obfuscation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2023, pp. 5608–5613.

[113] B. Knott, S. Venkataraman, A. Hannun, S. Sengupta, M. Ibrahim, and L. van der Maaten, "CrypTen: Secure multi-party computation meets machine learning," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 34, 2021, pp. 4961–4973.

[114] X. Luo, Z. Chen, M. Tao, and F. Yang, "Encrypted semantic communication using adversarial training for privacy preserving," *IEEE Commun. Lett.*, vol. 27, no. 6, pp. 1486–1490, Jun. 2023.

[115] L. Hu, S. Tan, H. Wen, J. Wu, J. Fan, and S. Chen, "Interference alignment for physical layer security in multi-user networks with passive eavesdroppers," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 3692–3705, 2023.

[116] J. Lin, Q. Li, J. Yang, H. Shao, and W.-Q. Wang, "Physical-layer security for proximal legitimate user and eavesdropper: A frequency diverse array Beamforming approach," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 3, pp. 671–684, Mar. 2018.

[117] Z. Li, J. Li, Y. Liu, X. Liang, K. G. Shin, and Z. Yan, "Exploiting interactions of multiple interference for cooperative interference alignment," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7072–7085, Nov. 2021.

[118] Y. Wang, W. Yang, P. Guan, Y. Zhao, and Z. Xiong, "STAR-RIS-assisted privacy protection in semantic communication system," *IEEE Trans. Veh. Technol.*, vol. 73, no. 9, pp. 13915–13920, Sep. 2024.

[119] R. Kaewpuang, M. Xu, W. Y. B. Lim, D. Niyato, H. Yu, and J. Kang, "Cooperative resource management in quantum key distribution (QKD) networks for semantic communication," *IEEE Internet Things J.*, vol. 11, no. 3, pp. 4454–4469, Feb. 2024.

[120] U. Khalid, M. S. Ulum, A. Farooq, T. Q. Duong, O. A. Dobre, and H. Shin, "Quantum semantic communications for metaverse: Principles and challenges," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 26–36, Aug. 2023.

[121] E. Wong, L. Rice, and J. Z. Kolter, "Fast is better than free: Revisiting adversarial training," 2020, *arXiv:2001.03994*.

[122] S. Chen, N. Carlini, and D. Wagner, "Stateful detection of black-box adversarial attacks," in *Proc. 1st ACM Workshop Security Privacy Artif. Intell.*, 2020, pp. 30–39.

[123] E. Erdemir, T.-Y. Tung, P. L. Dragotti, and D. Gündüz, "Generative joint source-channel coding for semantic image transmission," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2645–2657, Aug. 2023.

[124] J. Chen, D. You, D. Gündüz, and P. L. Dragotti, "CommIN: Semantic image communications as an inverse problem with INN-guided diffusion models," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2024, pp. 6675–6679.

[125] S. Jiang et al., "Reliable semantic communication system enabled by knowledge graph," *Entropy*, vol. 24, no. 6, p. 846, 2022.

[126] P. Yao and D. Barbosa, "Typing errors in factual knowledge graphs: Severity and possible ways out," in *Proc. Web Conf. (WWW)*, 2021, pp. 3305–3313.

[127] C. K. Thomas and W. Saad, "Neuro-symbolic causal reasoning meets signaling game for emergent semantic communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 4546–4563, May 2024.

[128] N. Gao, Q. Huang, C. Li, S. Jin, and M. Matthaiou, "EsaNet: Environment semantics enabled physical layer authentication," *IEEE Wireless Commun. Lett.*, vol. 13, no. 1, pp. 178–182, Jan. 2024.

[129] E. Jorswieck, S. Tomasin, and A. Sezgin, "Broadcasting into the uncertainty: Authentication and confidentiality by physical-layer processing," *Proc. IEEE*, vol. 103, no. 10, pp. 1702–1724, Oct. 2015.

[130] H. Tan, N. Xie, and A. X. Liu, "An optimization framework for active physical-layer authentication," *IEEE Trans. Mobile Comput.*, vol. 23, no. 1, pp. 164–179, Jan. 2024.

[131] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, "ProvChain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in *Proc. 17th IEEE/ACM Int. Symp. Cluster Cloud Grid Comput. (CCGRID)*, 2017, pp. 468–477.

