Reconfigurable Intelligent Surfaces-Assisted Task-Oriented Communications for AI-Driven Vertical Applications

Shuaishuai Guo*, Jie Zhou*, Zhengyang Li*, Peng Zhang†, and Shuang Zhang⋆
*Control Science and Engineering, Shandong University, China
†Weifang University, China
⋆ Hebei Normal University, China
Email: shuaishuai_guo@sdu.edu.cn, jiezhou@mail.sdu.edu.cn, zhengyang_li@mail.sdu.edu.cn, sduzhangp@163.com, zshuang@hebtu.edu.cn

***Abstract*—The integration of reconfigurable intelligent surfaces (RIS) into next-generation wireless networks offers a promising avenue for enhancing communication performance across diverse vertical applications. This paper explores the utilization of RIS to support task-oriented communications, driven by the increasing need for intelligent data processing in AI-driven scenarios such as autonomous vehicles and smart factories. Unlike traditional communication systems focused on exact data recovery, task-oriented communications facilitated by RIS prioritize the precise execution of specific tasks, potentially transforming the wireless landscape by reducing data transmission burdens and enhancing system efficacy. We propose a task-specific RIS approach aimed at improving the task inference performance of legitimate receivers while thwarting eavesdroppers. Experimental studies demonstrate the effective configuration of RIS to meet the demands of cooperative task inference by dynamically adapting to the wireless environment. This research underscores the pivotal role of intelligent surfaces in achieving high-performance, task-specific communications, highlighting their significance in future wireless infrastructures.**

***Index Terms*—Reconfigurable intelligent surfaces, task-oriented communications, vertical applications**

I. Introduction

Vertical sectors, such as manufacturing, healthcare, automotive, and more, require tailored wireless communication solutions that cater to specific industry needs [1]. While the fifth-generation (5G) mobile communication technology offers transformative potential across various sectors with its promise of high speeds, low latency, and increased connectivity, deploying 5G technology requires significant investment in new infrastructure, including the installation of a large number of small cells due to 5G's shorter range compared to previous technologies. The high costs associated with this extensive infrastructure overhaul can be a barrier for many operators, especially in less populated areas where the return on investment is lower. Operators are under pressure to make costly investments in new infrastructure while facing uncertainty about future revenue streams from 5G services.

Reconfigurable intelligent surfaces (RIS) emerge as a promising low-cost solution to mitigate these challenges and shape the future direction of wireless communications. RIS can enhance the performance of wireless networks by intelligently manipulating electromagnetic waves, thus improving signal propagation and coverage without the need for additional transmitters. This capability allows for a more cost-effective augmentation of 5G and beyond networks, reducing the number of required base stations and the associated infrastructural costs. Furthermore, RIS technology supports the specific needs of various verticals by enabling highly customizable wireless environments [2]. This adaptability makes RIS an invaluable asset in realizing the full potential of 5G and beyond technology across diverse applications, ensuring more sustainable and economically viable network enhancements. Therefore, integrating RIS with 5G and beyond infrastructures not only addresses the immediate challenges of network deployment but also sets a foundational technology for future advancements in the telecommunications sector.

RIS technology has been primarily developed and deployed to enhance overall network performance without specific considerations for individual tasks. Their main roles have included improving signal coverage [3], boosting energy efficiency [4], reducing error rate [5], and enhancing data throughput across the board [6]. These applications do not differentiate between the types of data being transmitted or the specific requirements of the tasks at hand. With the advent of sophisticated applications powered by artificial intelligence (AI), such as autonomous vehicles, smart manufacturing, and personalized healthcare, there is a growing need for communication technologies that can prioritize and optimize data transmission based on the specific requirements of different tasks. For instance, autonomous vehicles require ultra-reliable and low-latency communication for safety-critical detection tasks, while telemedicine might prioritize high-quality generation tasks for augmented reality. AI-driven applications often generate and require processing of massive amounts of data with varying levels of importance and urgency.

