Advanced Learning Algorithms for Integrated Sensing and Communication (ISAC) Systems in 6G and Beyond: A Comprehensive Survey

Nguyen Cong Luong, Thien Huynh-The, *Senior Member, IEEE*, Thai-Hoc Vu, *Member, IEEE*,
Duc Van Le, *Senior Member, IEEE*, Huy T. Nguyen, *Member, IEEE*,
Nguyen Duc Hai, *Student Member, IEEE*, Gia-Vuong Nguyen, Nguyen Duc Duy Anh,
Dusit Niyato, *Fellow, IEEE*, Dong In Kim, *Life Fellow, IEEE*, and Quoc-Viet Pham, *Senior Member, IEEE*

***Abstract*—Integrated sensing and communication (ISAC) allows the same hardware platform and resources to function sensing and communication simultaneously, which reduces the hardware size and addresses the spectrum congestion concerns. However, the sharing of the hardware and resources of the sensing and communication functions raises resource managements. Traditional optimization approaches are developed based on rigid mathematical models in ISAC systems. However, they face computation complexity and may not achieve the desired performance under the dynamics of the ISAC system environments. Machine learning with ability in learning features/patterns of data as well as approximating mathematical models has recently proposed to effectively solve the complicated ISAC problems. In this survey, we thus provide a comprehensive literature review on applications of learning algorithms for ISAC systems. Particularly, we review learning approaches proposed for emerging issues in ISAC systems, including beamforming designing/tracking, waveform design, spectrum allocation, time allocation, and power allocation, angle of arrival (AoA)/angle of departure (AoD) estimation, signal classification, and security issues. Moreover, we present applications of advanced learning methods for wireless sensing, which is considered to be an emerging sensing service of the next-generation networks. We conclude the survey with highlighting technical issues of learning algorithms and discussing future research directions.**

***Index Terms*—Integrated sensing and communication, machine learning, resource management, waveform design, wireless sensing.**

Received 29 December 2024; revised 23 May 2025; accepted 23 June 2025. Date of publication 1 July 2025; date of current version 2 January 2026. This work was supported in part by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIT) under Grant 2021R1A2C2007638; in part by the Ministry of Science and ICT (MSIT), South Korea, through the ICT Creative Consilience Program supervised by the Institute for ICT Planning and Evaluation (IITP) under Grant IITP-2020-0-01821; in part by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant 102.02-2023.43; in part by Singapore Ministry of Education (MOE) Tier 1 under Grant RG87/22 and RG24/24; in part by the NTU Centre for Computational Technologies in Finance (NTU-CCTF), and the RIE2025 Industry Alignment Fund—Industry Collaboration Projects (IAF-ICP) administered by A*STAR under Award I2301E0026; and in part by Alibaba Group and NTU Singapore through Alibaba-NTU Global e-Sustainability CorpLab (ANGEL). *(Corresponding author: Dong In Kim.)*

Nguyen Cong Luong, Nguyen Duc Hai, and Nguyen Duc Duy Anh are with the Phenikaa School of Computing, Phenikaa University, Hanoi 12116, Vietnam (e-mail: luong.nguyencong@phenikaa-uni.edu.vn; 21010560@st.phenikaa-uni.edu.vn; 21011488@st.phenikaa-uni.edu.vn).

Thien Huynh-The is with the Department of Computer and Communications Engineering, Ho Chi Minh City University of Education and Technology, Ho Chi Minh City 71307, Vietnam (e-mail: thienht@hcmute.edu.vn).

Thai-Hoc Vu is with the Institute of Information Technology, Digital Transformation, Thu Dau Mot University, Binh Duong 820000, Vietnam (e-mail: hocvt@tdmu.edu.vn).

Duc Van Le is with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW 2052, Australia (e-mail: duc.le1@unsw.edu.au).

Huy T. Nguyen is with the Smart and Autonomous Systems Research Group, Faculty of Information Technology, School of Technology, Van Lang University, Ho Chi Minh City, 70000, Vietnam (e-mail: huy.nt@vlu.edu.vn).

Gia-Vuong Nguyen is with the Department of Artificial Intelligence and Robotics, and the Department of Convergence Engineering for Intelligent Drone, Sejong University, Seoul, 05006, South Korea (e-mail: vuongng.cce@gmail.com).

Dusit Niyato is with the College of Computing and Data Science, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg).

Dong In Kim is with the Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon 16419, South Korea (e-mail: dongin@skku.edu).

Quoc-Viet Pham is with the School of Computer Science and Statistics, Trinity College Dublin, Dublin 2, D02 PN40 Ireland (e-mail: viet.pham@tcd.ie).

Digital Object Identifier 10.1109/COMST.2025.3584333

I. INTRODUCTION

INTEGRATING sensing and communication (ISAC) [1], [2], [3] has emerged as a key pillar for 6G and beyond to support various emerging applications such as Internet-of-Things (IoT) networks [4], [5], vehicle-to-everything (V2X) communications [6], [7], [8], connected autonomous systems [9], human activity sensing [10], and unmanned aerial vehicle (UAV) networks [11], [12], [13]. For example, future V2X communications are expected to provide accurate localization service while achieving high data rate transmission under high mobility environments [6] or human activity sensing allows a base station or an access point to use wireless signals for human activity recognition [14] or health monitoring [15]. The key advantage of the ISAC is to allow using the same hardware, i.e., digital processing platform and antenna, and resources, i.e., spectrum and power resource, to execute the sensing function and communication function simultaneously. The ISAC thus addresses the low hardware size and low-cost requirements of future devices as well as spectrum congestion concerns of the future wireless systems [16]. As a result, ISAC has been proposed for several areas as illustrated in Fig. 1.

However, the sharing of hardware and resources of the communication function and sensing function as well as the dynamics of the ISAC system environments impose resource management issues including waveform design, beamforming

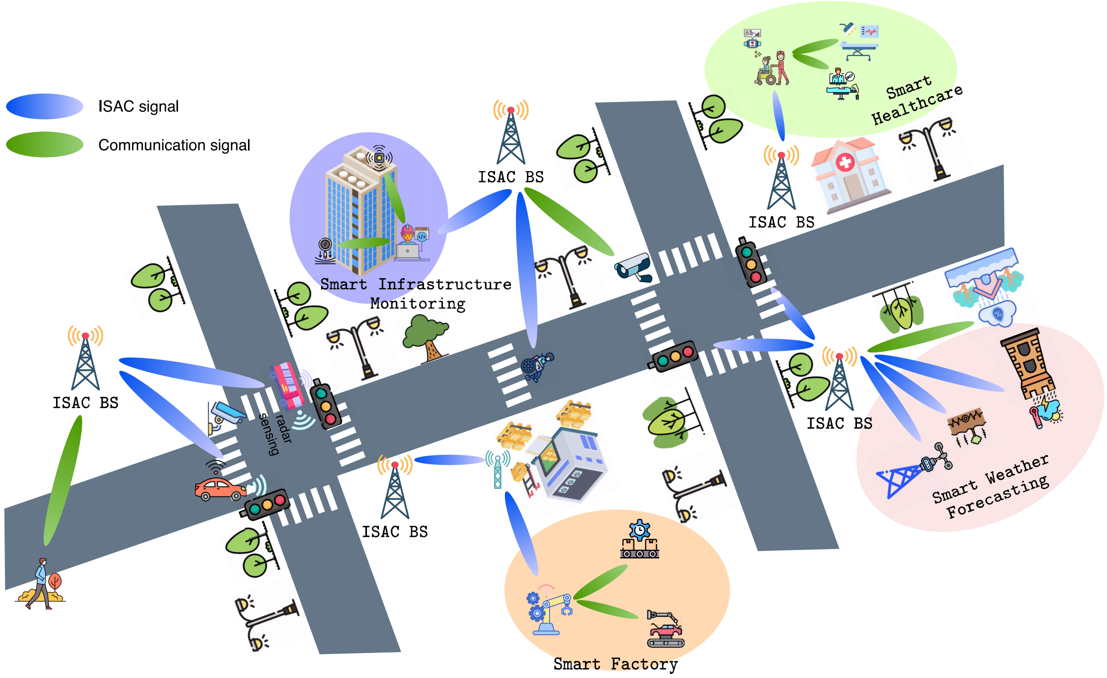


Fig. 1. An illustration of how ISAC technology serves as a foundational element for 6G networks and leverages integrated sensing and communication to create intelligent systems across various domains. From autonomous vehicles and smart infrastructure monitoring to precision healthcare and disaster response, ISAC empowers applications with real-time data and adaptive capabilities.

design, spectrum allocation, time allocation, and power allocation. Traditional optimization approaches such as successive convex approximation (SCA) [17] and heuristics [18] have been proposed to solve the issues. However, they are often computation-intensive since the ISAC system optimization problems typically consist of multiple coupled variables, non-convex objective functions, and constraints. Moreover, these traditional approaches are mostly developed based on static and rigid mathematical models, while the real ISAC system environments are uncertain and dynamic [19], [20]. Therefore, these approaches incur high computational complexity and do not always achieve the desired performance. Novel approaches are needed to address these issues and to enable improved ISAC system performance in real-world deployments.

Recently, learning algorithms based on machine learning (ML) techniques have been proposed to solve highly complicated problems in different areas. Compared to the traditional optimization approaches, learning algorithms have major advantages as follows. First, they can transfer complex optimization tasks into the training phase. Then, the trained models can be deployed to adaptively provide the real-time resource allocation decisions at run time [21]. This addresses the computation challenges and dynamic issues of the traditional approaches. Second, learning algorithms own generalization properties [22], which help provide resource allocation solutions properly even in unknown/uncertain system conditions [23]. Third, learning algorithms effectively learn features lying within the data [23], [24], [25], which can be used to learn the structure/pattern of the complex channels and helps to reduce the complexity of the beamforming and waveform designs. Finally, learning algorithms can capture complex data representations [26], [27], [28], which facilitates signal classification as well as parameter estimation. Given the advantages, learning algorithms have recently been proposed to solve various issues in ISAC systems including *(i)* beamforming designing/tracking, *(ii)* waveform design, *(iii)* spectrum allocation, time allocation, and power allocation, *(iv)* angle of arrival (AoA)/angle of departure (AoD) estimation and signal classification, *(v)* security issues, and *(vi)* wireless sensing. This requires a comprehensive survey on learning algorithms for ISAC systems.

There are several magazines and surveys on ISAC. However, they do not provide a comprehensive survey on applications of learning algorithms for ISAC. For example, fundamental limits of the ISAC systems are given in [1], [2], [29]. The work in [30] provides a comprehensive survey on the background and performance trade-offs between sensing and communication in ISAC systems. Meanwhile, some surveys on waveform design, i.e., transmit signal design, for ISAC systems are found in [31] and [29]. A more comprehensive survey on waveform design, antenna array design, and clutter suppression for ISAC systems is found in [32]. However, the issues of beamforming design and resource management as well as the use of learning algorithms are not considered for ISAC systems. A survey on channel models in ISAC systems is found in [33]. A recent work in [34] provides a comprehensive survey on propagation characterization and channel modeling in THz-based ISAC systems. The discussions of the THz-based ISAC systems are also provided in [35]. Two surveys on resource management in ISAC systems are found in [19] and [36]. However, there are few learning algorithm approaches discussed in the survey. Otherwise, there are existing surveys on learning algorithms. However, they are not specific to ISAC. For example, the work in [21] provides a comprehensive survey on applications of reinforcement learning algorithms for wireless systems. A survey on learning, i.e., deep learning and deep reinforcement learning (DRL) algorithms, to robotics, is given in [37]. Some recent surveys have discussed potential applications of potential roles of learning algorithms for ISAC such as [23], [38], and [39]. However, these works lack a comprehensive survey on recent learning approaches for ISAC. Table I highlights the differences between our survey and existing surveys. As seen, this is a unique work that provides a comprehensive survey on applications of learning algorithms for ISAC.

This motivates us to provide a comprehensive review of the applications of learning algorithms to address issues in ISAC systems. The contributions of this survey are as follows:

- We review and discuss a number of learning approaches proposed for the ISAC transmitter design, including beamforming optimization and tracking, waveform design, spectrum allocation, time allocation, and power allocation.
- We review and discuss learning approaches for the ISAC receiver design, including AoA/AoD estimation and signal classification.
- We review and discuss a number of advanced learning algorithms that have been recently used for wireless sensing, which is an emerging service leveraging communication signals of Wi-Fi or cellular networks for target sensing. Discussions on using learning algorithms for recent issues, e.g., security, joint sensing and computation offloading, and joint sensing and federated learning, are also provided.
- We provide comparisons among the existing approaches and highlight their shortcomings.
- We highlight technical issues of implementing the learning algorithms and discuss promising research directions related to ISAC.

The remainder of the paper is organized as follows. Section II introduces performance metrics and challenges of ISAC. Section III reviews the applications of learning algorithms for beamforming design and tracking. Section IV discusses learning algorithms for waveform design. Section V provides reviews of learning algorithms for AoA/AoD estimation and signal classification. Section VI reviews learning algorithms for resource allocation. Section VII discusses learning algorithms for wireless sensing. Section VIII discusses the applications of learning algorithms for other emerging issues. Section IX concludes this paper, highlights key challenges, and future work.

Fig. 2 shows the organization of this survey. Fig. 3 shows the percentages of related works for different issues of ISAC. As observed, learning approaches for the beamforming design have the highest percentages, followed by the waveform design and resource allocation. The reason may be due to the sharing of resource between the sensing and communication. Thus, more works are investigated for the beamforming design and resource allocation to address the trade-off between the sensing and communication performance. Abbreviations frequently used in this work are listed in Table II. It is noted that fundamentals with key concepts and performance metrics of ISAC can be referred to recent works such as [1], [2], [29], and [30]. It is further noted that the terms of "sensing" and "radar" can be used interchangeably due to the fact that they mostly work the same in principle. Thus, ISAC has the different names as joint communication and radar sensing (JCAS) [32], [45], [46], joint radar and communications (JRC) [47], [48], [49], dual-functional radar communications (DFRC) [50], [51], [52].

