Will 6G Be Semantic Communications? Opportunities and Challenges From Task Oriented and Secure Communications to Integrated Sensing

Yalin E. Sagduyu, Tugba Erpek, Aylin Yener, and Sennur Ulukus

Abstract

This paper explores opportunities and challenges of task-oriented communications (TOC), also referred to as goal-oriented communications, and semantic communications (SemCom) for next-generation (NextG) communication networks through the integration of multi-task learning (MTL). This approach employs deep neural networks representing a dedicated encoder at the transmitter and multiple task-specific decoders at the receiver, collectively trained to handle diverse tasks including semantic information preservation, source input reconstruction, and integrated sensing and communications (ISAC). Performance is improved with this MTL approach compared to conventional communication and sensing schemes. To extend the applicability from point-to-point links to multi-receiver settings, we envision the deployment of decoders at various receivers, where decentralized learning addresses the challenges of communication load and privacy concerns, leveraging federated learning or split learning techniques that distribute model updates or split segments of models across decentralized nodes, respectively. However, the efficacy of this approach is contingent on the robustness of the employed deep learning models. We scrutinize potential vulnerabilities stemming from adversarial attacks during both training and testing phases. These attacks aim to manipulate both the inputs at the encoder at the transmitter and the signals received over the air on the receiver side, highlighting the importance of fortifying semantic information against potential multi-domain exploits. Overall, the joint and robust design of TOC, SemCom and ISAC in the MTL framework emerges as the key enabler for context-aware, resource-efficient, and secure communications ultimately needed in NextG network systems.

Introduction

The landscape of wireless communications is undergoing a profound transformation, driven by the promise of next-generation (NextG) communications technologies and their potential applications. This transformation is strengthened by the ability to convey *semantic information* as a fundamental cornerstone of NextG communications, introducing a paradigm shift that transcends conventional data transfer by embedding network communications with contextual depth and intelligence [1], [2], [3].

Task (goal)-oriented communications, characterized by the utilization of deep neural networks (DNNs) at both the transmitter and receiver, aims to convey semantic information that is pertinent to the significance of a task, catering to the dynamic information needs of emerging applications powered by NextG communications [4], [5]. Task-oriented communications (TOC) involves training an encoder at the transmitter for source coding, channel coding, and modulation operations jointly with a decoder at the receiver to accomplish a specific task, such as classification (using the data samples of the transmitter) without the need to transfer and reconstruct all data samples to the receiver [5], [6]. A potential task to complete at the receiver is to check whether the semantic information is preserved using a machine learning (ML) classifier. If the output matches the meaning of information at the transmitter, it is verified that the correct meaning is conveyed to the receiver. In addition to this semantic preservation task, *semantic communications* (SemCom) involves also an information reconstruction task at the receiver. These two tasks are executed together by jointly training an encoder at the transmitter and two decoders (one for each task) at the receiver [7].

By combining the two objectives of reconstructing information and preserving semantic meaning (captured by successful completion of a task at the receiver), *Multi-task learning* (MTL) lays the foundation for SemCom that transcends conventional data transfer [7]. The building blocks for transition from conventional communications to SemCom is illustrated in Fig. 1. This paradigm shift is poised to shape the trajectory of emerging 6G applications, such as augmented reality/virtual reality (AR/VR), Internet of Things (IoT), smart cities, vehicle-to-everything (V2X) networks, and autonomous driving.

- In AR/VR applications, SemCom can enhance immersion by enabling users to actively engage with digital environments.

Digital Object Identifier:
10.1109/MNET.2024.3425152
Date of Current Version:
18 November 2024
Date of Publication:
8 July 2024

Yalin E. Sagduyu (corresponding author) and Tugba Erpek are with Nexcepta, Gaithersburg, MD 20878 USA; Aylin Yener is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA; Sennur Ulukus is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA.