By integrating content awareness directly with RIS technology, it becomes possible to develop intelligent surfaces that can dynamically adjust their properties to prioritize or enhance communication based on the specific AI task being

ISBN 978-3-903176-65-2 © 2024 IFIP

performed. This could mean altering the way signals are reflected and focused to prioritize certain types of data over others or to enhance the signal quality for particular receivers engaged in critical tasks. Developing RIS with capabilities to support task-specific communications can lead to several improvements. First, optimizing communications for specific tasks can significantly reduce the bandwidth and energy consumption by not uniformly amplifying all types of traffic. Second, task-specific enhancements can improve the performance and reliability of critical applications, thereby increasing the overall effectiveness of AI-driven systems. Third, it allows for greater customization of services and infrastructure, providing tailored solutions that meet the exact needs of different vertical applications.

The concept of task-specific RIS opens up new avenues for research, particularly in how these surfaces can be designed to understand and interact with varying data types and communications needs dynamically. Research can explore how machine learning algorithms can be integrated into RIS to make real-time decisions about how to manipulate electromagnetic waves to benefit specific tasks. Pursuing this direction also involves addressing several challenges, such as developing algorithms that can quickly analyze the type of data being transmitted and make immediate decisions on how the RIS should react. Additionally, there would be challenges related to interoperability between different AI systems and RIS, as well as ensuring the privacy and security of communications.

By focusing on making RIS supportive of task-specific needs, especially in AI-driven applications, the technology can become even more revolutionary, pushing the boundaries of what is currently possible in wireless communications and enabling more advanced, efficient, and effective solutions tailored to the needs of the future. This paper aims to explore the potential of RIS-assisted task-oriented communications for AI-driven vertical applications, highlighting key research directions, challenges, and opportunities in this emerging field.

II. The Role of RIS in AI-Driven Vertical Applications and Its Opportunities

AI-driven vertical applications represent a significant advancement in various industries, including manufacturing, healthcare, automotive, and beyond, as illustrated in Fig. 1. These applications leverage AI technologies to automate processes, enhance decision-making, and improve overall efficiency. In such dynamic environments, efficient and reliable communication systems are essential to support the exchange of data and commands between interconnected devices and systems. RIS is a promising technology to enhance and secure task-oriented communications in AI-driven vertical applications. Comprising passive surfaces embedded with numerous low-cost, programmable elements capable of manipulating electromagnetic waves, RIS offers a transformative solution to address the communication challenges in such environments. By intelligently adjusting the phase and amplitude of reflected signals, RIS can improve the task inference performance of legitimate receivers while e thwarting eavesdroppers.

Traditional communication systems prioritize exact data recovery, which may not align with the task-oriented communication needs of dynamic AI applications [7]–[9]. Furthermore, traditional systems may lack the flexibility to adapt to changing communication requirements in real-time, posing challenges in environments with diverse communication needs. Integrating RIS into task-oriend communications offers a solution to prioritize and optimize data transmission based on the specific requirements of different tasks. For instance, in smart manufacturing environments, RIS can dynamically adjust signal reflections to optimize communication between robotic systems for cooperative inference, enhancing overall efficiency and productivity. RIS can adaptively manipulate the wireless channel to optimize transmission for specific tasks, leading to improved efficiency and adaptability in industrial settings.

The integration of RIS into AI-driven vertical applications holds tremendous potential to revolutionize communication systems in various industries. By prioritizing task-oriented communications and leveraging the capabilities of RIS technology, AI-driven vertical applications can achieve higher levels of efficiency, reliability, and adaptability. This transition presents a strategic response to the numerous challenges identified, transforming potential barriers into opportunities for profound advancements in communication technology.

As RIS settings are directly optimized for task inference, these systems address critical needs for reduced latency and enhanced bandwidth management. This is particularly vital in cooperative inference scenarios, where synchronization across multiple nodes is crucial for efficient model training and execution. By reducing communication delays, RIS-enhanced systems enable more effective real-time data processing and decision-making capabilities.

Furthermore, integrating RIS with generative models opens new avenues for improving data compression and communication techniques. This enhancement is essential for applications where bandwidth is limited but high data fidelity is required, such as in streaming high-definition video or transmitting detailed medical images. Through these advanced techniques, RIS can maintain or improve the quality of data while efficiently managing the network's bandwidth. These are appealing in telemedicine relying on augmented reality.