[132] Y. Wang, Z. Su, N. Zhang, J. Chen, X. Sun, and Z. Ye, "SPDS: A secure and auditable private data sharing scheme for smart grid based on blockchain," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7688–7699, Nov. 2021.

[133] U. Jayasinghe, G. M. Lee, T.-W. Um, and Q. Shi, "Machine learning based trust computational model for IoT services," *IEEE Trans. Sustain. Comput.*, vol. 4, no. 1, pp. 39–52, Jan.–Mar. 2019.

[134] S. Gyawali, Y. Qian, and R. Q. Hu, "Deep reinforcement learning based dynamic reputation policy in 5G based vehicular communication networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6136–6146, Jun. 2021.

[135] J. Wang, X. Jing, Z. Yan, Y. Fu, W. Pedrycz, and L. T. Yang, "A survey on trust evaluation based on machine learning," *ACM Comput. Surveys*, vol. 53, no. 5, pp. 1–36, 2021.

[136] N. B. Truong, G. M. Lee, T.-W. Um, and M. Mackay, "Trust evaluation mechanism for user recruitment in mobile crowd-sensing in the Internet of Things," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 10, pp. 2705–2719, Oct. 2019.

[137] N. H. Sultan, V. Varadharajan, L. Zhou, and F. A. Barbhuiya, "A role-based encryption (RBE) scheme for securing outsourced cloud data in a multi-organization context," *IEEE Trans. Services Comput.*, vol. 16, no. 3, pp. 1647–1661, May/Jun. 2023.

[138] Y. Xue, K. Xue, N. Gai, J. Hong, D. S. L. Wei, and P. Hong, "An attribute-based controlled collaborative access control scheme for public cloud storage," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 11, pp. 2927–2942, Nov. 2019.

[139] R. Xu, J. Joshi, and P. Krishnamurthy, "An integrated privacy preserving attribute-based access control framework supporting secure deduplication," *IEEE Trans. Depend. Secure Comput.*, vol. 18, no. 2, pp. 706–721, Mar./Apr. 2021.

[140] T. Xue, Y. Wen, B. Luo, G. Li, Y. Li, and B. Zhang, "SparkAC: Fine-grained access control in spark for secure data sharing and analytics," *IEEE Trans. Depend. Secure Comput.*, vol. 20, no. 2, pp. 1104–1123, Mar./Apr. 2023.

[141] Z. Bao, D. He, M. K. Khan, M. Luo, and Q. Xie, "PBidm: Privacy-preserving blockchain-based identity management system for Industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 1524–1534, Feb. 2023.

[142] J. Xu, K. Xue, H. Tian, J. Hong, D. S. L. Wei, and P. Hong, "An identity management and authentication scheme based on redactable blockchain for mobile networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6688–6698, Jun. 2020.

[143] M. Shen, H. Liu, L. Zhu, K. Xu, H. Yu, and X. Du, "Blockchain-assisted secure device authentication for cross-domain Industrial IoT," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 5, pp. 942–954, May 2020.

[144] J. Chen, Z. Zhan, K. He, R. Du, D. Wang, and F. Liu, "XAuth: Efficient privacy-preserving cross-domain authentication," *IEEE Trans. Depend. Secure Comput.*, vol. 19, no. 5, pp. 3301–3311, Sep./Oct. 2022.

[145] Y. Wang, Z. Su, J. Li, N. Zhang, K. Zhang, and K.-K. R. Choo, "Blockchain-based secure and cooperative private charging pile sharing services for vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 1857–1874, Feb. 2022.

[146] N. Fotiou and G. C. Polyzos, "Decentralized name-based security for content distribution using blockchains," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2016, pp. 415–420.

[147] H. Wang, Q. Wang, and D. He, "Blockchain-based private provable data possession," *IEEE Trans. Depend. Secure Comput.*, vol. 18, no. 5, pp. 2379–2389, Sep./Oct. 2021.

[148] G. Ateniese, B. Magri, D. Venturi, and E. Andrade, "Redactable blockchain or rewriting history in bitcoin and friends," in *Proc. IEEE Eur. Symp. Security Privacy (EuroS & P)*, 2017, pp. 111–126.

[149] A. Bahramali, M. Nasr, A. Houmansadr, D. Goeckel, and D. Towsley, "Robust adversarial attacks against DNN-based wireless communication systems," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS)*, 2021, pp. 126–140.