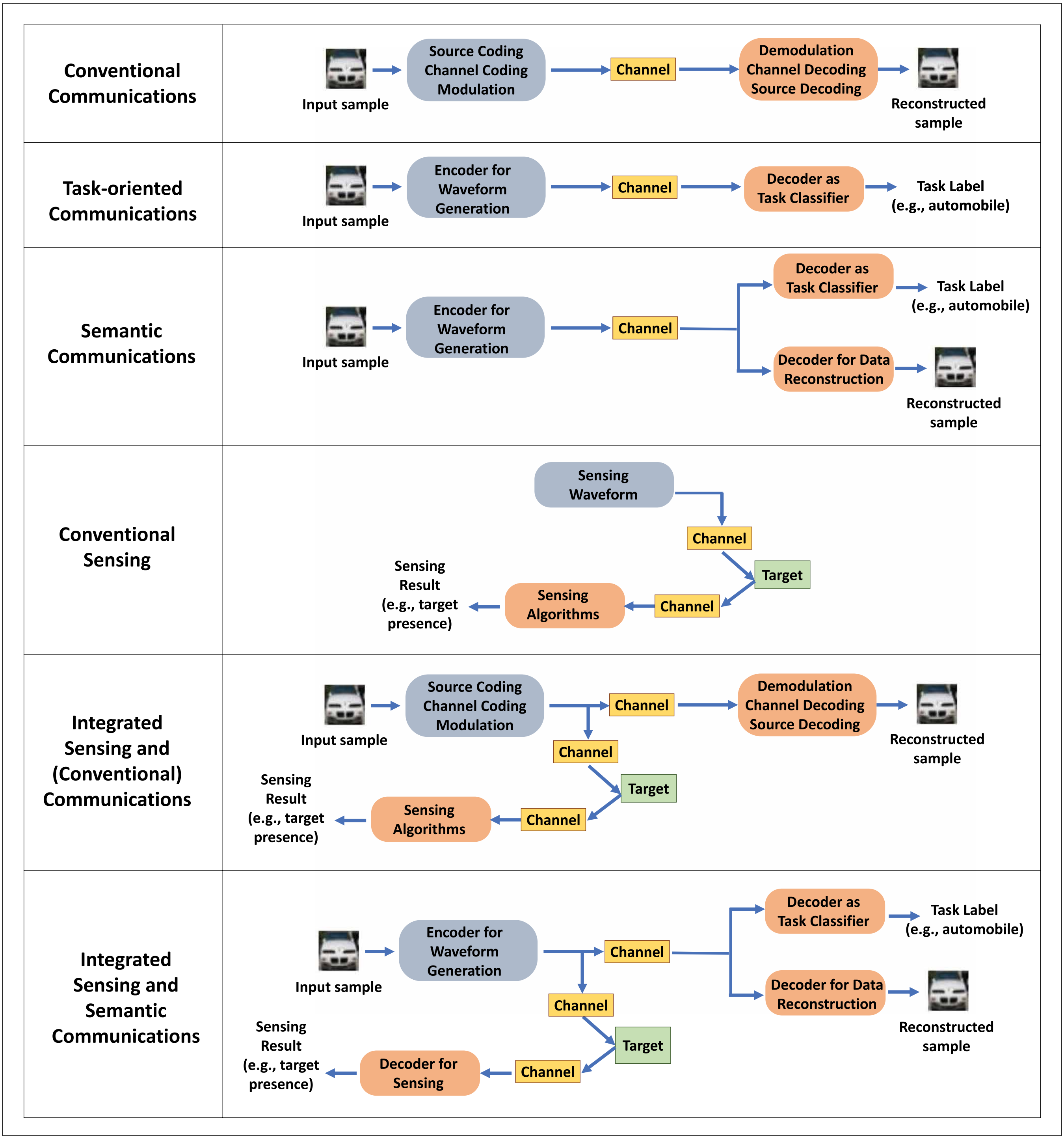
Conventional Communications
Input sample
Source Coding Channel Coding Modulation
Channel
Demodulation Channel Decoding Source Decoding
Reconstructed sample
Task-oriented Communications
Input sample
Encoder for Waveform Generation
Channel
Decoder as Task Classifier
Task Label (e.g., automobile)
Semantic Communications
Input sample
Encoder for Waveform Generation
Channel
Decoder as Task Classifier
Task Label (e.g., automobile)
Decoder for Data Reconstruction
Reconstructed sample
Conventional Sensing
Sensing Waveform
Channel
Target
Channel
Sensing Algorithms
Sensing Result (e.g., target presence)
Integrated Sensing and (Conventional) Communications
Input sample
Source Coding Channel Coding Modulation
Channel
Demodulation Channel Decoding Source Decoding
Reconstructed sample
Channel
Target
Channel
Sensing Algorithms
Sensing Result (e.g., target presence)
Integrated Sensing and Semantic Communications
Input sample
Encoder for Waveform Generation
Channel
Decoder as Task Classifier
Task Label (e.g., automobile)
Decoder for Data Reconstruction
Reconstructed sample
Channel
Target
Channel
Decoder for Sensing
Sensing Result (e.g., target presence)

FIGURE 1. From conventional communications and sensing to TOC, SemCom, and ISAC.

Preserving semantic information, like patient attributes in telemedicine and machinery failure points in remote maintenance, ensures that conveyed information retains its context, enhancing AR/VR interactions.

- In IoT and smart city applications, where interconnected sensors and actuators demand exchange of contextual information such as target types in surveillance feeds or operational states of machinery, SemCom can maintain the integrity of information across diverse applications, from smart homes to industrial IoT.
- In V2X network and autonomous driving applications, SemCom can contribute to the safety and efficiency of vehicular communications. Collaborative validation of semantic information across tasks, like classifying traffic signs or road conditions, fosters a holistic understanding and cooperative decision-making among vehicles.

MTL of SemCom can be enriched by *integrated sensing and communications* (ISAC). As illustrated in Fig. 1, transmitted signals are leveraged for dual purpose of sensing the environment [8]. While a decoder can detect and

As we explore TOC, SemCom, and ISAC, this paper navigates the intricate landscape of NextG networks, providing insights and strategies to propel the context-awareness, reliability, resource efficiency, and security of communications in the era of 6G.

classify signals reflected from potential targets, a sensing loss is combined with reconstruction and semantic losses for integrated sensing and SemCom, ensuring resource efficiency by eliminating separate probe signals for environmental sensing [9]. The sensing component in integrated sensing and conventional communications can be designed via signal processing algorithms such as cyclostationary and energy detectors, or via ML. In integrated sensing and SemCom, the sensing component is formulated as another decoder DNN that is jointly trained with the encoder at the transmitter (for waveform generation) and the two decoders at the receiver (for task classification and information reconstruction).

Extending beyond traditional point-to-point links, there is a need to adopt *multi-receiver communications* with different tasks to be completed at different receivers [10]. By deploying decoders at various receivers for different tasks, this extension needs to embrace decentralized learning to address challenges associated with communication load and privacy. *Federated learning* can distribute the MTL process across network nodes for collaborative learning [11], while *split learning* can distribute segments of models across nodes.

Amidst these advancements, the security challenges associated with TOC, SemCom and ISAC cannot be overlooked. DNNs are known to be highly vulnerable to even slight variations in training and test data [12]. These variations can be deliberately added as small perturbations by adversaries, leading to *adversarial machine learning* (AML) threats that can disrupt the training and testing of the underlying DNNs. These stealth but powerful attacks can effectively manipulate both inputs at the encoder and signals received over the air, underscoring the critical importance of defending against *multi-domain attacks* [5], [7], [13].

As we explore TOC, SemCom, and ISAC, this paper navigates the intricate landscape of NextG networks, providing insights and strategies to propel the context-awareness, reliability, resource efficiency, and security of communications in the era of 6G. Our novel contributions are summarized as follows:

- We present the MTL approach that integrates communications (information recovery) with task execution involving semantic information validation and target sensing. This approach outperforms conventional communication and sensing schemes in terms of semantic information preservation, information reconstruction, and target sensing under AWGN and Rayleigh channels.
- We discuss potential extensions of this MTL approach to multi-user settings and decentralized operations.
- We highlight the potential vulnerabilities of this MTL approach to stealth attacks built upon AML, and discuss the need for defense mechanisms.