In terms of reliability, latency, and throughput, RIS significantly bolsters the performance of end-to-end trained communication models. These improvements are crucial in sectors such as smart manufacturing and urban traffic management, where robust communication networks are essential for maintaining continuous and reliable operations. The ability of RIS to tailor signal processing to the specific requirements of these applications enhances overall system performance and operational efficiency.

Additionally, the adaptability afforded by RIS technology facilitates the development of sophisticated feedback mechanisms. These systems can dynamically adjust to real-time environmental changes and communication demands, ensuring that the network remains robust and effective under diverse

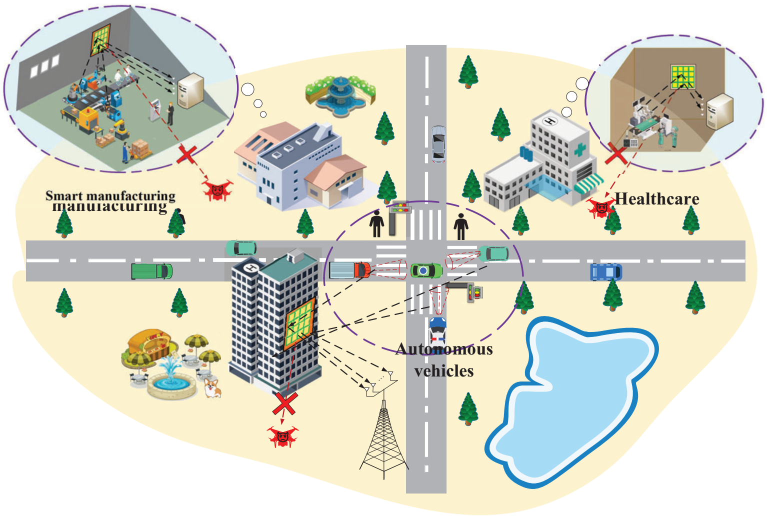


Fig. 1. The applications of RIS in smart manufacturing, healthcare, and autonomous vehicles for supporting and securing task-oriented communication.

conditions. This adaptability is particularly beneficial in dynamic environments where communication needs can fluctuate rapidly.

III. Challenges and Possible Solutions

The integration of RIS into AI-driven vertical applications offers numerous opportunities while also posing significant challenges. This section elucidates the major research hurdles and prospects associated with advancing RIS-assisted task-oriented communications.

A. Problems to be Solved

Addressing the integration of RIS into diverse AI-driven applications necessitates overcoming several formidable challenges. A primary issue is the real-time and dynamic optimization of RIS configurations, considering the varying requirements of different tasks in terms of latency, bandwidth, and reliability. Moreover, tailoring signal processing techniques specific to tasks facilitated by RIS demands an in-depth understanding of the task requirements and the capabilities of RIS. Such tasks present scalability challenges, particularly in expansive deployments such as smart cities or industrial complexes.

Further, the development of feedback mechanisms capable of effectively utilizing RIS to adapt communication strategies is crucial. This requires not only sophisticated algorithmic innovation but also capabilities for real-time data analysis. The challenge is ensuring these systems remain responsive and efficient under a range of conditions, which may include diverse and potentially adverse environmental factors.

The variability of these environmental factors also complicates the robustness of end-to-end trained systems that integrate RIS, necessitating extensive simulation and empirical testing to ensure efficacy across various scenarios. Additionally, different AI applications demand specific signal characteristics, which RIS can tailor to enhance signal propagation paths and conditions. This customization is vital in settings such as smart factories or smart grids, where optimal operations depend on precise signal attributes.

End-to-end training aims to optimize the entire communication chain for specific tasks. However, such systems are particularly sensitive to the conditions of the physical layer. Improving the adaptability of the physical layer through RIS can significantly enhance the overall performance under diverse operational conditions.