[150] R. Xia, D. Liu, J. Li, L. Yuan, N. Wang, and X. Gao, "MMNet: Multi-collaboration and multi-supervision network for sequential deepfake detection," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 3409–3422, 2024.

[151] H. Oz, A. Aris, A. Levi, and A. S. Uluagac, "A survey on ransomware: Evolution, taxonomy, and defense solutions," *ACM Comput. Surveys*, vol. 54, no. 11s, pp. 1–37, 2022.

[152] T. McIntosh, A. Kayes, Y.-P. P. Chen, A. Ng, and P. Watters, "Ransomware mitigation in the modern era: A comprehensive review, research challenges, and future directions," *ACM Comput. Surveys*, vol. 54, no. 9, pp. 1–36, 2021.

[153] F. Wang, Y. Hong, and X. Ban, "Infrastructure-enabled GPS spoofing detection and correction," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 9, pp. 9462–9475, Dec. 2023.

[154] H. Du, J. Wang, D. Niyato, J. Kang, Z. Xiong, and J. Zhang, "Semantic communications for wireless sensing: RIS-aided encoding and self-supervised decoding," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2547–2562, Aug. 2023.

[155] L. Sun, Y. Yang, M. Chen, and C. Guo, "Disentangled information bottleneck guided privacy-protective joint source and channel coding for image transmission," *IEEE Trans. Commun.*, vol. 72, no. 11, pp. 6992–7005, Nov. 2024, doi: 10.1109/TCOMM.2024.3406381.

[156] G. Chen, G. Nan, Z. Jiang, H. Du, R. Shi, and Q. Cui, "Lightweight and robust wireless semantic communications," *IEEE Commun. Lett.*, vol. 28, no. 11, pp. 2633–2637, Nov. 2024.

[157] E. Bourtsoulatze, D. B. Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2019, pp. 4774–4778.

[158] T. Marchioro, N. Laurenti, and D. Gündüz, "Adversarial networks for secure wireless communications," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2020, pp. 8748–8752.

[159] Z. Li, W. Chen, Q. Wu, H. Cao, K. Wang, and J. Li, "Robust beamforming design and time allocation for IRS-assisted wireless powered communication networks," *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2838–2852, Apr. 2022.

[160] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[161] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[162] R. Xu, G. Li, Z. Yang, M. Chen, Y. Liu, and J. Li, "Covert and reliable semantic communication against cross-layer privacy inference over wireless edge networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2024, pp. 1–6, doi: 10.1109/WCNC57260.2024.10570800.

[163] M. Chehimi, C. Chaccour, C. K. Thomas, and W. Saad, "Quantum semantic communications for resource-efficient quantum networking," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 803–807, Apr. 2024.

[164] X. Li, Y. Pei, X. Yue, Y. Liu, and Z. Ding, "Secure communication of active RIS assisted NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 4489–4503, May 2024.

[165] H. Du, G. Liu, D. Niyato, J. Zhang, J. Kang, and Z. Xiong, "Generative AL-aided joint training-free secure semantic communications via multi-modal prompts," in *Proc. IEEE ICASSP*, 2024, pp. 12896–12900.

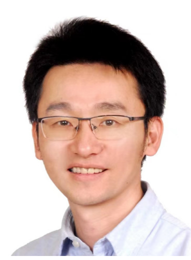
[166] H. Zeng, Z. Su, Q. Xu, and D. Fang, "USV fleet-assisted collaborative data backup in marine Internet of Things," *IEEE Internet Things J.*, vol. 11, no. 22, pp. 36308–36321, Nov. 2024.