The remainder of the paper is organized as follows. The sections "Task-Oriented Communications," "Semantic Communications," and "Integrated Sensing and Communications" present TOC, SemCom, and ISAC, respectively, in the MTL framework. The section "Multi-User Communications and Distributed Learning" describes extension to distributed learning. The section "Security Vulnerabilities" presents vulnerabilities to AML threats. The section "Conclusion and Future Research Directions" concludes the paper and describes future research directions.

Task-Oriented Communications

Information systems rely on reliable and efficient processing of information encompassing generation, feature extraction, compression, and transfer of different data modalities. Traditionally, wireless systems are built on the principle of ensuring reliable communications subject to channel impairments. Consequently, they often overlook the significance of transferred information in relation to the underlying tasks or goals.

To unlock the full potential of 6G applications, *TOC* offers a transformative approach that prioritizes the successful execution of specific tasks over the meticulous reconstruction of transmitted information. The primary objective is not necessarily reconstructing data samples but successfully completing specific tasks (e.g., classifying the received signals to correct labels) at a receiver by utilizing the data at the transmitter.

The system model is shown in Fig. 2. In TOC, the transmitter's operations, encompassing source coding, channel coding and modulation, are conceptualized as an encoder, namely a DNN, tasked with generating and transmitting low-dimensional feature vectors. Unlike the conventional receiver chain, which focuses on information reconstruction, the task-oriented receiver employs a dedicated decoder (another DNN) to directly perform the assigned task, such as classifying the received signals, while obviating the need for the resource-intensive process of reconstructing the input samples.

TOC reduces the number of transmissions, latency, and energy consumption by limiting the size of encoder output. These gains are especially critical for emerging 6G applications requiring ultra-low latency for split-second decisions such as needed in AR/VR and V2X systems, and high energy efficiency for battery-powered devices such as deployed in IoT systems.

For all results in this paper, we assume that the transmitter has images from CIFAR-10 dataset. There are 60,000 color images, each of size 32×32×3, with a split of 50,000 and 10,000 samples for training and testing. The transmitter encodes these images and transmits their latent representations of reduced dimension over a wireless channel. This dimension is controlled by the encoder's output size, n_c, which is an important design parameter that controls the number of channel uses, the input size of decoders, and the latency in communications and sensing. Lower n_c leads to higher compression ratio (between the input image and transmitted symbols) and lower latency, while larger n_c improves sensing, task and reconstruction accuracy. Requirements for task and sensing accuracy, reconstruction loss

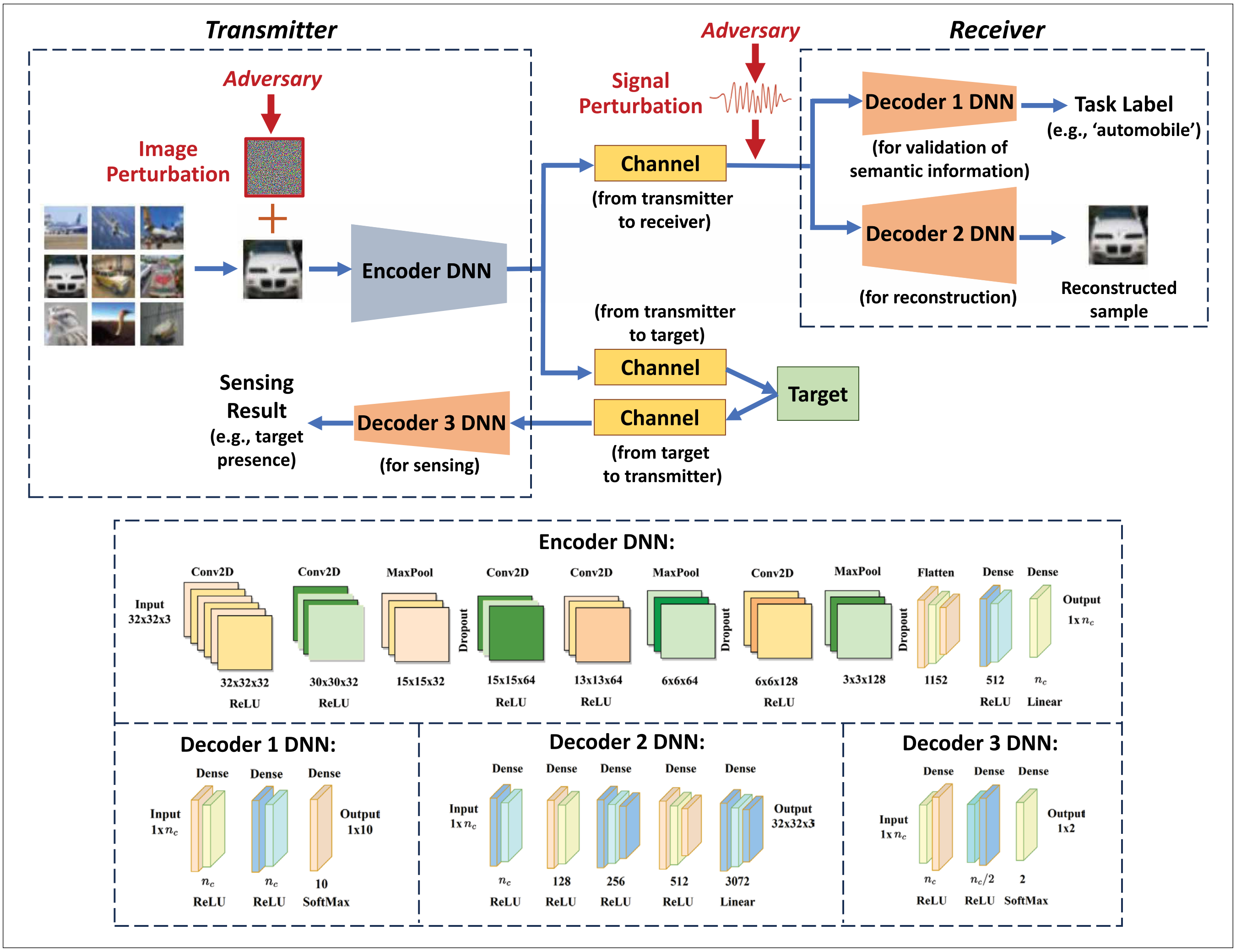

FIGURE 2. System model of MTL for TOC, SemCom, and ISAC.