Moreover, as the deployment of end-to-end task-oriented grows, security and privacy emerge as critical concerns, especially in sensitive industrial environments. References such as [10] and [11] highlight the need for robust security mechanisms that protect against potential threats and vulnerabilities, ensuring the confidentiality and integrity of transmitted data. Addressing these security challenges is essential for the broader acceptance and implementation of RIS in task-oriented communications.

B. Possible Solutions

In light of the formidable challenges identified in the integration of RIS into AI-driven vertical applications, a series of innovative solutions can be articulated to address these issues. Each solution targets specific aspects of the challenges to enhance the efficacy and robustness of remote task inference systems in vertical applications.

To tackle the challenge of dynamic and real-time optimization of RIS configurations, the development of intelligent control algorithms is paramount. These algorithms can adaptively modify RIS settings in response to fluctuating network conditions and application demands. Utilizing advanced machine learning techniques, such as reinforcement learning [12], these algorithms can learn optimal configurations from historical data and real-time feedback, thus ensuring that the RIS always operates at peak efficiency.

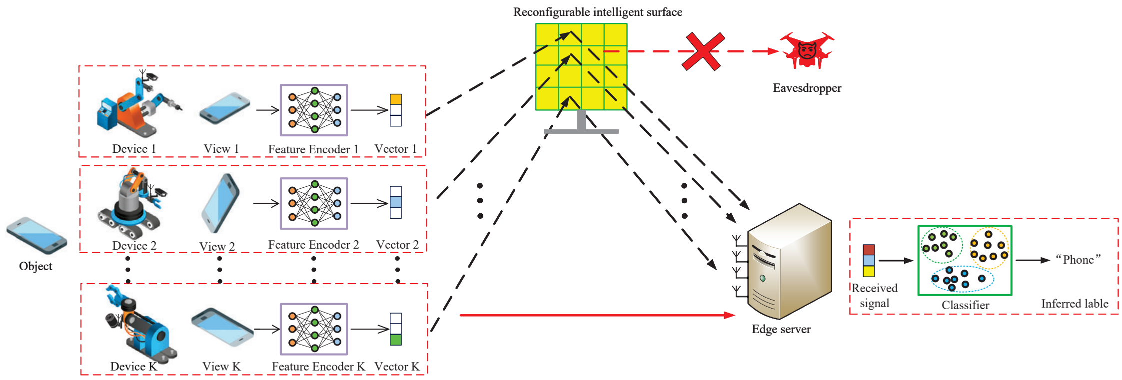


Fig. 2. RIS-assisted multi-device classification task-oriented communication in smart manufacturing, targeting high-classification accuracy and security.

Addressing the scalability issues in extensive deployments, such as in smart cities or industrial complexes, requires the implementation of scalable architecture for RIS management. This architecture should support hierarchical control mechanisms that can manage multiple RIS panels efficiently, distributing the computational load and reducing the latency associated with configuration adjustments.

For the tailored signal processing required by specific AI-driven applications, developing modular signal processing units that can be customized for various applications can be a viable solution. These units would use programmable logic to adapt their processing strategies based on the input from the AI systems and the feedback received from the environment, thus optimizing the signal characteristics for each specific task.

To improve the responsiveness and efficiency of communication systems under diverse and adverse conditions, integrating predictive analytics into the feedback mechanisms is essential. These analytics can anticipate potential disruptions based on environmental data and adjust the RIS configurations preemptively to mitigate any impact on communication quality.

Furthermore, rigorous testing protocols need to be established to assess the robustness of end-to-end trained systems that incorporate RIS. Simulation tools that can model complex environments and network conditions will play a crucial role in this process, helping to identify weaknesses before deployment and allowing for the refinement of systems.

Security and privacy also remain critical as RIS-assisted task-oriented communication system deployments expand. Developing dedicated security protocols for RIS-enabled networks can help safeguard against breaches and ensure data integrity. These protocols should include advanced encryption methods and secure authentication processes to protect data transmission across potentially vulnerable networks [10], [11].

Finally, fostering a multidisciplinary collaboration among experts in wireless communications, cybersecurity, AI, and domain-specific applications will be crucial. Such collaboration can ensure that the solutions are not only technically feasible but also aligned with practical requirements and industry standards, facilitating smoother integration and adoption.