[167] T. Han, J. Tang, Q. Yang, Y. Duan, Z. Zhang, and Z. Shi, "Generative model based highly efficient semantic communication approach for image transmission," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2023, pp. 1–5.
[168] S. Tonmoy et al., "A comprehensive survey of hallucination mitigation techniques in large language models," 2024, *arXiv:2401.01313*.
[169] W. Wang, Y. Yang, and F. Wu, "Towards data-and knowledge-driven AI: A survey on neuro-symbolic computing," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Oct. 17, 2024, doi: 10.1109/TPAMI.2024.3483273.
[170] Y. Wang, Y. Pan, M. Yan, Z. Su, and T. H. Luan, "A survey on ChatGPT: AI-generated contents, challenges, and solutions," *IEEE Open J. Comput. Soc.*, vol. 4, pp. 280–302, 2023.
[171] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 282–310, 1st Quart., 2021.
[172] J. B. Perazzone, P. L. Yu, B. M. Sadler, and R. S. Blum, "Artificial noise-aided MIMO physical layer authentication with imperfect CSI," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2173–2185, 2021.
[173] J. Xie et al., "A survey of blockchain technology applied to smart cities: Research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2794–2830, 3rd Quart., 2019.
[174] E. Parhizkar, M. H. Nikravan, R. C. Holte, and S. Zilles, "Combining direct trust and indirect trust in multi-agent systems," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2020, pp. 311–317.
[175] J. Wang, Z. Yan, H. Wang, T. Li, and W. Pedrycz, "A survey on trust models in heterogeneous networks," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 4, pp. 2127–2162, 4th Quart., 2022.
[176] A. Sekhari, J. Acharya, G. Kamath, and A. T. Suresh, "Remember what you want to forget: Algorithms for machine unlearning," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 34, 2021, pp. 18075–18086.
[177] H. Wei, W. Ni, W. Xu, F. Wang, D. Niyato, and P. Zhang, "Federated semantic learning driven by information bottleneck for task-oriented communications," *IEEE Commun. Lett.*, vol. 27, no. 10, pp. 2652–2656, Oct. 2023.
[178] H. Xu, T. Zhu, L. Zhang, W. Zhou, and P. S. Yu, "Machine unlearning: A survey," *ACM Comput. Surveys*, vol. 56, no. 1, pp. 1–36, 2024.
[179] A. Golatkar, A. Achille, and S. Soatto, "Eternal sunshine of the spotless net: Selective forgetting in deep networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, pp. 9301–9309.
[180] W. Wang, Z. Tian, C. Zhang, and S. Yu, "Machine unlearning: A comprehensive survey," 2024, *arXiv:2405.07406*.
[181] C. Wu, S. Zhu, and P. Mitra, "Federated unlearning with knowledge distillation," 2022, *arXiv:2201.09441*.
[182] Y. Liu, L. Xu, X. Yuan, C. Wang, and B. Li, "The right to be forgotten in federated learning: An efficient realization with rapid retraining," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2022, pp. 1749–1758.
[183] X. Lu, K. Zhu, J. Li, and Y. Zhang, "Efficient knowledge base Synchronization in semantic communication network: A federated distillation approach," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2024, pp. 1–6.
[184] M. Min, W. Wang, L. Xiao, Y. Xiao, and Z. Han, "Reinforcement learning-based sensitive semantic location privacy protection for VANETs," *China Commun.*, vol. 18, no. 6, pp. 244–260, 2021.
[185] H. Liu, B. Wang, R. Meng, S. Han, and X. Xu, "Adaptive privacy budget-based differential privacy co-training for wireless semantic communication," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2024, pp. 1–6.
[186] P. Jauernig, A.-R. Sadeghi, and E. Stapf, "Trusted execution environments: Properties, applications, and challenges," *IEEE Security Privacy*, vol. 18, no. 2, pp. 56–60, Mar./Apr. 2020.
[187] 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.
[188] M. Davies, A. Wild, G. Orchard, Y. Sandamirskaya, G. A. F. Guerra, and P. Joshi, "Advancing neuromorphic computing with Loihi: A survey of results and outlook," *Proc. IEEE*, vol. 109, no. 5, pp. 911–934, May 2021.
[189] W. Luo et al., "Recent progress in quantum photonic chips for quantum communication and Internet," *Light Sci. Appl.*, vol. 12, no. 1, p. 175, 2023.
[190] W. Yang, X. Chi, L. Zhao, Z. Xiong, and W. Jiang, "Task-driven semantic-aware green cooperative transmission strategy for vehicular networks," *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 5783–5798, Oct. 2023.
[191] J. Wu, "Development paradigms of cyberspace endogenous safety and security," *Sci. China Inf. Sci.*, vol. 65, no. 5, 2022, Art. no. 156301.

Shaolong Guo is currently pursuing the Ph.D. degree with the School of Cyber Science and Engineering, Xi'an Jiaotong University, Xi'an, China. His research interests include security and privacy protection in deep learning, semantic communications, and network games.