and latency should be considered while selecting n_c. For example, when n_c is 20, the dimension of transferred information is reduced from 3072 to 20 with the compression ratio of 0.65%. We consider both AWGN and Rayleigh channel models. The decoder at the receiver classifies the received signals to one of the 10 labels.

The encoder-decoder pair is modeled as an end-to-end DNN that is trained by considering channel and data characteristics to optimize task performance. The optimization success is quantified via ML metrics, such as the loss function in a classification task. We consider categorical cross-entropy as the loss function for training. The DNN architectures shown in Fig. 2 are determined by gradually increasing the number of layers and layer sizes until the average classification accuracy cannot be further improved for task completion.

Fig. 3 shows how the task accuracy improves when the signal-to-noise ratio (SNR) of communication channel or n_c increases. While the task accuracy is higher under AWGN compared to Rayleigh channel in general, the performance gap closes with higher SNR or n_c. TOC increases the task accuracy compared to conventional communications scheme, where we consider separate source (de)coding, channel (de)coding and (de) modulation blocks at the transmitter-receiver pair. JPEG-2000, Reed-Solomon error correction, and 16-QAM are used for source coding, channel coding and modulation, respectively. The receiver first reconstructs the images and then classifies them using a DNN. JPEG-2000 cannot compress images reliably to the level provided by the TOC that can significantly reduce the number of channel uses as an additional benefit of TOC.

We present the MTL approach that integrates communications (information recovery) with task execution involving semantic information validation and target sensing.

Semantic Communications

In traditional communication systems, the central objective is the reliable transmission of messages, accounting for channel impairments to minimize a reconstruction loss, e.g., symbol error rate. This involves the design of individual or collaborative operations at the transmitter and receiver. DNNs play a pivotal role in this framework, acting as *autoencoders* to capture transmitter and receiver operations, including source (de)coding, channel (de)coding and (de)modulation at the transmitter-receiver pair. The goal is to minimize the end-to-end reconstruction loss.

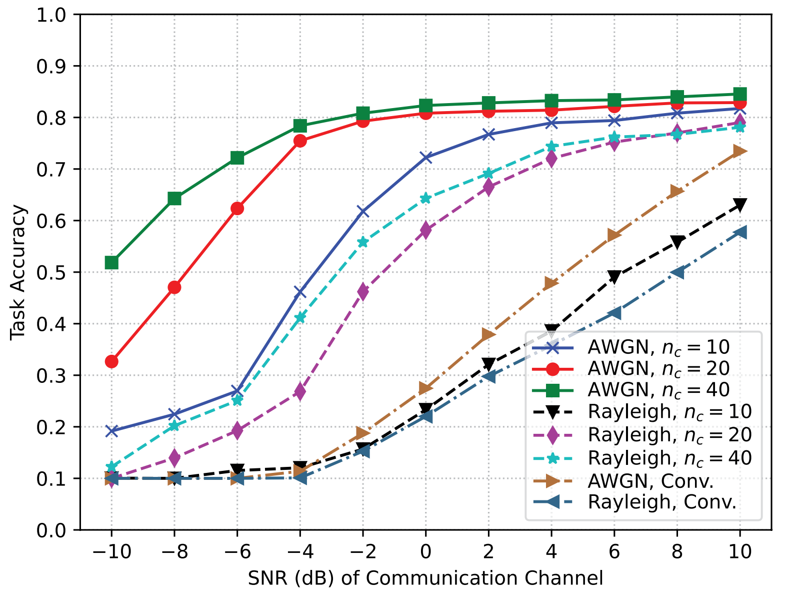

FIGURE 3. Task accuracy versus SNR (dB) of communication channel (TOC is evaluated for different values of encoder output size, n_c, and conventional communications is referred to as "Conv.").

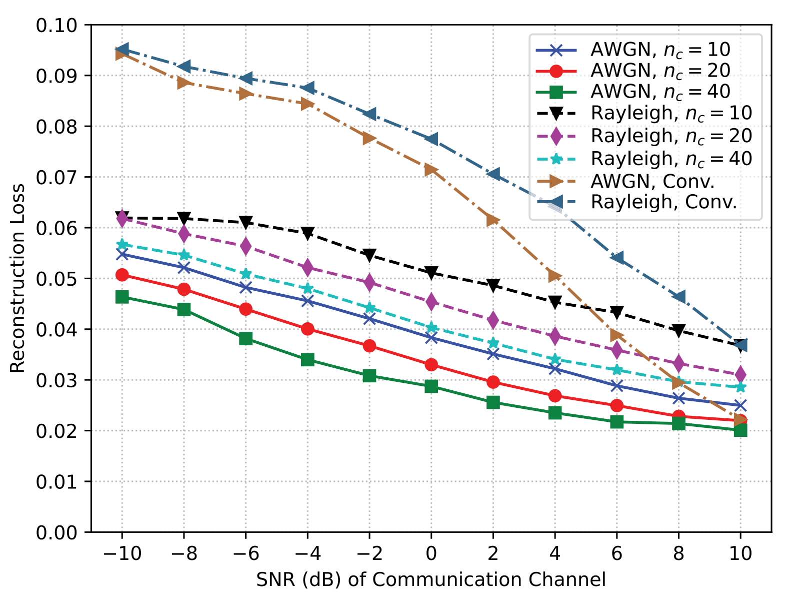

FIGURE 4. Reconstruction loss versus SNR (dB) of communication channel (SemCom is evaluated for different values of encoder output size, n_c, and conventional communications is referred to as "Conv.").