By implementing these solutions, the potential of RIS in revolutionizing communication systems across various AI-driven verticals can be fully realized, leading to more adaptable, efficient, and secure task-oriented networks.

C. *Potential Applications*

RIS-assisted task-oriented communications offer a wide range of potential applications across various industries. These applications leverage RIS technology to optimize communication systems for specific tasks, thereby enhancing efficiency, reliability, and performance.

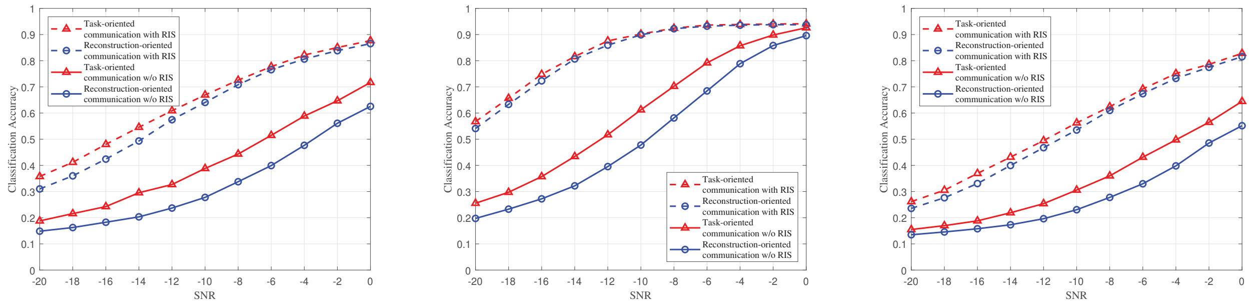
One prominent application area is autonomous vehicles, where RIS-assisted communications can play a crucial role in optimizing communication between vehicles and infrastructure components. By dynamically adjusting signal reflections and steering, RIS can facilitate reliable and low-latency cooperative inference between autonomous vehicles, enabling seamless coordination and cooperation in complex traffic scenarios.

In smart manufacturing environments, RIS-assisted communications can enhance coordination and control of robotic systems and machinery. By optimizing signal propagation and coverage, RIS can improve communication reliability and efficiency, leading to increased productivity and throughput in manufacturing processes.

Another promising application area is healthcare systems, where RIS-assisted communications can facilitate real-time monitoring and decision-making in clinical environments. By ensuring reliable and low-latency task-related communications between medical devices and systems, RIS can enable remote diagnostics, telemedicine services, and patient monitoring, ultimately improving patient care outcomes. Furthermore, task-specific RIS technology can be applied in various other industrial scenarios, such as energy management, logistics, and infrastructure monitoring.

IV. Experiment and Discussion

In this section, we detail the experimental outcomes of our study on RIS-assisted task-oriented communication strategies, utilizing real-world datasets and simulated wireless environments. Our investigation employs a case study approach,



(a) 1-NN. (b) MAP. (c) NN.

Fig. 3. Evaluation of the RIS on classification performance using three different classifiers: 1-NN, MAP, and NN. The graphs display classification accuracy across a range of SNR from -20 dB to 0 dB, comparing task-oriented communication with and without RIS against reconstruction-oriented communication with and without RIS.

focusing on multi-view cameras used for remote classification or recognition tasks. As depicted in Fig. 2, cameras mounted on robots within a smart manufacturing setting capture images from multiple angles to facilitate object detection. To mitigate the adverse effects of challenging wireless conditions typical in industrial environments, a RIS is deployed to enhance remote multi-device cooperative inference. Additionally, we address security concerns by positing the presence of an eavesdropper capable of performing task inference. In our experimental setup, it is assumed that the transceivers and eavesdropper possesses complete knowledge of the training dataset, labels, and communication channels. Although in practical scenarios the channel information for the eavesdropper may remain undisclosed, our analysis assumes full information availability to ascertain the efficacy of security measures under controlled conditions.