II. INTEGRATED SENSING AND COMMUNICATION: PERFORMANCE METRICS AND CHALLENGES

In this section, we present performance metrics commonly used in ISAC systems for resource management design. Then, we discuss challenges including resource trade-offs, dynamic channel modeling, interference management, and security and privacy concerns-paving the way for innovative solutions in ISAC implementation.

A. Performance Metrics for ISAC

Performance metrics are used to evaluate Quality-of-Services (QoSs) of the sensing and communication functions of ISAC. ISAC systems typically use the following

1) Beampattern: The beampattern of an ISAC system determines the spatial distribution of its signal power, which typically is used for the target sensing [53]. Beampattern design often involves approximating a desired pattern using least squares. For a transmitter with N_t uniform linear array (ULA) antennas, the mean squared error (MSE) for beampattern approximation is

$$\mathrm{MSE} = \sum_{k=1}^{K} \left| \mathfrak{B}(\theta_k) - \mathsf{a}^{\mathrm{H}}(\theta_k)\mathbf{R}_{\mathrm{x}}\mathsf{a}(\theta_k) \right|^2, \tag{1}$$

where $\mathfrak{B}(\theta_k)$ is the desired beampattern level at the k-th azimuth angle grid among all K grids, $(\cdot)^{\mathrm{H}}$ denotes the conjugate-transpose operation, $\mathbf{R}_{\mathrm{x}}$ is the covariance matrix of transmit waveforms [53], and $\mathbf{a}(\theta_k)$ is the transmit steering vector defined as

$$\mathsf{a}(\theta_k) = \left[1, e^{j2\pi\Delta\sin(\theta_k)}, \ldots, e^{j2\pi(N_t-1)\Delta\sin(\theta_k)}\right] \tag{2}$$

where Δ is the normalized distance between adjacent array elements. It can be observed that minimizing the beampattern allows sensing the target with high accuracy.

2) Signal-to-Cluster-Plus-Noise Ratio (SCNR): In cluttered environments, SCNR often is used for the sensing function that measures the power ratio of the desired signal to combined noise and clutter [53]. A higher value of SCNR implies the higher sensing QoS. Mathematically, it is expressed by

$$\mathrm{SCNR} = \frac{\left\|\alpha_{\mathrm{rcs}}\mathsf{a}_{\mathrm{rcs}}^{\mathrm{H}}(\mathrm{x})\mathbf{B}\right\|^2}{\sum_{c=1}^{C}\left\|\alpha_c\mathsf{a}_c^{\mathrm{H}}(\mathrm{x})\mathbf{B}\right\|^2 + \sigma^2}, \tag{3}$$

where α_c and α_{rcs} are the complex coefficients involving the cascaded complex gains of the target/clutter c and the radar cross-section of the target/clutter c, respectively; $\mathbf{a}_c^{\mathrm{H}}(\mathrm{x})$ is the array response vector between the base station and clutter c, while $\mathbf{a}_{\mathrm{rcs}}^{\mathrm{H}}(\mathrm{x})$ is the array response vector between the base station and the target; $\mathbf{B}$ denotes the transmit beamforming matrix (or vector in single-stream cases); and σ^2 is the noise variance of radar link. SCNR optimization employs methods

TABLE I
SUMMARY OF THE RELATED SURVEYS / ARTICLES

Existing surveys	**Fundamentals of ISAC**	**Beamforming design**	**Waveform design**	**Channel modeling and estimatio**	**Resource allocation**	**AoA/AoD estimation**	**Wireless sensing**	**Security issues**	**Learning approaches**
[32]	✓	✓	✓						
[19]	✓				✓				
[30]	✓	✓	✓		✓				
[31]	✓		✓						
[36]	✓	✓	✓		✓				
[29]	✓		✓						
[23]		✓			✓				✓
[40]	✓	✓							✓
[41]	✓	✓	✓		✓				
[42]	✓	✓	✓	✓		✓	✓		
[39]	✓	✓	✓	✓			✓		
[43]								✓	
[44]	✓							✓	
[33]				✓					
[34]	✓		✓	✓					
[38]		✓						✓	
Ours	✓	✓	✓	✓	✓	✓	✓	✓	✓

like null space design, eigenbeamforming, or clutter-plus-noise covariance-based techniques to enhance radar detection.

3) Signal-to-Interference-Plus-Noise Ratio (SINR): SINR is a common metric used to evaluate the communication function, which is defined by

$$\mathrm{SINR} = \frac{\left|\mathrm{h}_m^{\mathrm{H}}\mathrm{f}_m\right|^2}{\sum_{i=1,i\neq m}^{M+1}\left|\mathrm{h}_m^{\mathrm{H}}\mathrm{f}_i\right|^2+\sigma^2}, \tag{4}$$

where $\mathbf{h}_m^{\mathrm{H}}$ denotes the channel between the transmitter and m-th user, $\mathbf{f}_m$ is the transmit beamforming associated with m-th user. SINR can be maximized by optimizing beamforming vector $\mathbf{f}_m$ with some common methods, such as zero-forcing beamforming, minimum mean-squared error beamforming, and interference alignment.

4) Cramer-Rao Lower Bound (CRLB): The CRLB defines the theoretical accuracy limit for parameter estimation in sensing tasks (e.g., localization, velocity, AoA estimation) [54]. It provides a lower bound on the variance of unbiased estimators. For a parameter ϱ, the variance of any unbiased estimator $\hat{\varrho}$ satisfies:

$$\mathrm{Var}(\varrho) \geq \mathrm{J}^{-1}(\varrho), \tag{5}$$

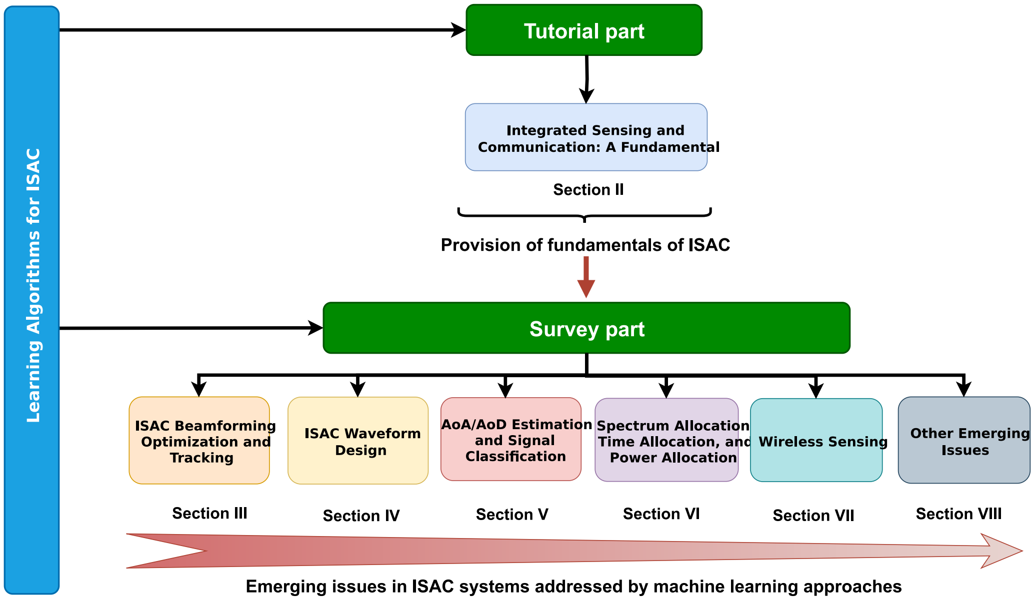


Fig. 2. Structure of the survey.

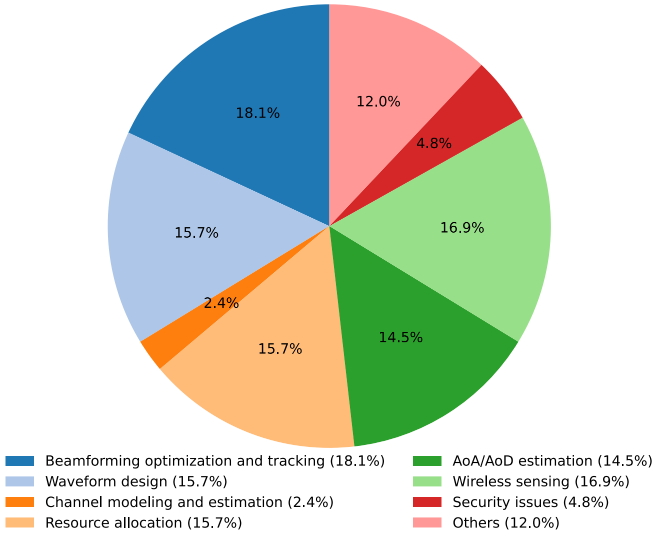


Fig. 3. Percentages of related works for different issues of ISAC.

where $\mathbf{J}(\varrho)$ is the Fisher information matrix (FIM) [3], which is a matrix that quantifies the amount of information about ϱ contained in the observation y. It is noted that a lower bound on the variance means that the sensing function cannot achieve better estimation accuracy than the CRLB under ideal conditions. The FIM can be expressed as follows:

$$\mathbf{J}(\varrho) = \mathrm{E}\left[\nabla_{\varrho} \log \mathrm{p}(y|\varrho) * \left(\nabla_{\varrho} \log \mathrm{p}(y|\varrho)\right)^{T}\right], \tag{6}$$

where $\mathrm{E}[\dots]$ is the expectation operator, ∇_{ϱ} is the gradient operator with respect to ϱ, and $\mathrm{p}(y|\varrho)$ denotes the probability density function of the received signal y given the true parameter ϱ. The CRLB bounds the variance of any unbiased estimator of ϱ to ensure that it cannot be smaller than the corresponding diagonal element of the inverse FIM. Here, ϱ may represent the range, Doppler, or AoA that need to be estimated.

TABLE II
LIST OF ABBREVIATIONS FREQUENTLY USED IN THIS PAPER

Abbreviation	Description
AI	Artificial Intelligence
AoA/AoD	Angle of Arrival/Angle of Departure
ASRIS	Active Reconfigurable Intelligent Surfaces
CFAR	Constant False Alarm Rate
CNN(s)	Convolutional Neural Network(s)
CRLB	Cramer-Rao Lower Bound
DL	Deep Learning
DDQN	Double Deep Q-Network
DFRC	Dual-Function Radar-Communication
DQN	Deep Q-Network
GCN(s)	Graph Convolutional Network(s)
HBF	Hybrid Beamforming
ISAC	Integrated Sensing and Communication
LSTM	Long Short-Term Memory
ML	Machine Learning
MADDPG	Multi-Agent Deep Deterministic Policy Gradient
MD-STAR	Multi-Dual Simultaneously Transmitting and Reflecting RISs
MDP	Markov Decision Process
MI	Mutual Information
MRL	Meta Reinforcement Learning
MSE	Mean-Squared Error
OTFS	Orthogonal Time Frequency Space
PGA	Projected Gradient Ascent
PPO	Proximal Policy Optimization
RGB	Red Green black (color model)
SAC	Soft Actor-Critic
SCA	Successive Convex Approximation Algorithm
SINR	Signal-to-interference-plus-noise ratio
STAR-RIS	Simultaneously Transmitting and Reflecting Intelligent Surface
UAV	Unmanned Aerial Vehicle

B. Challenges in ISAC for 6G

Due to sharing the hardware and resources, ISAC systems faces numerous challenges as follows.

1) Sensing and Communication Trade-Off: ISAC systems must balance limited resources, such as bandwidth and power, between sensing and communication tasks. Allocating more resources to one often degrades the other, particularly in applications like autonomous driving and smart manufacturing, where both functions are critical [55]. Static, non-adaptive methods are ineffective in optimizing the trade-off dynamically in changing environments [2]. Learning algorithms enable real-time resource allocation and beamforming optimization by learning from interactions to optimize both sensing and communication. Neural networks predict environmental factors like interference levels, dynamically adjusting resource usage to maintain quality for both tasks. This adaptability ensures efficient use of resources across varying conditions.

2) Dynamic Channel Modeling and Prediction: Highly dynamic environments, such as urban areas with fast-moving vehicles, challenge ISAC systems in maintaining reliable sensing and communication. Traditional channel models are insufficient for millimeter-wave and terahertz bands, which are sensitive to obstacles and atmospheric changes [56]. Accurate, real-time channel modeling and prediction are crucial for optimizing resource allocation and beamforming [38]. Learning-based approaches, including recurrent neural networks (RNNs) and graph neural networks (GNNs), analyze large datasets to identify non-linear patterns in channel behavior. These models predict channel state information (CSI) and adapt system parameters to varying conditions.

3) Interference Management and Signal Decomposition: ISAC systems operating in shared spectrum environments face significant interference, especially in dense urban or industrial settings [57]. Traditional signal processing techniques often fail to differentiate overlapping signals in real-time, degrading both sensing and communication performance. Deep learning (DL) techniques, including convolutional neural networks (CNNs) and autoencoders, are highly effective to process raw signals separately, while autoencoders can extract relevant features from noisy data, decomposing signals into distinct components. Generative adversarial networks (GANs) effectively denoise signals, modeling interference statistics and removing unwanted components. Moreover, RL dynamically adjusts system parameters like frequency and power, hence ensuring robust interference management.