In addition to information reconstruction, *SemCom* emerges as a paradigm shift, striving to preserve the intrinsic meaning of the information conveyed to the receiver. In SemCom, the training loss for the autoencoder encompasses not only the reconstruction loss of traditional communications but also introduces the *semantic loss* to capture the preservation of meaning during information transfer, ensuring that the conveyed information retains its intended significance.

As shown in Fig. 2, an encoder is employed at the transmitter to handle source coding, channel coding, and modulation, whereas the receiver involves two decoders. One decoder focuses on reconstructing information by performing joint demodulation, channel decoding, and source decoding. Simultaneously, another decoder evaluates whether semantic information is preserved, effectively performing an ML task such as image classification. In MTL, the reconstruction loss is combined with a semantic loss by jointly training three DNNs, namely one encoder and two decoders. This method adeptly captures latent feature spaces, facilitating the reliable transfer of compressed feature vectors with minimal channel usage, all while maintaining a low semantic loss.

The exploration of SemCom spans diverse data types, ranging from text, image, and video to speech and audio. The goal is to maintain the integrity of meaning across different forms of information, aligning seamlessly with the applications of 6G networks that will benefit significantly from the preservation of semantic meaning, enriching the user experience and enabling more contextually aware communication.

Building upon the demonstration use case from the section "Task-Oriented Communications." we add mean-squared error as the reconstruction loss (to recover image samples at the receiver) in addition to the semantic loss (to validate if the semantic information is conveyed) in MTL, where we can set individual weights for losses. The DNN architecture to reconstruct data samples is shown in Fig. 2.

Fig. 4 shows how the reconstruction loss decreases when SNR or n_c increases under both AWGN and Rayleigh channel models, leading to more reliable information recovery. SemCom decreases the reconstruction loss compared to the conventional communications scheme. MTL can help the task accuracy for semantic information validation reach the level presented in the section "Task-Oriented Communications."

Integrated Sensing and Communications

The integration of sensing functionalities emerges as a pivotal component in NextG communications, expanding beyond traditional data transmissions to include the *ISAC* paradigm. This section formulates sensing as a task in MTL for SemCom, where tasks extend from preserving semantic information and reconstructing input samples to environment sensing, namely detecting a potential target.

ISAC leverages the dual use of communication signals for sensing and eliminates the need for additional probe signals, thereby conserving bandwidth, reducing delay, and enhancing energy efficiency. In 6G applications, this resource-efficient integration of sensing into communications unveils transformative possibilities. One key example is in V2X networks. The dual functionality of communication and sensing allows vehicles to exchange crucial information about their positions and intentions, while sensing and monitoring surroundings for potential hazards, identifying changing road conditions, and detecting objects such as other vehicles, pedestrians, and traffic signs.

ISAC presents two distinctive avenues: (i) existing communication waveforms are repurposed for sensing, as in WiFi sensing, (ii) a more forward-thinking alternative is to design a communication waveform with a dual purpose. In the latter approach, the waveform is crafted not solely for information exchange but also with a specific focus on sensing, especially in tasks like target detection. This strategy ensures that the designed waveform is tailored to excel in both communication and sensing tasks, maximizing their effectiveness in diverse scenarios.

To optimize the overall performance, we consider a DNN-driven design of communication waveform shown in Fig. 2 for both information transfer and sensing. Employing an encoder, the transmitter conducts joint operations encompassing source coding, channel coding, and modulation to generate signals. The receiver employs another DNN, serving as a decoder, for joint operations including demodulation, channel decoding, and source decoding to reconstruct data samples. In the meantime, the transmitted signal has a dual role, facilitating communication with the receiver and enabling sensing capabilities. In the presence of a target, the signal is reflected from the target and potentially received by the transmitter or another node, leading to two possible cases of sensing. In *monostatic sensing*, the transmitter performs the sensing operation based on the signals reflected from the target. In *bistatic sensing*, the transmitter of signals and the receiving end of reflected signals are separated such that the receiving end could be the receiver of communication signals or any other receiver deployed for environment sensing.

In all these cases, an additional decoder is employed for sensing. This decoder can be tasked with detecting the target's presence and determining its range. Through MTL, all involved DNNs undergo joint training, considering data, channel, and potential target characteristics. In MTL, an additional decoder can be incorporated at the receiver, operating as a task classifier. This decoder assesses the fidelity of task completion such as image classification based on received signals, enhancing the incorporation of semantics into ISAC.

Building upon the MTL framework from the section "Semantic Communications," we measure the sensing loss as categorical cross-entropy in addition to reconstruction and semantic losses (discussed in sections "Task-Oriented Communications" and "Semantic Communications") such that all the underlying DNNs are jointly trained according to the weights assigned to the MTL losses. The DNN architecture for sensing is shown in Fig. 2. Considering the mono-static sensing case, Fig. 5 shows how the sensing accuracy (namely, the probability of successfully detecting a target) increases with the SNR of channels to and from the target, whereas the task accuracy and reconstruction loss can reach the levels presented in sections "Task-Oriented Communications" and "Semantic Communications," respectively, through MTL. Sensing accuracy is higher when n_c is low even at low SNR and the performance gap is smaller compared to task accuracy with increasing n_c since the sensing task is less complex compared to TOC and is also less sensitive to channel degradation. The MTL approach for ISAC increases the target sensing accuracy compared to the conventional sensing scheme for which we consider an energy detector to sense the signals reflected from the targets.