Yuntao Wang (Member, IEEE) received the Ph.D. degree in cyberspace security from Xi'an Jiaotong University, Xi'an, China, in 2022, where he is currently an Assistant Professor with the School of Cyber Science and Engineering. His research interests include security and privacy in intelligent IoT, network games, and blockchain.

Ning Zhang (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Waterloo in 2015. He is currently an Associate Professor and the Canada Research Chair with the Department of Electrical and Computer Engineering, University of Windsor. His research interests include connected vehicles, wireless networking, and security. He serves/served as an Associate Editor for IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, IEEE INTERNET OF THINGS JOURNAL, and IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING. He also serves/served as the TPC/General/Symposium Chair for numerous conferences and workshops, such as IEEE ICC, GLOBECOM, VTC, INFOCOM Workshop, and MOBICOM Workshops. He is also a Distinguished Lecturer of IEEE ComSoc, a Highly Cited Researcher (Web of Science), and the Vice Chair of IEEE Technical Committee on Cognitive Networks and IEEE Technical Committee on Big Data.



Zhou Su (Senior Member, IEEE) has published technical papers, including top journals and top conferences, such as IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE/ACM TRANSACTIONS ON NETWORKING, and INFOCOM. He received the Best Paper Award of the International Conference IEEE ICC2020, IEEE BigdataSE2019, and IEEE CyberSciTech2017. He is an Associate Editor of IEEE INTERNET OF THINGS JOURNAL, IEEE OPEN JOURNAL OF THE COMPUTER SOCIETY, and *IET Communications*.

Tom H. Luan (Senior Member, IEEE) received the B.E. degree in electrical and computer engineering from Xi'an Jiaotong University, China, the master's degree in electrical and computer engineering from the Hong Kong University of Science and Technology, Hong Kong, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada. From 2013 to 2017, he was a Lecturer of Mobile and Apps with Deakin University, Australia. From 2017 to 2022, he was a Professor with the School of Cyber Engineering, Xidian University. Since 2022, he has been with the School of Cyber Science and Engineering, Xi'an Jiaotong University, China, as a Professor. He has published over 210 peer reviewed papers in journal and conferences, including IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and INFOCOM. His research mainly focuses on the LLM, AI-agent, metaverse, and protocol and security design and performance evaluation of edge computing and digital twin network. He won the 2017 IEEE VTS Best Land Transportation Best Paper Award and the IEEE VTC 2023 Best Paper Award.

Zhiyi Tian (Member, IEEE) received the B.S. degree in information security and the M.S. degree in computer technology from Sichuan University, China, in 2017 and 2020, respectively, and the Ph.D. degree from the University of Technology Sydney, Australia, in 2024. His research interests include security and privacy issues in deep learning and semantic communications. He has been actively involved in the research community by serving as a Reviewer for prestige journals, such as *ACM Computing Surveys*, IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, and *ACM Transactions on Knowledge Discovery from Data*, and international conferences, such as IEEE ICC and IEEE GLOBECOM.

Xuemin (Sherman) Shen (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990. He is currently a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include network resource management, wireless network security, the Internet of Things, AI for networks, and vehicular networks. He received the "West Lake Friendship Award" from Zhejiang Province in 2023, the President's Excellence in Research from the University of Waterloo in 2022, the Canadian Award for Telecommunications Research from Canadian Society of Information Theory in 2021, the R. A. Fessenden Award from IEEE, Canada, in 2019, the Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) in 2019, the James Evans Avant Garde Award from the IEEE Vehicular Technology Society in 2018, the Joseph LoCicero Award in 2015 and the Education Award in 2017 from the IEEE Communications Society (ComSoc), and the Technical Recognition Award from the Wireless Communications Technical Committee in 2019 and the AHSNTechnical Committee in 2013. He also received the Excellent Graduate Supervision Award from the University of Waterloo in 2006 and the Premier's Research Excellence Award from the Province of Ontario, Canada, in 2003. He served as the Editor-in-Chief for IEEE INTERNET OF THINGS JOURNAL, IEEE NETWORK, and *IET Communications*. He is the Past President of the IEEE Communications Society. He was the Vice President for Technical and Educational Activities, the Vice President for Publications, the Member-at-Large on the Board of Governors, the Chair of the Distinguished Lecturer Selection Committee, and a member of the IEEE Fellow Selection Committee of the ComSoc. He is a Registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, a Chinese Academy of Engineering Foreign Member, an International Fellow of the Engineering Academy of Japan, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society.