4) Privacy and Security in Shared Data: ISAC systems allow the sensing function and communication function to share signals for the target sensing and data communication. Also, the ISAC systems process sensitive data (e.g., location or health metrics), making them vulnerable to unauthorized access and various forms of cyberattacks, including those by eavesdroppers and jammers [58]. The dynamic and uncertain behavior of attackers, coupled with the mobility of sensing targets, further exacerbates security challenges. Advanced learning algorithms including RL can be used to well learn the behavior of the attackers. Moreover, federated learning (FL) can be used to enhance privacy and security by enabling local model training without transmitting raw data, thereby reducing exposure risks [59].

III. LEARNING ALGORITHMS FOR ISAC BEAMFORMING OPTIMIZATION AND TRACKING

In reality, ISAC systems face inherent trade-offs in beamforming optimization and tracking because they share the same radio resources for both transmitting data to users and sensing the environment. Allocating more power, antenna resources, or time/frequency to communication often reduces the resources available for sensing, impacting detection range and accuracy. Similarly, beam patterns optimized for narrow communication links might not be ideal for wide-area sensing or dynamic target tracking. Furthermore, signals used for one function can act as interference for the other. Achieving optimal performance in ISAC requires carefully balancing these competing demands in resource allocation and beam design.

The design and real-time adaptation of beamforming strategies in ISAC systems present significant challenges that motivate the exploration of learning-based approaches. Traditional model-based optimization and signal processing techniques often face limitations when applied to the complexities inherent in ISAC environments. Learning algorithms offer compelling advantages to overcome these hurdles:

- *Handling High-Dimensionality and Non-Convexity:* ISAC beamforming involves complex, often non-convex, high-dimensional optimization problems balancing sensing and communication. Learning algorithms can effectively approximate solutions and learn direct mappings from inputs to beamformers, potentially outperforming traditional solvers in complexity and speed after training.
- *Adaptability to Dynamic Environments and Real-Time Tracking:* Wireless environments and target locations change rapidly. Learning algorithms, particularly reinforcement learning and online methods, enable fast adaptation and tracking by learning policies that adjust beams in real-time based on environmental feedback, overcoming the latency limitations of iterative optimization.
- *Robustness to Model Imperfections and Reduced Reliance on Accurate CSI:* Traditional methods rely heavily on accurate CSI and system knowledge, which are often unavailable or imperfect. Data-driven learning algorithms can learn effective strategies directly from operational data, reducing sensitivity to model errors and incomplete CSI.
- *Implicit Management of Sensing-Communication Trade-offs:* Explicitly formulating and solving the intricate trade-offs between sensing and communication performance is challenging. Learning algorithms can implicitly learn to balance these competing objectives through carefully designed reward or loss functions during the learning process.

A. Beamforming Optimization

Beamforming, a critical technique in ISAC systems, aims to focus energy in a specific direction to enhance signal quality and reduce interference. Traditional beamforming methods, e.g., SCA [17], often rely on prior knowledge of the CSI, which can be challenging to obtain accurately in dynamic environments. Learning-based approaches are emerged as promising alternatives, leveraging ML techniques to optimize beamforming in real-time based on the collected data. The existing works of learning-based beamforming optimization are classified into three groups named communication performance-oriented optimization, sensing performance-oriented optimization, others. Specifically, those approaches in the first class aim to prioritize the communication function by maximizing the deliverable information for users while ensuring adequate sensing performance whereas the second class aims to prioritize the sensing function by optimizing the target sensing performance without compromising the communication functionalities. The remaining approaches focus on other performance metrics such as bit/package error rate, energy efficiency, security-related metrics.

1) Communication Performance-Oriented Optimization: Various kinds of learning algorithms are applied to maximize the communication rate while guaranteeing the certain quality of sensing tasks. Considering a THz ISAC systems, [60] presents a modulation technique leveraging a dynamic delay alignment (DDA) with active reconfigurable intelligent surfaces (ASRIS)-equipped dual-function base station (BS). A joint optimization problem to maximize sum rate by optimizing transmit beamforming, ASRIS reflection/transmission, vehicle's mobility correlation parameters, and radar receive filter is formulated while keeping the worst radar SNR requirement. A multiagent deep deterministic policy gradient (MADDPG) algorithm [61], in which the elements of ASRIS are implemented as agents in the multi-agent scenario to handle the resource allocation tasks, is proposed to overcome the non-convex nature of the problem. This is due to the fact that the MADDPG algorithm can offer several advantages in adaptability, scalability, and robustness to partial observability. By leveraging the capabilities of MADDPG, this approach aims to enhance system performance in THz communications and sensing. The simulation validates the benefits of ASRIS and the superiority of MADDPG compared to different optimization methods in terms of sum rate from 3% to 27%. However, the ML-based MADDPG algorithm may incur substantial communication overhead for system state exchange, seeking the reducing communication complexity for effective deployment. In addition, the above mentioned works consider the system under perfect CSI, which is impractical for real scenarios.

Unlike works assuming perfect CSI, [64] addressed intermittent CSI in MISO ISAC systems by jointly optimizing CSI decisions and beamforming for communication rate maximization and radar error minimization. The deep reinforcement online learning (DROL) framework is proposed to overcome causality and the mixed integer nonlinear programming (MINLP) by enabling CSI decision-making without relying on current estimates. Despite its effectiveness (reduced CSI overhead, 90% performance improvement vs. exhaustive), the proposed scheme's performance is sensitive to channel model accuracy for prediction/re-estimation. In [65], the effectiveness of RIS in enhancing radar-communication coexistence is assessed via a spectrum-sharing framework between an RIS-assisted cellular system and a MIMO radar. A joint optimizing transmission precoder, RIS phase shift, and radar transmit waveform is formulated to maximize the mutual information (MI) while addressing the challenges of maintaining operational fairness and limiting interference to the radar system. A low-complexity MRL algorithm is developed to enhance spectrum utilization [66], [67]. Simulation confirms MRL/RIS benefits for interference control (0.95 detection probability) and speed (1.5-5x faster). However, the static environment assumption limits applicability in dynamic scenarios, posing an adaptation challenge.

Leveraging the benefits of STAR-RIS, [62] presents a novel algorithm for the multi-dual STAR-RISs (MD-STAR) in the scenario of ISAC systems to provide full-plane coverage for the general distribution of communication users and sensing targets. A sum-rate maximization problem in optimizing active beamforming of dual-function radar-communication (DFRC)-enabled BSs and passive beamforming of MD-STAR in ISAC systems is considered subject to maximum position error bound (PEB) and hardware limitations of MD-STAR. A two-layered multi-agent federated Q-learning (TMFQ) algorithm is proposed in which the inner-layer Q-learning focuses on the solution of BSs and MD-STAR, whereas the outer-layer Q-learning optimizes hyper-parameters including the learning rate and discount rate of the inner-layer one. Federated learning is employed to facilitate information exchange between agents in the inner Q-learning. Empirical evidence suggests the superiority of TMFQ in achieving 10% with federated learning and 5% rate improvement without federated learning compared to benchmark approaches. Nonetheless, implementing MD-STAR with the required reconfigurable elements and control capabilities can be challenging and may introduce additional hardware costs. The system model of RIS-assisted ISAC for high-speed environments is illustrated in Fig. 4. As an advanced solution to the above work, the work in [68] adopts the meta-learning techniques to address the inherent non-convexity of the algorithm, which is originally caused by a joint transmit beamforming and phase shift optimization problem. In [68], a STAR-RIS-aided ISAC system to assist a base station in transmitting communication signals and conducting sensing tasks is investigated. A joint optimization problem for transmit beamforming and phase shift is formulated to maximize the communication data rate and illumination power of targets. A meta soft actor-critic (meta-SAC) [69] algorithm, which is a fusion of the SAC algorithm and meta-learning techniques, is proposed to address the non-convex nature of the problem. The key feature of this work relies on the fact that combining meta-learning with SAC can accelerate convergence and improve performance under dynamic conditions [67]. The simulations show that meta-SAC outperforms traditional DRL methods in terms of the achievable rate of mobile users, highlighting its potential for enhancing ISAC system performance. The performance of the proposed approach may be sensitive to channel variations and uncertainties. Robustness to channel fluctuations will be crucial for real-world scenarios. In addition, STAR-RIS deployments require reconfigurable components and control

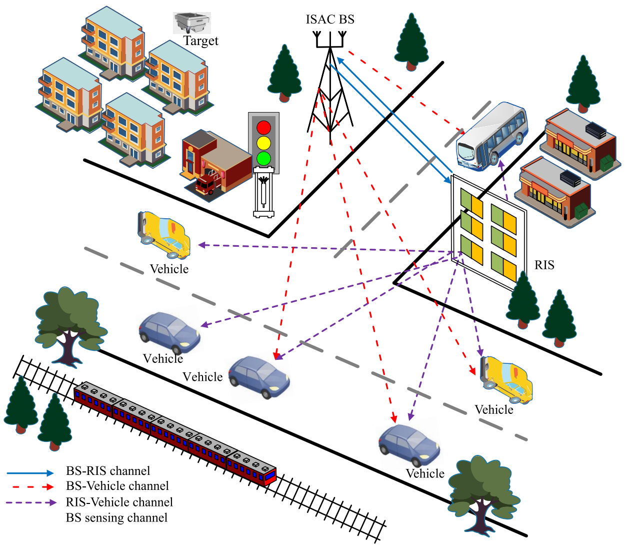


Fig. 4. A system model with RIS-assisted ISAC for high-speed environments, as proposed by [60], [62], [63]. The multiple-antenna BS has capability of communicating with and locate vehicle targets within its coverage area. These signals can travel directly or indirectly to the vehicles via RIS to enhance the ISAC system's performance by dynamically adjusting the phase and amplitude of reflected signals. The RIS is a key element in ISAC, mitigating signal blockage and high propagation loss, thus minimizing positioning errors while meeting various communication demands.

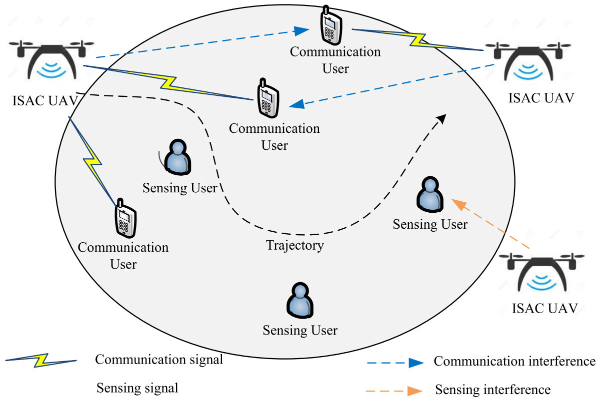


Fig. 5. A system model UAV-enabled ISAC systems, as proposed by [72]. Equipped with a UPA antenna, the UAV needs to periodically reposition itself to maintain optimal communication with mobile users due to their dynamic and unpredictable movements. To track these users, the UAV utilizes radar signals and adjusts its location accordingly as they move.

capabilities that can be challenging and may incur additional hardware costs.

Accounting for UAV deployment, the authors in [70] develop a groundbreaking approach for enhancing ISAC systems by leveraging a UAV-mounted RIS. A communication rate max-min problem is formulated by jointly optimizing the beamforming vector, RIS phase shift and UAV trajectory with respect to the power budget and radar SNR requirements. PPO-based [71] approach is developed that not only addresses the challenges of non-convex problem but also provides a tractable solution [71]. Results highlight its potential to improve signal quality and connectivity coverage in ISAC systems. In a different scenario, [72] presents a DRL-based approach for optimizing UAV-assisted wireless communication networks. The problem of user movement tracking, beamforming, and UAV trajectory optimization is addressed, where the authors focus on maximizing data transmission rates for users in a dynamic environment with respect to the UAV's power constraints and limited flight range. The dual-layer deep unfolding [73] network, coupled with DRL, is developed to enable the concurrent fine-tuning the user movement tracking, beamforming design and UAV trajectory as their adjustments are occurred on different timescales. The results verify the superiority of the proposed approaches in achieving nearly identical sum-rate performance, closely approaching the 99% efficiency of the weighted minimum mean-squared error (WMMSE) algorithm. The major shortcoming in this work is that only one UAV is considered, making the system model less realistic. The system model of UAV-enabled ISAC systems is illustrated in Fig. 5.

2) Sensing Performance-Oriented Optimization: Different from the above approach, the following works [63], [74], [75], [76] took priority in optimizing the sensing performance without compromising the communication performance. To simultaneously reduce computational complexity and performance degradation, the authors in [74] propose an unsupervised learning-based approach for beamforming design in RIS-aided ISAC systems. By employing lightweight structures in [77] and tailored channel samples, an ISAC beamforming neural network (IBF-Net) model is developed to effectively optimize transmit beamformer and RIS phase shift matrix to maximize the sensing SNR of the echo signal subject to the lower bound of communication SNR while reducing computational complexity. This learning framework can eliminate the need for labeled data, further simplifying the design process. Computational experiments reveal that the proposed approach yields satisfying performance and substantially reduced computational complexities. By contrast, the effectiveness of the proposed method in dynamic environments, where channel conditions change rapidly, may need further investigation. To address the characteristics of high speed environments, the authors in [63] investigate an RIS-assisted 6G V2X system under the support of ISAC technology to provide simultaneous positioning and communication capacity for vehicle target. A problem of minimizing positioning error is formulated while maintaining the minimum communication capacity constraint. A flexible deep deterministic policy gradient (FL-DDPG) algorithm is developed to solve the high-dimensional non-convex optimization problem. DDPG [78] is selected as the core algorithm due to its ability to adapt to fast-fading channels, mitigate channel estimation errors, and favor optimal action selection through an ϵ-greedy policy [78]. The proposed FL-DDPG algorithm demonstrates its superiority in enhancing positioning accuracy by a minimum of 89% and improving the achievable rate of the vehicle target by nearly 3 times compared to the traditional methods. A system model with multiple BSs is necessary to capture real-world complexities.