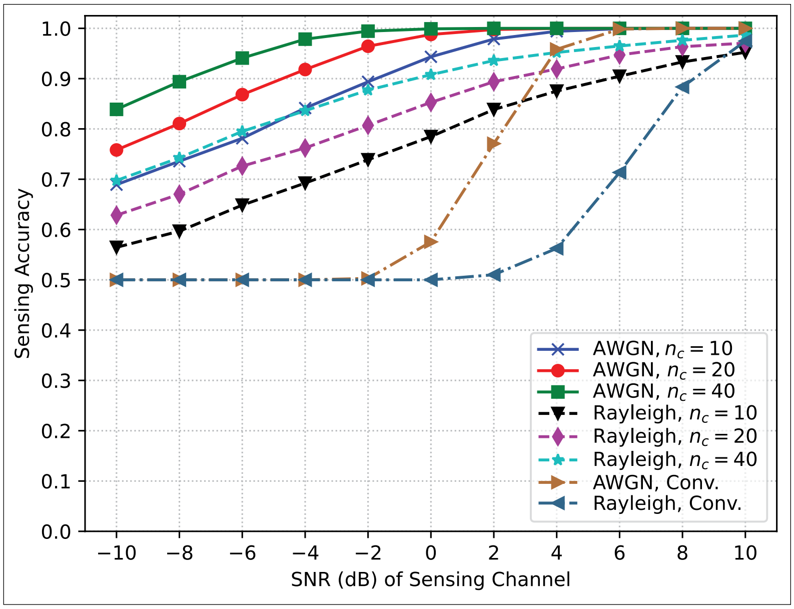

FIGURE 5. Sensing accuracy versus SNR (dB) of sensing channel that corresponds to channels to and from the target (integration of sensing into SemCom is evaluated for different values of encoder output size, n_c, and conventional sensing is referred to as "Conv.").

Multi-User Communications and Distributed Learning

The extension of the presented MTL approach to multiple users introduces a paradigm shift, transforming the traditional point-to-point communications into a collaborative model via a distributed framework. This multi-user approach not only accommodates the diverse information needs of different users but also facilitates the validation of semantic information by multiple receivers, each potentially tasked with distinct objectives.

Extension multiple receivers. The studies of SemCom have initially focused on preserving the meaning of information for a single receiver. In the multi-user context, this approach can be expanded to cater to diverse recipients, each assigned a unique task. Each receiver, equipped with a specific decoder tailored to its task, validates the semantic information conveyed. This extension allows for the simultaneous fulfillment of multiple tasks, ranging from classification to contextual understanding, leveraging semantic information to its fullest potential.

The need for *multi-user SemCom* arises from the diverse information requirements inherent in 6G applications, where users are envisioned with various semantic tasks crucial for comprehensive information understanding and collaborative decision-making. 6G's vision encompasses a multitude of applications, such as AR/VR, IoT, smart city, V2X networks and autonomous driving, requiring a communication system that can cater to varied and nuanced semantic information needs. Multi-user SemCom becomes essential to ensure that each participant receives information tailored to their specific tasks, fostering a more intelligent and responsive 6G ecosystem.

MTL in a collaborative setting. Ensuring that each receiver's semantic task is adequately addressed while preserving global semantic coherence poses a significant technical challenge. To achieve multi-user SemCom, the concept of MTL can be extended to a *distributed and collaborative learning* setting. Each receiver's task becomes a distinct facet of the overarching communication objective. The encoder at the transmitter is trained to accommodate these diverse tasks, and each decoder at the receiver is tailored to validate the semantic information within the context of its designated task.

Transition to federated and split learning. To address the distributed nature of multi-user SemCom, the evolution from MTL to federated and split learning becomes pivotal. Federated learning enables decentralized training across multiple receivers, optimizing the learning process by leveraging local datasets at each receiver while collectively enhancing semantic information across the entire network. Its privacy-preserving nature ensures that sensitive information remains on user devices, addressing data sharing concerns. Enabling decentralized model training optimizes resource utilization, seamlessly integrating with edge computing for enhanced responsiveness in applications like AR/VR, IoT, smart city, V2X networks, and autonomous driving. Federated learning's adaptability to diverse data distributions ensures robust and scalable models, crucial in multi-user SemCom. Additionally, the reduction in exchanged network data enhances efficiency, a vital aspect in 6G.

Shared and task-specifics components of DNNs can be partitioned across different nodes via split learning. While federated learning aggregates model updates from nodes to a global model without sharing raw data, split learning distributes segments of a model across nodes, sharing only intermediate activations for collaborative training. In MTL, split learning can be particularly advantageous by allowing different parts of the MTL model to be trained on different nodes, each focusing on specific tasks of the overall learning objective.

Security Vulnerabilities

With the reliance on DNNs, SemCom becomes susceptible to *AML threats*. This section explores adversarial, poisoning, and backdoor attacks in *multi-domain* attack scenarios: manipulation of the input to the transmitter's encoder and manipulation of the over-the-air signal reaching the receiver.

Adversarial attacks target the robustness of the DNNs by strategically tampering with test data. Adversaries may manipulate the input to the encoder at the transmitter or interfere with signals received by the decoder over the air. In both cases, subtle perturbations can be introduced to deceive the learning process, leading to the distortion of semantic information, information reconstruction, and sensing results. We can distinguish targeted and untargeted adversarial attacks. In a targeted adversarial attack, adversaries manipulate the model's output with the goal of causing misclassification into a specific, predefined target class. In an untargeted adversarial attack, the goal is to cause any misclassification without steering the decision towards a predefined target.

Poisoning attacks focus on compromising the integrity of DNNs by injecting malicious data into training process. Adversaries may tamper with the training dataset of the transmitter at the input of the encoder or may strategically interfere with and manipulate over-the-air signals that are collected for training. These attacks aim to distort the learning representations, resulting in biased semantic features and leading to incorrect task executions at multiple decoders.