A. Datasets and Neural Networks Structures

To evaluate the practical efficacy of our proposed methodologies, we undertook classification tasks leveraging the ModelNet10 dataset [13] to facilitate multi-device edge inference. The ModelNet10 dataset is composed of ten categories of computer-aided design (CAD) objects, including items such as sofas, bathtubs, and beds. Each category is represented through twelve distinct views per sample. Notably, there exists a disparity in the distribution of training samples across categories, varying from 106 to 889, whereas the test samples range from 50 to 100 in each category. We employed the pretrained ResNet-18 [14] architecture as the foundational feature encoder for each participating device.

B. Communication System Settings

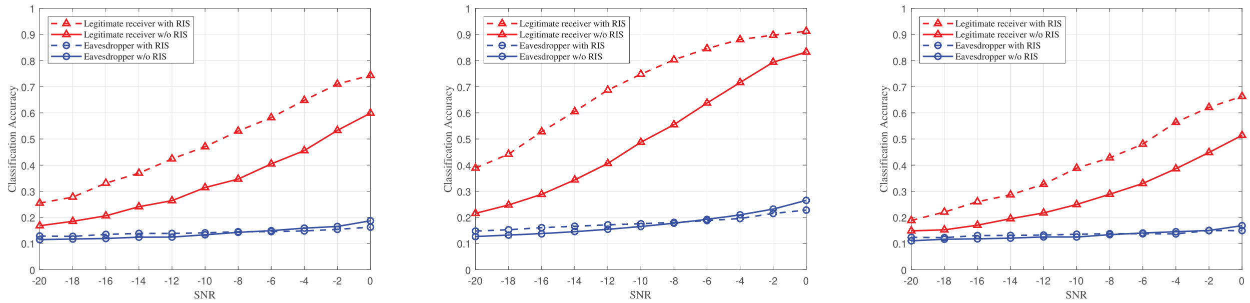
In our experimental setup, we consider a scenario involving three devices. For the sake of generalization, we model the communication channel between each device and the edge server as a Rayleigh fading channel, with the signal-to-noise ratio (SNR) varying between -20 dB and 0 dB. Regarding the data dimensionality, we set the feature vector length to 12 for the ModelNet10 dataset, with each device acquiring 4 features. Each device is equipped with 4 transmit antennas, while the edge servers and potential eavesdroppers possess 4 receive antennas each. The number of RIS units is 4. To underscore the efficacy of RIS-assisted task-oriented communication frameworks, we evaluate various classifiers including the maximum a posteriori probability (MAP) classifier, the nearest neighbor (1-NN) classifier, and the neural network (NN) classifier.

C. Precoding and RIS-Configurations

Diverging from conventional precoding and reconfigurable intelligent surface (RIS) configurations that focus on reconstruction-oriented communications, we refine these configurations by implementing a maximum coding rate reduction (MCR^2) criterion [9], [15]. This criterion aligns with the objectives of classification tasks, aiming to maximize the inter-class sample distances while minimizing the intra-class distances. In contrast, the conventional approach utilizes precoding and RIS configurations to minimize reconstruction loss, primarily targeting the minimization of the mean-square error (MMSE).

D. Experiment Results and Discussion

In Fig. 4, we compared the impact of RIS on classification accuracy using three different classifiers. The experimental results fully demonstrate the significant advantages of task-oriented communication and RIS technology in improving classification accuracy. Specifically, under three different classifiers (1-NN, MAP, and NN), the task-oriented communication showed higher classification accuracy compared to the reconstruction communication method, especially in low SNR environments. This is because the task-oriented communication method aligns the targets of learning and communication, thereby improving the precise execution of the task. Moreover, with the assistance of RIS, the performance of all three classifiers has been further improved. Taking the MAP classifier in Fig. 4(b) as an example, in the low SNR range of -20dB to 0dB, the reconstrution method with RIS assistance can improve classification accuracy from about 20% to about 55%, while the task-oriented communication with RIS assistance can improve accuracy from about 25% to nearly



(a) 1-NN. (b) MAP. (c) NN.

Fig. 4. Impact of RIS on classification accuracy, analyzed using three distinct classifiers: 1-NN, MAP, and NN. This analysis presents classification accuracy for legitimate receivers and eavesdroppers across a SNR spectrum from -20 dB to 0 dB, illustrating the effects of implementing RIS versus no RIS conditions..