In [75], an AI-based framework for 6G holographic MIMO integrated sensing, localization, and communication (ISLC) is studied to maximize sensing utility, accounting for user distance, beampattern, losses, dense locations, and SINR.

A VAE-based mechanism is used for sensing (user location), and a sequential neural network scheme is used for resource allocation (beamforming). The AI framework proves more effective than LSTM, achieving 34.02% power savings while maintaining high performance. Multiple BS deployment scenarios should be analyzed to account for multicell effects. In contrast, [76] studies an ISAC transceiver with a faulty uniform linear array for simultaneous single-target detection/localization and MISO communication. A differentiable model-based learning approach is proposed for joint transmitter and sensing receiver optimization, utilizing an unsupervised loss for label-free impairment compensation and exploring semi-supervised strategies. Semi-supervised learning achieves performance comparable to supervised learning with only 1.2% labeled data, but is not applicable to multi-target scenarios.

3) Others: Instead of considering the common data rate as a comparable metric for communication performance, the authors in [79] formulate a joint optimization problem that integrates the average SER minimization of multi-user communications, the detection probability maximization and the root mean-squared error (RMSE) minimization of angular estimation of the target into the single optimization framework. A novel end-to-end approach for designing symbol-level precoding (SLP)-based dual-functional ISAC systems is introduced to leverage the waveform design degrees of freedom (DoFs) in both temporal and spatial domains, where the MLP and LSTM networks are employed for this purpose. Numerical analysis confirms the feasibility and effectiveness of the proposed deep-learning-based optimization in enhancing the angular estimation of 23.5% decrease in RMSE of ISAC systems. However, this approach contains a limitation in dealing with latency and computational requirements of deep learning models that may limit their suitability for real-time applications. On the other hand, the availability and quality of labeled data for ISAC systems may pose challenges in real-world deployment.

Among the few articles that consider energy aspects in ISAC systems, the authors in [80] present the integration of simultaneous wireless information and power transfer (SWIPT) techniques to integrated sensing, communications, and powering (ISACP) systems in 6G massive MIMO wireless networks. Particularly, the energy efficiency maximization problem is formulated by optimizing power allocation to mobile users in an FL architecture while keeping the maximum tolerable AoA estimation error as sensing performance. The system models, training data collection method, and actor-critic based multi-agent FL mechanism is set up that shows the effectiveness of the proposed approach in addressing the energy-constrained challenges of ISACP in 6G networks. Coordinating multiple agents and exchanging information in an FL framework may introduce additional communication overhead that requires efficient communication protocols and techniques to minimize overhead and ensure scalability.

Taking into account the security domain for ISAC, the works [81] develop a groundbreaking approach for the cybersecurity algorithms to protect the legitimate users in ISAC system. Specifically, Smart ISAC (S-ISAC) system, which incorporates a unique generative adversarial network combined with a differentiable Kolmogorov-Smirnov loss function, named KSGAN, is proposed to address the challenges of achieving infinitesimal CRLB for sensing and mitigating adversarial machine learning (AML) attacks. GANs demonstrate success in identifying anomalous data patterns induced by adversarial attacks [82]. The proposed KSGAN can effectively identify AML attacks on range-Doppler heatmap features, enabling accurate target vehicle parameter estimation using Constant False Alarm Rate. The numerical analysis demonstrates the superior accuracy of KSGAN in detecting AML, i.e., achieving a True Positive Rate (TPR) from 0.924 to 1, compared to standalone GANs, i.e., TPR from 0.894 to 0.95, and the enhanced performance of MIMO S-ISAC beamforming over standalone ISAC systems, average spectral efficiency of 0.98 compared to 0.92. In such cases, the performance of the learning models may be sensitive to the quality and quantity of training data that presents the challenges in obtaining sufficient and representative data in certain scenarios.

B. Beamforming Tracking

Beamforming prediction and beam tracking are essential components of ISAC systems, particularly in scenarios with dynamic channels and high-mobility users. While beamforming optimization focuses on finding the optimal beamforming weights at a given point in time, beam prediction and tracking aim to anticipate and follow the channel dynamics over time. In general, beamforming prediction involves forecasting the optimal beam direction or weights for future communication frames based on past channel observations and other relevant information. This allows the system to proactively adjust beamforming parameters, reducing the overhead associated with feedback-based methods. Beam tracking, on the other hand, involves continuously adjusting the beam direction or weights to follow the time-varying channel. This ensures that the transmitted signal remains aligned with the channel and minimizes interference.

1) Beamforming Prediction: Considering the unique features of vehicular networks, [83], [84] introduce their approaches to deal with the characteristics of high-speed systems. Notably, [83] propose transformer-based beamforming to overcome beam misalignment and channel fading in high-speed mobile communication. Building on the known capabilities of Transformer networks for optimization, a Transformer's merged-attention mechanism to optimize the mapping from multimodal BS sensor data to the optimal communication beamforming vector. This involves extracting and fusing features from the sensor data using 3D ResNet-18. The data demonstrates the superiority of the proposed multimodal in achieving high accuracy, i.e., 13% to 34%, compared to single-modal schemes in the high-speed environment. A new method is proposed in [84], where the authors study a DL-based beamforming prediction framework for ISAC beam-tracking in vehicular networks. This method bypasses explicit channel prediction to reduce signaling overhead and computational complexity. The proposed EPDnet architecture, which combines CNN and LSTM modules, is implemented to optimize the beamformer for enhanced sensing and communication performance. This optimization is guided by a network utility function based on joint CRBs, evaluating sensing performance subject to a minimum downlink communication sum-rate. Achieving favorable sensing without compromising communication is demonstrated; however, high-speed handover issues are critical for ensuring long-term QoS.

Exploring the capability of the roadside unit (RSU) in high-mobility vehicular networks, the works in [85] develop a groundbreaking approach by incorporating vision techniques to enhance the capabilities of vision-based algorithms, addressing limitations of traditional approaches that are unable to overcome in ISAC systems. For instance, the target-to-user association problem in next-generation V2I ICAS systems is examined, where the BS is equipped with a MIMO radar-assisted communication system. Specifically, the authors develop a DL-aided method to associate vehicular radar targets to communication users in the communication beamspace by modeling the beam prediction problem as a classification task over a fixed-size codebook of beamforming vectors. The renowned YOLO architecture is modified to infer both the radar targets' classes and the beamforming vectors' classes for each detected target. The results demonstrate the capability of performing both beam prediction and radar target detection, with nearly 48.7% improvement in probability of association over different antenna array sizes.

Taking advantage of the unique benefits of a multimodal approach, [88] makes a significant contribution by proposing a Multi-Modal Feature Fusion Network (MMFF-Net) as the core of proactive beamforming scheme for mmWave massive MIMO vehicular communication networks. By integrating multi-modal sensing and communication, the proposed approach can enhance beam alignment accuracy and improve communication data rate by obtaining comprehensive environmental features instead of relying solely on communication processing. The validation on the enriched Vision-Wireless dataset demonstrates that MMFF-Net can downgrade the angle prediction error below 32% for the overall process and 12% when vehicle approaches the RSU, respectively. Specifically, the achievable rate is gained over 25% while an outage probability is reduced more than 16% when a large antenna array is adopted. These results again confirm more accurate and stable angle prediction even in complex dynamic scenarios and adverse environmental conditions over benchmark schemes.

To enhance the awareness of capturing temporal correlations, [86] presents a novel end-to-end predictive beamforming (E2E-PB) approach for ISAC systems in V2I networks, that obtains the beamforming vector directly from the reflected signal samples and effectively bypasses the intermediate state parameter estimation step for enhancing the achievable rate. The proposed attention-based LSTM network is developed to capture temporal correlations in reflected signal samples, enabling direct optimization of the beamformer. The results reveal the superiority of the E2E-PB approach compared to traditional methods in terms of achievable rate, i.e., more than 23%, and show comparable performance when compared with the optimal beamforming design with perfect CSI, i.e., less than 5% error. A limitation of this work is its reliance on a single RSU model, which may not fully capture real-world complexities.

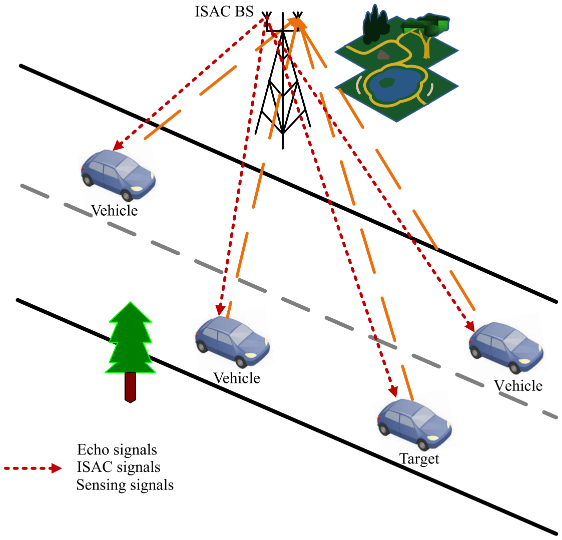


Fig. 6. A system model of ISAC-empowered V2I systems, as proposed by [86], [87]. The roadside unit, i.e., ISAC BS, uses distinct antenna arrays for transmitting downlink signals and receiving reflected echo signals. By integrating ISAC functionality, the system can determine the beamformer based on these reflected sensing signals from vehicles, eliminating the need for pilot signals and feedback overhead.

Leveraging federated transfer learning to preserve user privacy, the authors in [87] present a novel radar-aided federated transfer beam prediction (RaFT-BP) approach for beam prediction in MIMO communication systems to secure user's information privacy. By effectively utilizing radar data and deep learning techniques, the proposed approach enables accurate beam selection with limited training samples, which can reduce data dependency along with real-time training and prediction capability. Computational experiments reveal the superiority of RaFT-BP in achieving the 93.78% top-5 accuracy with 600 samples in the distributed node, enabling 11.9% to 33.2% beam selection accuracy improvement compared with baseline schemes. The performance of the RaFT-BP algorithm may be sensitive to the quality and availability of radar data and communication data. Ensuring data quality and sufficient data collection will be crucial for effective training and performance. The system model illustrated ISAC-empowered V2I networks is provided in Fig. 6.

Different from the above approaches, several works consider the solution to reduce the beam training overhead, that may degrade the system performance [89], [90], [91]. The authors in [89] make a significant contribution by proposing an approach to reduce beam training overhead in mmWave and THz communication systems by leveraging distributed sensing nodes equipped with RGB cameras. By extracting environmental semantics from captured images and transmitting semantic data rather than sending raw images to BS, the proposed solution significantly alleviates data transmission overhead that enhances the system's adaptability and responsiveness to dynamic environments, allowing for prioritization of relevant information. Experimental results on the DeepSense 6G dataset show that the proposed approach can help the BS in finding the optimal beam in over 75% of instances for both distributed nodes, verifying the accurate predicting optimal beams while reducing sensing data transmission overhead. However, the performance of the proposed approach may be sensitive to environmental factors such as lighting conditions, occlusions, and changes in the environment that require the adaptation of the semantic extraction algorithms to handle different environmental conditions.

In contrast, the work in [90] adopts the hybrid deep quantum-transformer networks (QTNs) for beam prediction in mmWave ISAC systems. By leveraging multimodal sensory data and deep learning, the proposed method effectively addresses the challenges of high beam training overhead in V2I communications. The experimental results demonstrate the superiority of the hybrid QTN model in achieving high accuracy, i.e., distance-based accuracy (DBA) score of 0.9124 for multimodal and 0.8832 for position-based data, and outperforming existing approaches, even in zero-shot testing scenarios. The low complexity and high performance of the QTN models highlight their potential to improve beam management in mmWave ISAC systems. The current approach is limited by the absence of effective error mitigation techniques for quantum noise and decoherence.

2) Beamforming Tracking: Considering the dynamic channel conditions in mmWave communication system, [92] makes a significant contribution by proposing the beam tracking scheme leveraging multi-point radar sensing to address the challenges of frequent beam switching and signaling overhead in high-mobility scenarios. By aggregating sensing information from multiple radars, a DNN model is developed to predict the beam direction from the learning channel nonlinear correlations. The experiments verify that the proposed approach effectively operates in blockage scenarios and outperforms baselines that can be gained up to 86% of the upper bound in terms of spectral efficiency. While this assumption works well for non-dense scenarios, a new approach is needed for dense urban environments.

The works in [93] specifically address the challenges of beam tracking in high-speed environments by targeting the ISAC RCG-Net beam tracking solution for IoV on complex road trajectories. By integrating CNN and GRU, the proposed approach leverages DNNs to learn and predict angle changes, enhancing beam tracking precision and addressing communication instability issues. The experimental data confirms the superiority of the RCG-Net solution in terms of high angle tracking accuracy, i.e., over 80% angle tracking accuracy within an error tolerance of 0.5 degrees, and communication performance compared to the traditional beam tracking solutions. The system model's limitation to a single RSU restricts the scope of this work.

C. Discussion

Solving beamforming design problems is inherently challenging. While learning algorithms, particularly DRL, prove effective in conventional communication systems, the complexity escalates considerably in ISAC systems because communication and sensing must share limited resources like power, bandwidth, and hardware. This resource sharing introduces intricate trade-offs and non-convex optimization challenges. To address these compounded difficulties and enable effective resource allocation, as observed from the existing works, more sophisticated learning paradigms are proposed and utilized such as meta learning (MRL), federated learning approaches combined with DRL, and deep unfolding techniques.