Backdoor (Trojan) attacks insert hidden triggers to the training data such as manipulating some training data samples and falsifying their labels. Adversaries may embed subtle backdoors by manipulating selected inputs to the encoder during training. Adversaries may also insert triggers (e.g., phase shifts) to some of wireless signals transmitted over a channel. These triggers are later activated selectively in test time. This stealth attack can result in manipulation of MTL outcomes at different decoders for task classification, information reconstruction, and sensing purposes.

Mitigation strategies are needed to protect the MTL framework for TOC, SemCom, and ISAC in response to the multi-faceted security threats.

- Proactive defenses, including adversarial training, randomized smoothing, and certified methods, enhance model resilience against adversarial attacks. Continual model audits and anomaly detection mechanisms contribute to identifying and mitigating adversarial manipulations.
- Defense against poisoning attacks involves data validation, robust outlier detection, and leveraging decentralized learning like federated learning to reduce the impact of poisoned data. Continuous monitoring of semantic representations ensures the integrity of learned features.
- Backdoor threat mitigation entails measures such as input validation, adversarial training, and anomaly detection for over-the-air signal manipulations. Regular updates and diversification of the training dataset reduce the risk of backdoor insertion.

Large Language Models (LLMs) hold the potential to protect the MTL tasks against AML attacks by cross-referencing and validating the consistency and multi-modal authenticity (including textual and visual information), thereby fortifying the resilience to adversarial strategies. LLMs can be also used to generate detailed descriptions of potential attack scenarios, allowing for the preemptive strengthening of defense mechanisms.

To demonstrate security vulnerabilities, we consider adversarial attacks that add small perturbations to the encoder inputs to cause potential errors in the classification task that checks whether the semantic information is preserved. We consider the Fast Gradient Sign Method (FGSM) to generate perturbations for an untargeted adversarial attack. By calculating the gradient of the model's loss with respect to the input, the adversary generates adversarial examples by perturbing the input data in the direction that maximizes the semantic loss, aiming to mislead the model's predictions for semantic information validation. For a stealthy attack, we suppose that the strength of perturbation added to input samples is limited by a given perturbation-to-signal ratio (PSR).

Fig. 6 shows that, as the PSR increases, the task accuracy quickly drops leading to excessive number of classification errors, while Gaussian noise is not as effective unless the PSR is very high. We set n_c to 20 for these results. The choice of n_c is critical in terms of controlling security vulnerabilities. As shown in Fig. 3, task accuracy increases with n_c before an adversarial attack is launched. With increasing n_c, the encoder output and the decoder input sizes (for task classification) increase. This leads to larger DNN sizes for the encoder-decoder pair. As the complexity of the underlying DNNs increases, it becomes more difficult for the adversary to fool them. Therefore, larger n_c can achieve better protection of task accuracy under adversarial attacks (at the expense of increased latency, computational complexity, and memory footprint). The task accuracy under attack improves (over all PSRs) up to 6.71% and 47.76%, when n_c increases from 10 to 20, under AWGN and Rayleigh channels, respectively, and up to 5.98% and 24.88%, when n_c increases from 20 to 40, under AWGN and Rayleigh channels, respectively.

This attack example is built upon a white-box attack, where the adversary possesses complete knowledge of the target model. In contrast, a black-box attack can be launched by involving no information about the model's internal structure. Gray-box attacks fall in between, with partial knowledge. These variations can be studied to assess the attack surface exploited by different types of adversaries, each possessing different levels of knowledge.

Conclusion and Future Research Directions

The convergence of SemCom with task orientation and integrated sensing in MTL holds the key to unlocking the transformative potential of context-aware, scalable, and secure communications for 6G. This synergistic alignment not only optimizes information transfer efficiency but also establishes a foundation for intelligent, context-aware communications. TOC prioritizes successful task completion over conventional information reconstruction, crucial for addressing the diverse demands of 6G applications. SemCom contributes an additional layer of reliable data reconstruction in addition to preserving semantic information. Sensing is added as another task in MTL by utilizing transmitted signals for both communication and sensing purposes. Distributed operation with multiple receivers via federated or split learning can support collaborative model training, addressing communication load and privacy concerns for a more adaptive and scalable system. However, security implications due to AML underscore the imperative to fortify MTL for TOC, SemCom, and ISAC against emerging threats.

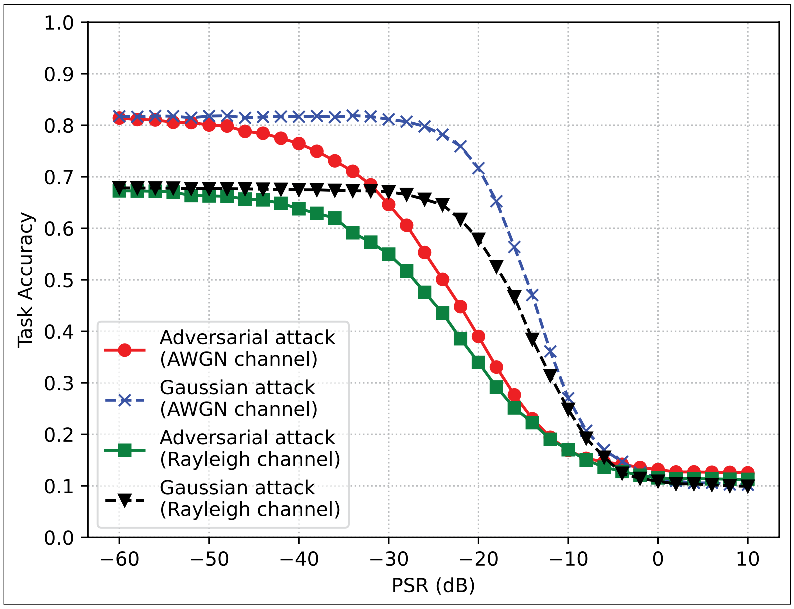

FIGURE 6. Task accuracy versus PSR (SNR = 3dB, n_c = 20) under adversarial attack.