58%. This indicates that RIS technology can effectively greatly improve the perception ability of remote inference systems in complex environments. Overall, RIS-assisted task-oriented communication can significantly improve classification accuracy under low SNR conditions, thereby greatly enhancing the performance of wireless communication systems in intelligent applications. This provides strong technical support for wireless communication in fields such as intelligent manufacturing and healthcare in the future.

The exploration results of the impact of RIS on the classification accuracy of the legitimate receiver and eavesdropper under different classifiers are demonstrated in Fig. 5. The results clearly show that with the assistance of RIS, the classification accuracy of legitimate receivers is significantly improved. This means that the introduction of RIS technology effectively improves the receiver's perception ability of signals, enabling them to more accurately distinguish different types of classes. It is worth mentioning that the classification accuracy of eavesdroppers with and without RIS assistance is almost the same, which means that the application of RIS technology has not increased the eavesdropper's ability to identify communication content, effectively ensuring communication security.

V. Conclusion

RIS-assisted task-oriented communications hold significant potential for revolutionizing communication systems in AI-driven vertical applications. By integrating RIS with AI techniques, it becomes possible to optimize communication performance and efficiency to meet the specific requirements of different tasks. Future research in this field offers numerous opportunities for innovation and advancement, with potential applications spanning various industries. Addressing key research challenges and exploring new directions will be crucial for realizing the full potential of RIS-assisted communications and advancing the field of wireless communications in AI-driven vertical applications.

References

[1] S. Guo, B. Lu, M. Wen, S. Dang, and N. Saeed, "Customized 5G and beyond private networks with integrated URLLC, eMBB, mMTC, and positioning for industrial verticals," *IEEE Commun. Stand. Mag.*, vol. 6, no. 1, pp. 52–57, Mar. 2022.

[2] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[3] S. Zeng, H. Zhang, B. Di, Z. Han, and L. Song, "Reconfigurable intelligent surface (RIS) assisted wireless coverage extension: RIS orientation and location optimization," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 269–273, Jan. 2020.

[4] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[5] J. Ye, S. Guo, and M.-S. Alouini, "Joint reflecting and precoding designs for SER minimization in reconfigurable intelligent surfaces assisted MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5561–5574, Aug. 2020.

[6] J. Zhao, J. Ye, S. Guo, Z. Bai, D. Zhou, and A. Mohamed, "Reconfigurable intelligent surface enabled joint backscattering and communication," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 771–782, Jan. 2024.

[7] J. Shao, Y. Mao, and J. Zhang, "Learning task-oriented communication for edge inference: An information bottleneck approach," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 197–211, Jan. 2021.

[8] D. Gündüz, Z. Qin, I. E. Aguerri, H. S. Dhillon, Z. Yang, A. Yener, K. K. Wong, and C.-B. Chae, "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5–41, Jan. 2022.

[9] C. Cai, X. Yuan, and Y.-J. A. Zhang, "Multi-device task-oriented communication via maximal coding rate reduction," *arXiv preprint arXiv:2309.02888*, 2023.

[10] 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.*, pp. 1–1, 2024.

[11] A. Zhang, Y. Wang, and S. Guo, "On the utility-informativeness-security trade-off in discrete task-oriented semantic communication," *IEEE Commun. Lett.*, pp. 1–1, 2024.

[12] W. Wang and W. Zhang, "Intelligent reflecting surface configurations for smart radio using deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2335–2346, Aug. 2022.

[13] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, "3d shapenets: A deep representation for volumetric shapes," in *Proceed. IEEE CVPR*, Boston, Massachusetts, USA, June 2015, pp. 1912–1920.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceed. IEEE CVPR*, Las Vegas, Nevada, USA, 2016, pp. 770–778.

[15] Y. Yu, K. H. R. Chan, C. You, C. Song, and Y. Ma, "Learning diverse and discriminative representations via the principle of maximal coding rate reduction," *Proc. NIPS*, vol. 33, pp. 9422–9434, Dec. 2020.