Different beamforming design issues in ISAC systems necessitate the use of specific learning algorithms. In the context of beamforming optimization, particularly for dynamic environments encountered in ISAC systems, MRL algorithms are a promising avenue, primarily owing to their inherent ability to achieve faster adaptation and efficient learning on new tasks. In addition, unfolding learning for ISAC beamforming optimization offers a powerful blend of model-based optimization principles and data-driven learning capabilities. It provides a computationally efficient way to solve complex, non-convex joint communication-sensing beamforming problems. In contrast, beamforming tracking benefits significantly from learning algorithms capable of capturing temporal correlations. LSTM networks, for instance, are well-suited to model the dynamics in sequential CSI or sensing data. Furthermore, the integration of multimodal features (e.g., visual, radar, and communication signals) and attention mechanisms offers a promising avenue. This is because accurately tracking and predicting the optimal communication beamforming vector requires processing and correlating diverse data sources, thereby optimizing the mapping from sensor data to the desired beamforming direction.

Despite the notable advantages that learning algorithms offer to ISAC systems, deep learning approaches often face practical limitations. Acquiring the substantial amounts of training data can be particularly challenging in specific ISAC deployment scenarios. Furthermore, the scarcity and quality of labeled data for ISAC tasks can significantly impede their real-world implementation. In the context of beamforming tracking, the performance of learning-based methods can be susceptible to dynamic environmental factors such as interference and signal blockages. Therefore, adapting these models to maintain robust performance across diverse and changing environmental conditions is a critical area for future research. Promising future directions include exploring the potential of emerging learning paradigms like Generative AI and diffusion models to address these challenges and further enhance the capabilities of ISAC systems.

IV. LEARNING ALGORITHMS FOR WAVEFORM DESIGN

Traditional communication and sensing systems usually require different performance for their transmit waveforms. Specifically, the communication-only waveforms are expected to enable reliable and low-latency transmission of their modulated information via the wireless channels, while the sensing-only waveforms are desired to possess specific properties to support effective extraction of accurate sensing target

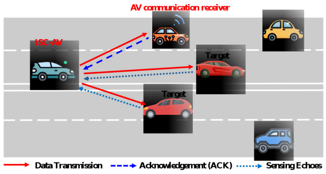


Fig. 7. An ISAC-AV system model is proposed by [97], in which an autonomous vehicle transmits a mmWave waveform for communicating with an AV communication receiver and detecting nearby target vehicles simultaneously.

information from the reflected echoes. Different from the independent design of such traditional communication and sensing waveforms, the ISAC waveform design aims at optimizing the transmit waveforms such that both communication and sensing performance can be achieved simultaneously. However, it is a challenging task due to the dynamics of the time-varying wireless environments and mutual interference between the sensing and communication subsystems. Conventional studies (e.g., [94], [95], [96]) usually formulate the ISAC design problem as the complex optimization problems which are in general compute-intensive due to the involved coupled variables, non-convex objective functions and constraints. Thus, solving these formulated optimization problems typically incurs high computational complexity, thus cannot always provide the optimal ISAC waveforms in real time, especially in the complex and dynamic environments. Meanwhile, the heuristics and approximate approaches may only obtain the sub-optimal solutions to the formulated optimization problems, resulting in the degradation of sensing and communication performance. Recently, the learning algorithms have been widely adopted for the ISAC waveform design since they can address the computation challenges of the optimization-based approaches by transferring the complex optimization tasks into the offline training phase. Then, the trained models are deployed to provide the real-time ISAC waveform solutions at run time. In this section, we review the recent studies on the learning-based ISAC waveform design approaches. Depending on the learning algorithms used for the proposed approaches, these studies can be divided into the following three categories: reinforcement learning, supervised learning, and deep unfolding.

A. Reinforcement Learning (RL)-Based Waveform Design

The studies in [97], [98], [99], [100] adopt various RL algorithms to design the waveforms in different ISAC systems. For instance, the study in [97] proposes a DRL-based approach to adapt the waveform structure of an autonomous vehicle (AV) equipped with a mmWave ISAC system, namely ISAC-AV as illustrated in Fig. 7. The waveform design problem is formulated as an MDP which considers the AV's data queue length and wireless link quality measured by SNR as a state to select the number of data frames of the waveform (i.e., the waveform structure). The objective is to balance the data transmission efficiency and sensing accuracy by designing the reward function as a weighted sum of the sensing target velocity estimation error, data queue length, and number of dropped packets due the queue full. A DRL-based approach which combines the DQN [101], double DQN [102] and dueling DQN [103] algorithms is proposed to learn the optimal policy of the formulated MDP. The extensive simulations show that the proposed approach can achieve up to 40% reward improvement and 50% average packet drop reduction, compared with three baselines based on the Q-learning, greedy, and deterministic policy algorithms. However, the motivation of combining the innovations from three different DRL algorithms is not well justified due to the lack of performance comparison with these DRL algorithms separately.

The authors in [98] formulate an optimization problem to find the optimal dual waveform transmitted for both radar and communication purposes in a MIMO ISAC system. The objective is to minimize the multi-user communication interference while producing the appropriate beampattern of the transmitted waveforms for accurately detecting the sensing targets. Solving the formulated optimization problem requires prior information about the target quantity and locations, which may be challenging to obtain in complex sensing environments. Thus, a model-free RL-based learning approach based on the SARSA algorithm is proposed to address this challenge. Specifically, the detection statistic of the received sensing signals is considered as the system state to select an action for determining the angle bins of the sensing targets which can maximize the target detection probability (i.e., the immediate reward). Then, the selected target angle bins are used as inputs to determine the optimal transmit waveform at the current state. The presented simulation results indicate that the proposed approach can achieve up to 80% sensing target detection probability higher than those of the omnidirectional and directional trade-off design approaches proposed in [50]. However, with the main focus on maximizing the sensing accuracy, the proposed approach does not help improve the communication performance compared to the baselines.

The authors in [99] propose a learning-based approach to optimize the co-design of the transmit waveform and receiver in a MIMO ISAC system. The proposed approach employs three fully connected networks (FCNs) to implement the waveform design module, sensing and communication receivers, respectively. Specifically, the FCN of the waveform design module plays the role of an RL agent which aims to select an action for determining the optimal transmit waveform at the beginning of each sampling period based on the initial waveform (i.e., the state). The reward function is defined as a weighted sum of the target detection and communication symbol decoding performance which are determined based on the outputs of the FCNs at the sensing and communication receivers. Numerical analysis results show that the proposed approach can achieve a higher target detection probability while maintaining a lower communication symbol error rate (SER), compared with the receiver-only design approach.

Moreover, the study in [100] develops a RL-based approach to guarantee the coexistence of an RIS-assisted radar and a MIMO communication system operating in the same frequency band. Specifically, an optimization problem based on MI of the communication system is formulated to jointly find the optimal transmit waveform, RIS phase shift matrix, and communication transmit precoder matrix under various constraints, such as the allowable interference, communication and sensing transmit signal powers. Then, the formulated optimization problem is converted into a MDP which considers the RIS phase shift and wireless channel gain as the state to select an action determining the optimal waveform, RIS phase shift matrix, and communication symbol precoder matrix to maximize the communication MI. Finally, a learning framework based on the DDPG algorithm is proposed to learn the optimal action selection policy. The proposed framework can achieve up to 62% communication MI (i.e., reward) improvement, compared with the learning framework based on the TDC3 algorithm.

Unlike the above studies [97], [98], [99], [100] that focus on designing the optimal ISAC waveforms, the authors in [104], [105] leverage the RL algorithms to optimize the resource allocation for transmitting such waveforms to further improve the sensing and communication performance. For example, the study in [104] considers an ISAC system with a MIMO automotive radar, which uses the two non-overlapping but interweaving transmit antenna subarrays for transmitting its waveform signals. A RL-based learning framework based on the SARSA algorithm is proposed for adaptively allocating the proper number of antennas for transmitting the sensing and communication signals, in response to the run-time change of the dynamic driving condition. Specifically, the proposed framework considers the current number of sensing antennas as a state to determine the optimal number of sensing antennas in the next control period. The reward function is designed to maximize the sensing direction of arrival (DoA) estimation accuracy while maintaining the received SNRs of all communication users above the desired thresholds. Various Monte-Carlo simulations are conducted to justify the superior sensing and communication performance of the proposed approach, compared with two baselines based on the random and uniform antenna allocation.

In addition, the authors in [105] considers an ISAC system which use different subcarriers of a time-varying wireless spectrum for transiting the sensing and communication signals under the non-stationary interference. Specifically, they propose a model-free online RL approach which takes the current wireless channel qualities of all subcarriers as the state. Then, an action is selected to allocate the sub-carriers and corresponding power levels for transmitting the sensing and communication waveforms. The reward function is designed such that the action selection can maximize MI-based sensing utility [106] while maintaining the MI-based communication utility above a threshold. Moreover, the proposed approach employs the two feed-forward neural networks (FNNs) to implement the RL agent for learning the optimal subcarrier allocation policy. The simulation results show that without the prior knowledge about the environment dynamic model, the proposed approach can achieve the similar or slightly better sensing and communication MI performance than the model-based learning approach proposed by [107]. However, as a model-free learning model, the proposed approach exhibits a slower learning convergence, thus requires a larger amount of training data to achieve the maximum sensing and communication performance.

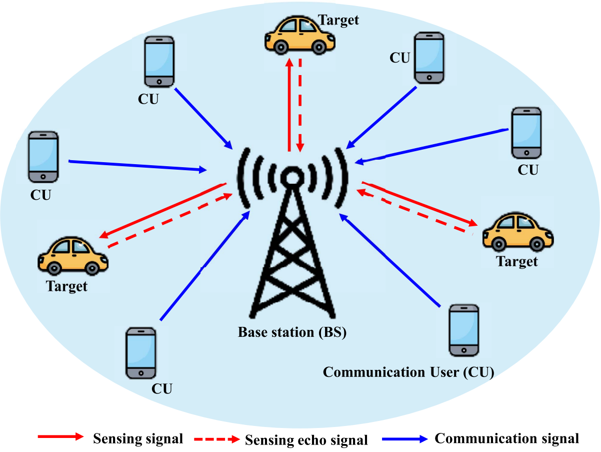


Fig. 8. An uplink ISAC system is proposed by [108], in which BS broadcasts the waveforms to transmit the data to the communication users (CUs) and then receives the echoes reflected from the sensing targets.

The studies in [108], [109], [110], [111] adopt various supervised learning (SL) models for the ISAC waveform design optimization. For example, the authors in [108] consider an ISAC system in which the BS broadcasts the waveforms and then receives the echoes reflected from the targets and the communication signals transmitted by the communication users (CUs) simultaneously, as illustrated in Fig. 8. The received sensing echoes and communication signals are fed into an target estimator and a communication symbol decoder for extracting the target information and decoding the CUs' data, respectively. An optimization problem is formulated to determine the optimal waveform and communication receive beampattern with the objective of jointly maximizing the communication and sensing rates. Then, this complex optimization problem is transformed into an equivalent problem with reduced complexity. Finally, a CNN-based approach, called ISACNN is developed to solve the reduced optimization problem. The simulation results indicate that the ISACNN can achieve the similar sensing and communication performance with reduced complexity, compared the baselines based on the optimization [112] and successive interference cancellation algorithms [113]. However, the ISACNN requires a comprehensive training dataset, which may be challenging in real-world ISAC systems.

Similarly, the authors in [109] propose an NN-based approach for the joint design of the waveform, transmitter and receiver in a MIMO ISAC system. The proposed approach employs six feed-forward NNs which are implemented for the communication encoder, radar beamformer, target presence detector, target angle estimator, uncertainty estimator, and communication receiver, respectively. Moreover, five loss functions are designed for training these NNs together in an end-to-end autoencoder (AE) framework to jointly optimize the design of the transmit waveform, target detector, and communication receiver. The simulation shows that the proposed NN-based learning approach can achieve a similar communication symbol error rate and target detection probability, compared with the benchmarks based on the traditional complex signal processing approaches for the transmitter, target detector and communication receiver, even with the hardware impairments of the transmit antenna array.

Differently, the studies in [110], [111] investigate the SL-based learning approaches for modulation and demodulation of the ISAC waveform signals. For instance, the authors in [110] develop an NN-based modulation approach in an ISAC system which transmits the data packets with the short block length. It also employs the noncoherent detector and MSE filter for the communication and radar receivers, respectively. Specifically, the authors employ a multilayer perceptron (MLP) model to implement an arbitrary encoder that takes a multi-bit information sequence as inputs to generate a zero-mean codeword for transmitting through the communication and sensing channels, such that the best communication decoding and target detection accuracy can be achieved. The developed MLP is a conditional NN [114], which can have different hyper-parameters with different inputs. Moreover, a weighted sum of the communication and sensing costs is defined as a loss function to train the developed MLP, in which the network weights can be changed to adjust the priority of the communication and sensing. The proposed MLP-based encoder achieves a lower bit error rate (BER) than the symbol-wise modulation based on the quadratic phase-shifted keying scheme. However, the proposed encoder approach can be only applied to the ISAC system in which the correlation between the wireless communication and sensing channels is known prior. Thus, it may not achieve a good performance in complex ISAC systems where the channel correlations cannot be accurately estimated.

Moreover, the authors in [111] employ the NN for the ISAC's intelligent receiver to identify the modulation types of its received signals. In particular, this study proposes a DNN-based automatic modulation recognition (AMR) approach which classifies the eight modulation types of the communication and sensing signals using the following two main steps. The first step adopts a time-frequency analysis algorithm, called smooth pseudo Wigner Ville distribution to transform the signal into a time-frequency image (TFI) with time-frequency characteristics. In the second step, the obtained TFI is fed into a DNN model to predict the modulation type of the input signal. Furthermore, the proposed AMR approach employs a neural architecture search (NAS) algorithm based on a differentiable neural structure search (DARTS) approach [115] to find the optimal structure of the DNN which is suitable for identifying the modulation type under complex electromagnetic environment conditions. The proposed DNN-based AMR approach can accurately identify the modulation types of the communication and radar signals, even when the SNR is less than 4 dB. Moreover, the proposed approach achieves the highest modulation type recognition accuracy and the lowest memory and computation overheads, compared with the four baseline approaches based on representative DNN models and automatic neural architecture search algorithms.