The convergence of SemCom with task orientation and integrated sensing in MTL holds the key to unlocking the transformative potential of context-aware, scalable, and secure communications for 6G.

There are various future research directions that can help maximize the transformative potential of the proposed approach in NextG systems. The use of more complex (e.g., V2V4Real and Berkeley DeepDrive datasets) and multi-modal datasets (e.g., integrating visual and textual information) can extend the fidelity of performance evaluation. *Timeliness* and *value* objectives can be incorporated along with preserving semantic information that can be enriched in different granularity levels [1]: (i) microscopic level: relative importance of different events, (ii) mesoscopic level: information attributes including freshness and value of information (e.g., importance or urgency in data vs. control packets, depending on applications), (iii) macroscopic level: system-level information including distortion and timing mismatch between the transmitted and reconstructed samples. The information freshness aspect, introducing metrics like the *age of information* (AoI) [14], is critical for 6G applications with fast decision-making mandates. In TOC and SemCom, novel metrics like *task age* or *age of semantic information* are needed for completing tasks and preserving semantic information in a timely manner [15]. The semantic knowledge base updates can be also integrated into the communication process to help with transmitting fewer symbols without semantic performance degradation. By sharing semantic knowledge base that maintains label information of each data sample, the receiver can check if the label is preserved during information transfer. Semantic knowledge base can also indicate how similar data samples of different labels are or if they belong to the same group. This can help penalize misclassifications to wrong labels or label groups. The imperative to enhance *privacy* in the presented MTL framework suggests several research avenues to prevent eavesdroppers from learning the outcomes of different tasks and semantic information. Another loss function can be added to MTL to penalize the detection by an eavesdropper. Alternatively, adversarial attacks can be used as a defense mechanism by adding perturbations to the signals at the encoder's output to fool an eavesdropper into missing task outcomes and semantic information.

References

[1] M. Kountouris and N. Pappas, "Semantics-empowered communication for networked intelligent systems," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 96–102, Jun. 2021.

[2] E. Uysal et al., "Semantic communications in networked systems," *IEEE Netw.*, vol. 36, no. 4, pp. 233–240, Oct. 2022.

[3] D. Gündüz et al., "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5–41, Jan. 2023.

[4] Y. Shi et al., "Task-oriented communications for 6G: Vision, principles, and technologies," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 78–85, Jun. 2023.

[5] Y. E. Sagduyu, S. Ulukus, and A. Yener, "Task-oriented communications for NextG: End-to-end deep learning and AI security aspects," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 52–60, Jun. 2023.

[6] H. Xie et al., "Deep learning enabled semantic communication systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2663–2675, 2021.

[7] Y. E. Sagduyu et al., "Is semantic communication secure? A tale of multi-domain adversarial attacks," *IEEE Commun. Mag.*, vol. 61, no. 11, pp. 50–55, Nov. 2023.

[8] F. Liu et al., "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[9] Y. E. Sagduyu et al., "Joint sensing and semantic communications with multi-task deep learning," 2023, *arXiv:2311.05017*.

[10] Y. E. Sagduyu et al., "Multi–receiver task-oriented communications via multi–task deep learning," in *Proc. IEEE Future Netw. World Forum (FNWF)*, Nov. 2023, pp. 1–6.

[11] M. Mortaheb et al., "Personalized federated multi-task learning over wireless fading channels," *Algorithms*, vol. 15, no. 11, p. 421, Nov. 2022.

[12] D. Adesina et al., "Adversarial machine learning in wireless communications using RF data: A review," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 77–100, 1st Quart., 2023.

[13] Y. E. Sagduyu et al., "Vulnerabilities of deep learning-driven semantic communications to backdoor (Trojan) attacks," in *Proc. 57th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2023, pp. 1–6.

[14] R. D. Yates et al., "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.

[15] Y. E. Sagduyu, S. Ulukus, and A. Yener, "Age of information in deep learning-driven task-oriented communications," in *Proc. IEEE INFOCOM Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, May 2023, pp. 1–6.

Biographies

Yalin E. Sagduyu (Senior Member, IEEE) (yalinsagduyu@ieee.org) received the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park, MD, USA. He is currently the Chief Technology Officer at Nexcepta. He has held research positions in academia and industry, including Virginia Tech, the University of Maryland, and Intelligent Automation, Inc. His research interests include wireless communications, networking, security, and machine learning. He received the IEEE HST Best Paper Award. He is also an Editor of IEEE Transactions on Communications and IEEE Transactions on Cognitive Communications and Networking.

Tugba Erpek received the Ph.D. degree in electrical and computer engineering from Virginia Tech. She was a Research Associate Professor with the Intelligent Systems Division, Virginia Tech National Security Institute, a Lead Scientist at Intelligent Automation, Inc., and a Senior Communications Systems Engineer at Shared Spectrum Company. She is currently a Principal Scientist at Nexcepta. Her research interests include wireless communications, networks, 5G and beyond communications, security and machine learning.

Aylin Yener (Fellow, IEEE) is currently a Roy and Lois Chope Chaired Professor with The Ohio State University. She received the 2014 IEEE Marconi Paper Award, the 2018 IEEE Communications Society WICE Outstanding Achievement Award, the 2019 IEEE Communications Society Best Tutorial Paper Award, and the 2020 IEEE Communication Theory Technical Achievement Award. She was the President of the IEEE Information Theory Society. She is the Director-elect for IEEE Division IX. She is also the Editor-in-Chief of IEEE Transactions on Green Communications and Networking.

Sennur Ulukus (Fellow, IEEE) received the Ph.D. degree from Rutgers University. She is currently the Anthony Ephremides Professor in information sciences and systems with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA. She received the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, the 2019 IEEE Communications Society Best Tutorial Paper Award, the 2020 IEEE Communications Society WICE Outstanding Achievement Award, and the TCGCC Distinguished Technical Achievement Recognition Award.