B. Deep Unfolding-Based Waveform Design

The recent studies in [116], [117], [118] employ the deep unfolding techniques to develop the model-driven NNs for optimizing the waveform design of ISAC. For instance, the study in [117] investigates a RIS-aided ISAC system which employs multiple reflection elements to establish the virtual line-of-sight links with its communication users. An optimization problem is to find the dual-functional waveform and RIS phase-shifts such that both the sensing and communication performance can be maximized. Then, an alternating direction method of multipliers (ADMM)-based iterative algorithm is proposed to decompose the formulated optimization problem into four subproblems. However, this iterative algorithm is compute-intensive due to involvement of complex matrix inversion operations. Thus, the authors further propose a low-complexity approach, called ADMM-NET which unfolds the ADMM-based iterative algorithm to a layer-wise NN architecture. The ADMM-NET reduces the training time and number of required training samples up to about 15% and 17%, respectively, compared with the ADMM-based iterative algorithm. Moreover, the ADMM-NET can achieve a BER of 10^{-2} when the communication SNR is 10dB.

The authors in [118] propose a deep unfolding-based waveform design approach for a wideband large MIMO ISAC system which operates in mmWave frequency bands. Specifically, an optimization problem is formulated to determine the optimal waveform that can provide a good trade-off between the sensing and communication performance. Similar to [117], the objective function also aims to minimize the weighted sum of MUI energy and waveform discrepancy subject to the transmit power constraints. However, solving the formulated optimization problem requires high computation overheads due to its NP-hard nature. Thus, to find the optimal solution with less computation complexity, a deep unfolding NN model is proposed to unroll the iterative projected gradient descent optimizer into a sparsely connected structure which can be trained using an unsupervised method. The simulation results demonstrate that with the SNR of 12 dB and sensing beampattern MSE of 10 dB, the proposed approach can achieve a communication sum rate of 14 bps/Hz which is 3x higher than that of the conventional branch-and-bound optimization-based waveform design approach.

Moreover, the ISAC systems in [117], [118] adopt an active sensing approach in which the sensing transmitter and receiver are co-located to transmit the waveform signals and receive the sensing echoes at the same time, respectively. The active sensing may suffer from high self-interference at the sensing receiver. Differently, the study in [116] considers an ISAC system with passive sensing, in which the sensing receiver is placed far from the transmitter. Specifically, the considered ISAC system transmits the same OFDM signal via multiple antennas as illustrated in Fig. 9. The receiver processes its received multipath signals among which the NLoS signals are reflected from the sensing targets for recovering the transmitted communication information and sensing target parameters. The receiver's signal processing problem is formulated as a

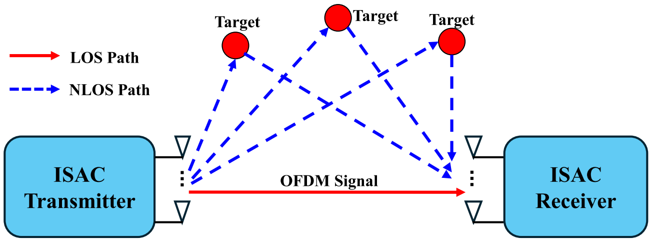


Fig. 9. An ISAC system with passive sensing and communication is proposed by [116], in which the sensing receiver is placed far from the transmitter and recovers the environment parameters using the received communication signals.

multi-parameter optimization problem to minimize the error between the real received signal and the reconstructed received signal. Then, an iteration algorithm is proposed to find the optimal solution of the formulated optimization problem. However, setting the proper values for the hyper-parameters of the proposed iteration algorithm is non-trivial. Improper hyper-parameter settings may degrade the sensing and communication performance as well as the algorithm convergence speed. Thus, a model-driven NN approach, called ISAC-NET is further proposed to learn the proper hyper-parameters. Compared with conventional communication modulation and target detection algorithms, the proposed ISAC-NET can reduce the communication SER and the normalized MSE of target velocity estimation by up to 10000x and 10x when the signal SNR is 20 dB.

C. Discussion

Existing studies have utilized three main types of learning algorithms for optimizing various aspects of ISAC waveform design and achieved outstanding communication and sensing performance improvements, compared with the traditional optimization-based ISAC design approaches. Specifically, RL is often adopted to adapt the configuration/structure of the transmit waveform and the transmitter's resource allocation in response to the changes of dynamic wireless environments, such that the good trade-off between the sensing and communication performance can be always achieved. In other words, RL is mainly applied to optimize the transmit waveform of dynamic ISAC systems. On the other hand, SL can allow the joint optimization of both the transmit waveform and the communication/radar receivers. Moreover, RL agents can learn the optimal waveform design solutions online by directly interacting with the environments over a long-term period. Differently, the SL-based approaches are typically utilized when the massive datasets can be collected for training the SL models in an offline fashion. Furthermore, both the RL-based and SL-based approaches are regarded as black boxes without interpretability and often require the long convergence time or/and massive training samples to achieve the desired ISAC performance. Meanwhile, unfolding learning recently emerged as a promising approach which can leverage both the learning capabilities and domain knowledge to build the model-driven NN models for obtaining the optimal and explainable ISAC waveform design solutions. In particular, the unfolding learning is well-suited for the complicated waveform design problems, especially when labeling the training data samples is challenging.

Applying the learning algorithms to optimize the ISAC waveform design suffers from two following major drawbacks. First, the learning algorithms often require comprehensive training datasets to fully understand the dynamics and uncertainty in the ISAC resource availability and time-varying wireless environments. Such datasets should cover all possible sensing and communication conditions of the applied ISAC systems to minimize the number of unseen environment states in the training phase. This requirement may limit application of the proposed learning algorithm-based ISAC waveform design approaches. This is because it may be time-consuming and labor-intensive to collect the large training datasets in real ISAC systems, especially with the highly-dynamic wireless environments and complex system design. Second, the learning algorithms may also require a long training time to learn the optimal ISAC waveform solutions. Specifically, the applied ISAC systems may suffer from the poor sensing and communication performance due to the non-optimal waveform solutions for a long period of time, when the online training methods (i.e., learning through direct interactions with the real ISAC systems) are adopted. Thus, it is critical to investigate new sample-efficiency learning algorithms for optimizing the ISAC waveform design with less training data samples and faster learning convergence rate. The generative diffusion model (DM) is a promising solution to achieve this goal. Specifically, the DMs have achieved successful application cases in wireless networks due to their remarkable ability in capturing complex data probability distributions and generating high-quality data [119]. For instance, the DMs can be used to generate the new data samples for training a SL-based ISAC waveform design algorithm based on a small number of collected data samples. Moreover, the DMs can be utilized as a policy network in the DRL algorithm to achieve a better trade-off between exploitation and exploration, thus reducing the required training time.

V. LEARNING ALGORITHMS FOR AOA/AOD ESTIMATION AND SIGNAL CLASSIFICATION

By merging sensing and communication into a unified framework, ISAC enables accurate environmental awareness and efficient information exchange for sensing, while simultaneously elevating communication capabilities like resource optimization. To do so, it requires two characteristics: 1) accurate AoA and AoD estimations and 2) strong capability in signal detection or classification. Particularly, the first characteristic aims to provide precise measurements of the direction of channels from which signals arrive at a receiver and the direction in which they depart from a transmitter. This enables ISAC systems to efficiently track objects, locate targets, and improve sensing capabilities. Meanwhile, the second characteristic plays a vital role in recognizing the presence of a signal amidst noise with detection activities and distinguishing the type or nature of either communication data or sensing signals from a mixture of signals acquired from the environment with classification activities. These activities help reduce the impact of interference caused by radar sensing on communication signals and vice versa.

Although there are several traditional techniques used for AoA and AoD estimations (e.g., beamforming-based estimation, maximum likelihood estimation, least squares, and phase interferometry) as well as signal detection and classification (i.e., matched filter, energy detection, correlation-based detection, waveform-based detection, and Fourier transform-based classification) [120], they often rely on mathematical models that struggle with dynamic environmental changes, where targets are typically in high mobility, and diverse noise interference. Moreover, finding the exact solutions by such methods is limited by the scale-up of ISAC components, like multi-objective detections or multi-antenna missions. These factors give rise to complex signal processing with high-rank matrix transformation. However, this can be addressed by the adoption of learning-driven techniques in ISAC, particularly DL in handling vast amounts of data and extracting complex patterns, due to its capability in universal approximations. Such approaches help enhance estimation precision but also address computational challenges, enabling the training of ISAC systems with offline data and then using an inference model to enhance the responsiveness in fast-changing wireless environments. Similarly, learning-driven signal detection and classification techniques bolster ISAC systems' ability to recognize and process signals with heightened accuracy, even in dynamic and noisy environments without requiring predefined algorithms. Given the critical role of ISAC and the advancements driven by learning algorithms, this review explores the latest advancements in ISAC, structured into two parts: the first discusses learning algorithms for AoA/AoD estimation in Section V-A, while the second focuses on signal detection and classification strategies in Section V-B. Section V-C concludes with insights into current research trends and future directions.

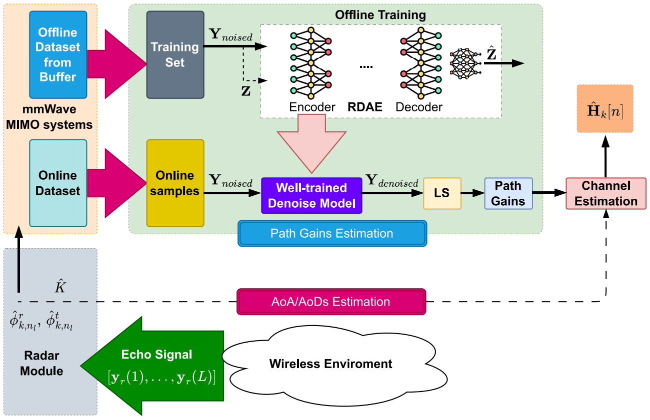


Fig. 10. Illustration of the workflow for the entire channel estimation with two main phases [120]. In the AoA/AoD estimation phase at the radar module, the radar processes echo signals y_r from L-length of transmitted waveforms to obtain the latest angle data ($\hat{\phi}^r_{k,n_l}$ for AoA, $\hat{\phi}^t_{r,n_l}$ for AoD) for $\hat{K}$ targets before inputting them into the mmWave MIMO system. For path gains estimation at the communication module, the vehicle buffer stores wireless signals $\mathbf{Y}_{noised}$ and $\mathbf{Y}_{denoised}$ as training data for the residual denoising autoencoder (RDAE) during the offline training stage, combined with input data $\mathbf{Z}$ through beamforming based on AoA/AoD. In the online stage, real-time signals are input into the mmWave MIMO system, outputting denoised data $\hat{\mathbf{Z}}$. Path gains and AoA/AoD outputs are combined to estimate the mmWave channel $\hat{\mathbf{H}}_k[n]$.

A. AoA/AoD Estimation

Among various issues of ISAC functions, accurate AoA/AoD estimation takes an important role in tracking targets as well as facilitating real-time decision-making in intelligent applications such as autonomous driving and location sensing. In [120], the authors develop a DL-aided MIMO estimator with two phases, as shown in Fig. 10. Technically, in the first phase, a novel subspace decomposition is considered to reduce the deviation between the nominal array manifold and the actual received data by finding AoA/AoD rooting rather than quantization, thus eliminating quantization error. In the second phase, a DL-based channel gain estimator trained with known angle information can denoise the received signal during gain estimation by combining a denoising autoencoder and a least-square estimator to filter out complex noise with minimal data and iterations. Simulation results show significant performance improvement for high-mobility mmWave channels with limited training resources. Focusing on MIMO radar angular resolution capable of separating multiple targets, the work in [121] proposes a DL-based supper resolution AoD estimation using a 2-dimension CNN to center the bearing angle zone and a spatial filtering pre-rotator to achieve a wider radar field of view. This detector, therefore, achieves a zoomed-in small bearing angle range without interference from out-of-beam targets, making it significantly superior to benchmark estimators. For example, it can achieve 0.4 degrees of resolution of 90% probability of resolution in 2 targets of classification at the transmit SNR of 7 dB if targets are located in the gird and the same performance with that of 12 dB.

To enhance the training behaviour, the work in [122] studies the impact of multiple target and snapshot sensing in an autoencoder approach for joint detection and location tasks, emphasizing the importance of parallel NN-based detection designs to not only train the autoencoder but also target a fixed false alarm rate for target detection. Besides, to separate sensing tasks during training, the number of radar targets is fed into the angle estimation network only if targets are present. The bit-wise mutual information is used as the key to measuring the quality of the communication together with modifying loss functions to optimize network performance through three stages: first, combining communication and angle estimation, then communication and detection, and finally, all three together. Unlike [122], the work in [123] focuses on the decomposition of loss functions to extract the respective SNR values and the number of sensing snapshots as well as to reflect the joint impact of communication, detection, and AoA estimation. This is achieved using bound-informed adaptations to ensure the neural network's capacity. The proposed method performs better than classical algorithms, such as the Neyman-Pearson-based power detector for object detection and ESPRIT for AoA estimation, particularly at low SNRs such as -5 dB SNR. However, the multi-snapshot estimation may not be effective for low-speed moving targets since the snapshot samples generated by this estimation cannot guarantee randomness during training and normalization of the expected magnitude of the gradients, causing overfitting or bias in learning.

In light of the sensing-integrated discrete Fourier transform spread OFDM function, the work in [124] introduces a DL-based ISAC receiver that features two NN blocks for improved performance in THz channels. The first block focuses on sensing, while the second extracts channel information from received signals. Testing with a subcarrier spacing of 15×2^n kHz, the approach shows enhanced communication and sensing capabilities, along with robustness to Doppler effects and multi-target estimation. Meanwhile, [125] presents a CNN framework that estimates the range and velocity of moving targets directly from radar range-Doppler maps, achieving a peak SNR improvement of 33 dB over D-periodogram, 21 dB over 2DResFreq, and 10 dB over VGG-19 at a transmit SNR of 30 dB. Additionally, [126] proposes two solutions for joint AoA and AoD estimation. The first uses complex-valued NN for improved channel matrix representation, while the second involves transforming the received channel matrix into smaller matrices. The first approach offers reduced computational complexity and precise target estimation, making it suitable for autonomous ISAC systems in dynamic environments.

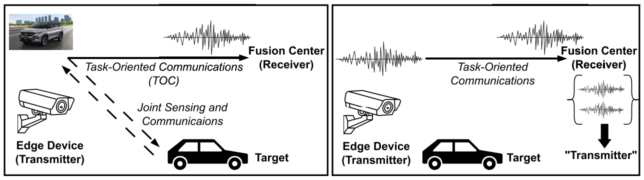
To provide consistent channel gains and better handling of Doppler shifts, the works in [127], [128] consider using OTFS waveforms to transform a time-varying multipath channel into a more stable two-dimensional channel. Specifically, in [127], a DL-based tap detector is developed to determine thresholds from estimates of normalized delay and Doppler shift for target tracking. This detector uses one CNN to discriminate targets based on delay from the input of the estimated matrix (i.e., downsampling semantic segmentation) and another CNN for extracting the normalized Doppler shifts related to the tap (i.e., upsampling semantic segmentation). By quantifying OTFS waveforms of frequency at 26 GHz and signal modulated by quadratic phase-shifted keying, the proposed DL-based detector can approximate the detection probability of the upper bound of ideal threshold determination and is much better than that of the practical one. Meanwhile, a deep unfolding network, called ADMM-Net, in [128] enhances AoA estimation accuracy by transforming the iterations of the alternating direction in the multipliers algorithm into multi-stages of DNNs with learnable parameters. ADMM-Net can achieve the lowest normalized MSE and the lowest running time over the existing algorithms.

B. Signal Detection/Classification

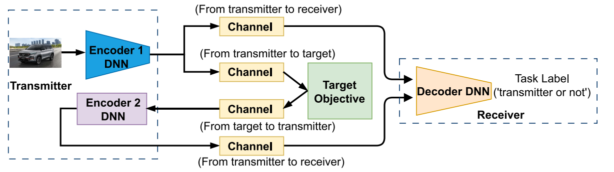
Integrating radar and communications functions into unified ISAC systems addresses spectrum scarcity but presents signal detection challenges, particularly with similar signal strengths. Inspired by this, the work in [129] proposes two cyclic profile-based learning frameworks using wideband cognitive radios to detect communication signals. Features are extracted, and a CNN/ANN identifies signals as radar or communication. Experimental results with six types of signals show effective separation of signal types at low SNR (as low as -10 dB) with upper 80% accuracy, indicating the ANN classifier's superiority over the CNN. In contrast, [130] improves signal classification accuracy using a compact CNN architecture inspired by LeNet-5, achieving comparable accuracy to random forests and RNN-LSTM methods with significantly fewer trainable parameters (19,669 vs. 726,890). The study in [131] presents two detectors that extract communication signals from radar without needing channel estimation. A fully connected DNN uses principal component analysis for symbol prediction, while an RNN-based LSTM processes time series data. For radar waveforms (LFM and FMCW) and Rician fading communication channels, numerical results show excellent SER performance under 16-quadrature amplitude modulation (QAM) and quadratic phase-shifted keying modulations compared to traditional zero-forcing detectors. Moreover, the DNN outperforms the RNN at high SNR, while the RNN is more robust at low SNR. Meanwhile, [132] optimizes the CNN architecture from the RadComNet framework to balance processing complexity and accuracy. Using advanced convolution techniques and attention mechanisms, RadComNet achieves 89.87% accuracy with a compact network (178,000 parameters) and a prediction time that's 1.6 times faster than Inception-V3.

Regarding ISAC systems, the work in [133] proposes a CNN-based ISAC classification framework, which includes the primary module for training signals in the phase and quadrature domains from multiple modulation types and the second module trained with QAM constellations to help the primary module differentiate between 16-QAM and 64-QAM. Focusing on the LFM signal commonly used in radar and the communication dataset RadioML 2016.10a, the proposed approach can achieve a classification accuracy of more than 91%. In [134], a concurrent radar-communication waveform recognition network (WaveNet) is introduced to address large and attenuated ISAC signals and reduce interference from regular Wigner-Ville distribution, focusing on learning underlying radio features from time-frequency representations. This is done by integrating three modules into deep CNN architecture: 1) cost-effective feature awareness, 2) grouped-of-kernel-wise residual connections, and 3) dual asymmetric channel attention. Compared to RadComNet, the proposed WaveNet can enhance the classification accuracy by 1.5% while narrowing network sizes from 178,000 to 60,000 parameters.

On the other front, some recent efforts focus on developing intelligent wireless frameworks to assist advanced ISAC. For example, the authors in [135] introduce a DNN-based convolutional architecture for joint sensing and task-oriented communication, as shown in Fig. 11 This architecture includes two encoders for the transmitter and one decoder for the receiver. The transmitter uses one encoder for joint source coding, channel coding, and modulation, and the other for processing reflected signals from the sensing target. The receiver's decoder separates and reconstructs features necessary for image classification. With 5 training epochs using 50,000 data samples (20% for testing), the framework shows 97% accuracy for joint sensing and 88% for task-oriented communication. Later, they propose a multi-task DL paradigm



(b) Step 1: Edge device takes an image by its camera and sends it to the fusion center via TOC while using it for target sensing. **(c) Step 2:** Edge device collects the signal reflected off the target and transmits it to the fusion center via TOC



(a) System model design for joint sensign and TOC : Image data, initially transmitted for object recognition, is utilized for both image classification and target sensing.

Fig. 11. Joint sensing and task-oriented communication systems with multi-modal data collection and delivery [135]. Herein, the main goal of step 1 is to eliminate the need for additional probe signals, thus saving bandwidth resources and reducing the latency associated with target sensing. Meanwhile, in step 2, target sensing data is re-encoded at the transmitter for compact communication to the fusion center via TOC. A decoder at the receiver fuses image and sensing data to identify active transmitters. Using three DNNs, the system jointly trains two encoders and a decoder for multi-modal integration, enabling image classification and signal detection for dynamic spectrum access, including detecting passive transmitters or those beyond the spectrum sensing range.

in [136] to generalize communication through four DNN blocks, with two blocks for each transceiver. On the transmitter side, the first block encodes input data for source coding, channel coding, and modulation, generating a latent space representation for ISAC. The second block handles sensing by detecting targets or classifying their range based on returned signals. On the receiver side, the first block is in charge of recovering source information, while the other block validates semantic content. Experimental results reveal that this multi-task learning approach optimizes target sensing and data transfer, preserving semantic information and high-fidelity features. Notably, it can be adapted for joint sensing and task-oriented communication by eliminating data reconstruction and using a common task classifier, making it ideal for real-time decision-making, context-aware data sharing, and efficient resource management.

C. Discussion

From the review of the existing studies, it is clear that learning-based approaches, developed by advanced DL techniques, play an important role in assisting ISAC functions. While learning-based models such as CNN are highly used for AoA/AoD estimation to adapt to dynamic changes in communication behaviour and sensing targets, DNN, ANN, CNN, and LSTM networks are preferred in improving the classification accuracy of communication signals from sensing signals. However, each type of DL technique requires specific conditions to operate while offering unique features for system performance improvements:

- *AoA/AoD estimation:* Exploiting DL solutions is useful for unstable wireless channels due to high target mobility. For example, CNN-based models offer diverse improvements, such as handling Doppler effects by translating communication channels into 2D maps for image processing tasks [124]. Training offline CNN estimators can denoise signals without complex optimization [120], making them robust for systems with limited resources. However, this requires strict angle preprocessing and a reliable dataset. Also, customizing loss functions [123] can accelerate CNN training. Especially, employing complex-valued architectures can provide richer representations to improve AoA and AoD estimation [126], allowing online estimation of many targets but remaining less focused.
- *Signal classification and detection:* Maintaining accuracy in classification and detection tasks while reducing network parameters is challenging. For example, using a CNN-based ISAC classifier in [133] can differentiate between 16-QAM and 64-QAM with over 91% accuracy for 12 modulation types at low SNR but requires many parameters. Similarly, DNN and RNN-LSTM detectors in [131] can extract communication signals without channel estimation, but their performance depends on dynamic changes from noise and multipath fading channels. Improving signal transformation using techniques like Fourier-based synchrosqueezing transform [134] can enhance classification accuracy and reduce network size. For higher levels, future ISAC systems will not only include DFRC but also scale to wider functions, like joint sensing and task-oriented communication [135] or multi-tasking of sensing classification, information reconstruction, and semantic extractions [136], enhancing real-time decision-making and resource management.

Further research should address the gaps in AoA/AoD estimation frameworks, particularly adaptive RL frameworks. For example, a reinforcement meta-learner design with a two-layer learning and sensing algorithm in [137] can provide the most informative sensing location by using SGD-based learning for minimizing average prediction error and optimizing sensing decisions with a reinforcement meta-learner [138]. Developing innovative solutions for efficient classifiers/detectors in active wireless environments is urgent. This requires new features like lightweight models and flexible plug-in modules to adapt to various ISAC multi-tasking needs, contributing to future learning-based classifier/detector standards.

VI. LEARNING ALGORITHMS FOR SPECTRUM ALLOCATION, TIME ALLOCATION, AND POWER ALLOCATION

Aside from AoA/AoD estimation and signal detection/classification, sharing resources between sensing and communication functions in ISAC systems raises concerns about the reasonable allocation of spectrum, time, and power resources. Specifically, the first factor concerns how to ensure efficient spectrum usage by allowing sensing and communication systems to coexist without rigid predefined allocations. The second factor refers to the way that ISAC balances sensing and communication in shared time slots, dynamically adjusting schedules to meet real-time operational demands.

The third factor queries how to design intelligent power control so that ISAC systems allocate transmission power optimally, reducing interference while maximizing signal strength for both functions. Although conventional allocation techniques for spectrum allocation (e.g., static spectrum assignment, graph-based spectrum optimization, or game-theoretic spectrum sharing), time allocation (e.g., time division, round robin scheduling, optimization-based scheduling), and power allocation (e.g., water-filling algorithms, convex optimization-based power control, or heuristic approaches like iterative power adaptation) offer foundational solutions, they struggle in dynamic environments with fluctuating interference and resource demands. Specifically, spectrum allocation methods depend on predefined channel allocations and interference constraints and face high computational complexity when scaling up to large networks. Time allocation methods often require extensive computation to determine which time slots are fixed and dynamic for varying sensing and communication demands. Meanwhile, power allocation techniques require precise CSI, typically struggle with non-linear interference effects in large-scale ISAC systems, and in some cases, cannot balance sensing and communication power dynamically, failing to generalize efficiently across varying requirements.

In recent years, integrating learning algorithms such as ML, DL, or RL into wireless-based systems has emerged as a viable option for efficient resource allocation solutions, and so have ISAC systems. More precisely, DL-based RL models can optimize spectrum allocation dynamically, allowing ISAC systems to learn optimal channel usage without predefined rules. RNNs and attention-based models improve time scheduling by predicting resource demands and adjusting transmission slots accordingly. For power control, deep Q-learning and policy gradient methods provide efficient, adaptive power distribution by analyzing environmental feedback and adjusting allocation strategies in real-time. Unlike traditional conventional allocation approaches, DL techniques scale efficiently, optimize under uncertainty, and enhance resource allocation across complex ISAC deployments. For example, as the target is closer to the ISAC system, fewer resources, i.e., spectrum and power, can be allocated to the sensing function, and more resources are allowed to the communication function.

In this section, we will survey recent technical learning algorithms applied to ISAC systems, covering spectrum usage prediction and frequency allocation in Section VI-A, power level adjustments in Section VI-B, time allocation optimization in Section VI-C, and antenna allocation issues in Section VI-D. We will conclude with key insights in Section VI-E.

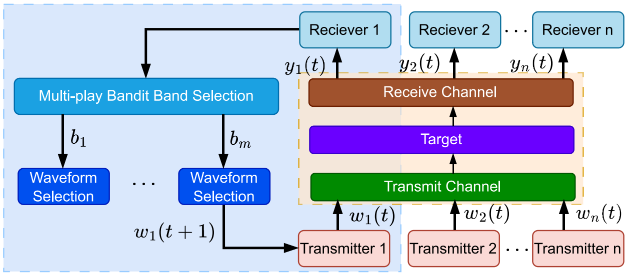


Fig. 12. Illustration of transmit/receive cycle for cognitive radio network where each node employs multiplayer bandit frequency selection paired with a "single-player" waveform selection algorithm [145] to choose a waveform and center frequency from library to transmit in each time step, thereby solving the problem of the interaction between the cognitive system and its environment. In particular, each node n independently selects a waveform $w_n(t)$, which is modulated by the environment and then returned as a waveform $y_n(t)$. Using the received energy in combination with the multiplayer bandit frequency b_m, the cognitive learner selects the next transmit waveform $w_n(t+1)$.

A. Spectrum Allocation

Instead of separating the function of radar and communication, exploiting their mutual benefit intelligently can flexibly deal with time-varying system demands from dynamic medium changes. The work in [139] proposes to use an RL-based framework to optimize both ranging and communication performance. In which, RL relies on a water-filling inner bound algorithm [140] to dynamically split bandwidth resources for exclusive communication and DFRC. In [141], a DRL-based approach addresses the challenge of non-cooperative coexistence between cognitive pulsed radar and proximity communication systems by first utilizing a double DQN to mitigate estimation bias and then exploiting a recurrent LSTM architecture to achieve stable training. Unlike [141], the work in [142] addresses the issue of radar spectrum sharing using the Thompson sampling in conjunction with a linear contextual bandit. Specifically, at each time step, the radar detects interference and creates a context vector for each agent to maximize rewards based on past actions, contexts, and rewards. The Thompson sampling algorithm then estimates the mean reward for each agent, along with an additional linear contextual bandit to model uncertainty and prevent overfitting due to limited observations. On the other hand, the work in [143] addresses mutual interference from cognitive radar observations by using a decentralized spectrum allocation approach. It uses an LSTM-based RL network to learn optimal actions from rewards in unknown mediums and double network to achieve the stable training cycle of the target network. Likewise, [144] develops an RL-based cognitive radar adaptive process by formulating MDP to enable the radar to learn from offline training data in each radar cycle and constructing the discount factor to model the selection for current rewards versus future rewards. Accordingly, the radar system with five contiguous 20-MHz frequency bands can effectively locate the interferer's position to maximize the use of adjacent bands to adapt to changes in target position and SINR quality to avoid the interference bands before the radar beam enters the interference area.

The learning-based approach in single-carrier ISAC systems lays the groundwork for improving radar sensing tasks, but multi-carrier radar transmission faces challenges related to frequency band usage and waveform selection. Inspired by this, an online ML method in [145] optimizes these aspects through a multiplayer multi-armed bandit (MAB) algorithm (see Fig. 12), which allocates frequency, and a single-player bandit model for waveform selection, enabling better radar tracking with reduced interference. Compared to the send-and-avoid and fixed power allocation schemes, this approach can converge to a better optimal solution under sublinear cumulative regret (5 times). Similarly, [146] proposes two RL solutions (independent coexistence and ISAC) for efficient spectrum sharing between communication and radar systems. For frequency shift keying (FSK), spectrum sensing predicts communication operations even under instability. For frequency-modulated continuous wave (FMCW), gNodeB identifies unused channels based on radar operations. Their simulations indicate that continuous wave radar maintains high accuracy in range and velocity estimation, even amidst interference at 40 GHz. In contrast, [147] tackles large-scale deployment with joint bandwidth and carrier frequency allocation through two methods: one using SCA and geometric programming for equal information acquisition, and another employing an RNN-based DL framework, which reduces computational time significantly. Differently, [148] addresses limited CSI collection by introducing cognitive multicarrier radar, using a partially observable MDP with PPO to extract environment states and guide the agent in decision-making, and a model-agnostic meta-learning approach to pre-train the policy network. Numerical results show that cognitive multicarrier radar achieves impressive learning capability and detection performance in congested spectrum environments, with an average detection probability of over 0.95 with 400 sampling times at the SNR of 10 dB.

B. Power Allocation

Power allocation is a common topic that aims to balance power distribution among communicating entities and so does ISAC. To jointly allocate frequency and power resources in dynamic ISAC systems, the work in [107] proposes an extended version of the model-based online learning algorithm by separating interference from the state transition function into interference caused by cooperative information sharing and disturbance. Compared to four benchmark scenarios (Markovian, deterministic, Poisson, and adversarial), the proposed algorithm demonstrates outstanding performance in terms of low and sub-linear constrained regret. Meanwhile, the work in [149] introduces a combinatorial online optimization framework using Lyapunov and model-free DRL to minimize power consumption while ensuring fairness under varying wireless channels and user arrivals. Simulation results show that this framework not only effectively combines the advantages of both methods but also significantly reduces transmission power consumption. In [150], a fast beam tracking framework for large-scale antenna array-based mmWave ISAC systems is proposed to enhance target location prediction using extended and unscented Kalman filters. In that, performance prediction, quantified by the posterior CRLB, serves as a metric for developing an RL approach to optimize sub-carrier and power allocation. Compared to the pilot-based scheme, the proposed beam tracking scheme achieves superior tracking performance and higher transmission rates, offering a better tradeoff under terminal interference.

Focusing on multi-cell multi-user mobile broadband networks, the work in [151] investigates a two-tier distributed spectrum sharing framework, where a multiagent algorithm with DDQN is proposed to handle power allocation, in conjunction with null-space based waveform projection method to mitigate interference from MIMO radar systems. Simulation results indicate that the proposed DDQN effectively adjusts power allocation to remain within the allowable radar interference limits. Different from [151], the work in [152] handles interference management issues by first transforming it into a functional optimization with stochastic constraints, then using transfer learning to extract knowledge or patterns from different domains/tasks, and finally distributing unsupervised learning models into base station to quantize and optimize system resource allocation in respond to a change in local CSI. Numerical results demonstrate that this method achieves spectral efficiency comparable to the weighted MMSE algorithm and doubles the performance of random and equal power allocation methods. On the other hand, the work in [153] studies a joint sub-channel assignment and power allocation in downlink multi-cell OFDM, where a DRL approach is proposed by exploiting DDQN to speed up estimation convergence for the sub-channel assignment policy and DDPG to map channel gain to power allocation action. By virtue of simulation outcomes, it is demonstrated that the proposed DRL approach achieves a stable loss function after 5000-time slots. Especially, the sum rate produced by the DRL-based approach provides a significant rate improvement of approximately 6 and 10 kbps when compared to two benchmark schemes of distributed learning and random, respectively.

Extending the scope with cooperative UAV networks, the work in [154] studies an ISAC-enabled multi-UAV cooperative system and proposes to solve the non-integer non-convex mixture optimization of joint transmission power, resource allocation, and UAV trajectory by combining DNN to approximate the policy and PPO-based MARL to improve agents training efficiency through learning historical patterns. This proposed learning scheme provides effective convergence performance after 1000 training episodes and an average probing score better than benchmark policies like random, single-agent learning, multi-agent deep deterministic gradient, and PPO-based Kullback-Leibler divergence. Meanwhile, the work in [155] formulates and proposes to solve the problem of maximizing the minimum weighted spectral efficiency among multi-ISAC aerial platforms by two sequential approaches. The first method uses a centralized SAC algorithm to directly sidestep the problem of sequential decision-making. The second method is to employ decentralized multi-agent SAC by defining each UAV as an agent of SAC and using two critic networks per agent to observe only its state and local medium information. Numerical results unveil the efficacy of the proposed SAC algorithm in obtaining high training speed and weighted spectral efficiency.

To achieve V2X applications, integrating UAVs with ISAC functions is crucial for accurate simultaneous location and mapping, surpassing GPS capabilities. The work in [156] proposes a device-free ISAC system model for precise control and coordination in distributed UAV manners. As shown in Fig. 13, each UAV senses vehicle users on urban roads, capturing high-precision locations and relaying information to

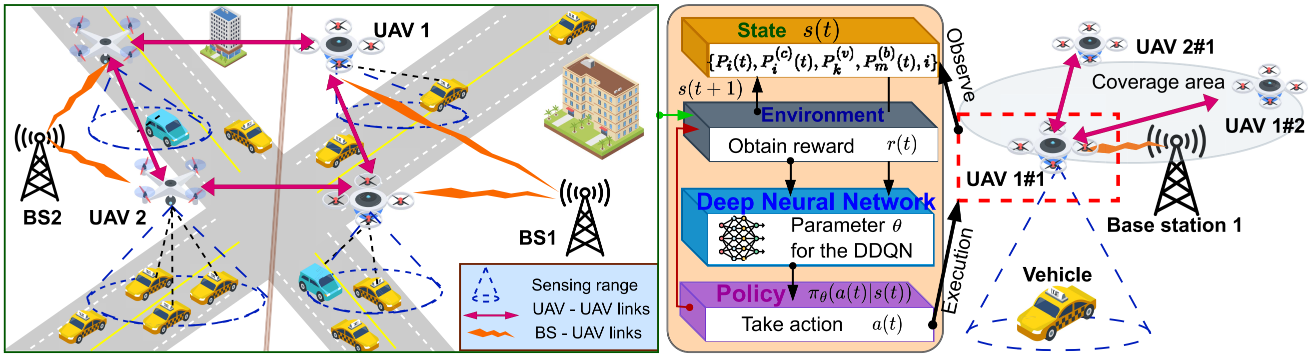


Fig. 13. Device-free ISAC system based on distributed UAVs and multiple access channels with mono-static sensing [156]. The left side depicts the specific application, where UAVs sense vehicles on urban roads to gather precise location data and provide sensing services while relaying the sensed information to the base station via MAC channel. The right side demonstrates how the designed DRL framework performs to improve the ISAC efficiency, where its training process stems from observing state space *s(t)* at time *t*, including the position of the *i*-th UAV $P_i(t)$, communication transmission power of the *i*-th UAV $P_i^{(c)}(t)$, position of the *k*-th UAV $P_k^{(v)}(t)$, and the base-station coordinate $P_m^{(b)}(t)$. Then, it outputs the reward function *r(t)* based on the environmental observation, followed by estimating the parameter of the evaluation network θ, and finally updating it with the action policy $\pi_\theta(a(t)|s(t))$ to make the decision *a(t)* based on *s(t)*. Once this training is completed, saved, and integrated into UAVs, UAVs can be used in actual scenarios.

base stations. Then, a triple objective optimization problem-data rate, radar mutual information, and energy efficiency-is formulated, involving communication power control, motion selection, distance between UAVs, and vehicle's sensed information. This problem is solved using a multi-agent RL solution to be decomposed into multiple reward functions, using centralized training and distributed execution for NNs to achieve swarm UAV collaboration and autonomous decision-making, with double DQN maximizing long-term rewards. So, what will be received by this proposal? The answer is that it can achieve 1.5 times higher data rates, double the number of sensing users served, and reduce the energy-saving gap by 0.5 compared to benchmarks. In [157], another aspect of RL using the Q-learning algorithm is considered to generate dynamic modulation patterns and adjust transmit power for radar signals under uncertain and potentially hostile conditions of electronic intelligent systems. Analysis of the time-frequency modulation signal provided by electronic intelligence systems shows that this approach can provide the low probability of intercept while using synthetic aperture radar imaging tasks for detection performance evaluation.

C. Time Allocation

Aside from spectrum and power allocation, addressing time allocation in ISAC systems also receives considerable attention, aiming to improve resource efficiency, enhance performance, and manage interference by scheduling strategic time slots for each function. Therefore, the authors in [158] introduce an intelligent real-time DFRC vehicle system to achieve flexible bandwidth utilization and quick responses to collision risks in bad weather. With a simple DQN design, the DRL algorithm determines the optimal policy for operating communication or radar mode with minimal miss detection probability. However, it relies on a perfect assumption of no false alarm or miss-detection events when autonomous vehicles operate in radar mode. Therefore, another efficient framework for ISAC with three tiers is introduced in [159] that uses MDP for adaptive radar/communication configuration based on current medium status, DRL techniques for optimal policy without prior medium information, and transfer learning to accelerate training when vehicles enter new environments. Following that, this framework can reduce the obstacle miss detection probability up to 67% compared to pure DRL methods. Similar to [158], [160] designs a GNN-based multi-agent message-passing DRL framework to balance performance with minimal prior vehicle model knowledge, along with a bit correctness rate metric to guide data transmission timing and recipients. Experimental results show that this GNN-based DRL not only outperforms non-learning algorithms but also robust learning communication with task reward only. In [161], the authors outline a novel constrained DRL approach using deep Q-learning to tackle the task of adaptive time allocation in tracking and communication of the ISAC system. By testing with a simple DQN design of two hidden layers and 64 neurons per layer, the proposed DRL can fully exploit the dwell time allocation protocol to speed the sum-rate performance up 20% compared to the fixed dwell time allocation benchmarks.

Towards large-scale deployment, the work in [162] introduces single and multi-agent RL frameworks for vehicle agents in ISAC systems. This aims to optimize radar coordination and minimize the age-of-information of data queues. In single-agent scenarios, MDP-based RL reduces the age of data packets while enhancing radar detection under challenging conditions. In the multi-agent scenario, a model-free policy-based RL method using medium access control for multi-agent settings prioritizes channel use, facilitating decentralized decision-making. Experimental results indicate strong performance of these RL algorithms with limited environmental info, surpassing benchmarks. To enable MIMO ISAC systems, [109] designs an autoencoder architecture comprising six blocks built in FNNs for joint communication and sensing (see Fig. 14), facilitating end-to-end learning. Numerical results demonstrate improved SER over baseline amid hardware impairments in the transmitter's antenna array.

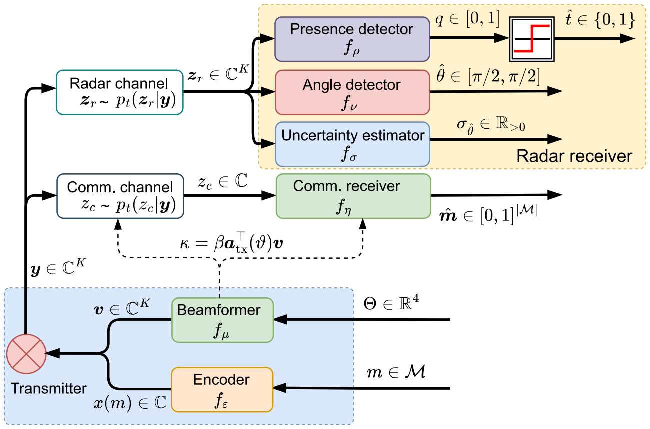


Fig. 14. A block diagram of a novel AE design for MIMO-ISAC systems shows the radar receiver co-located with the transmitter and communication receiver remotely positioned [109]. Initially, the transmitter relies on knowledge of angle information Θ and message m to generate the transmit beamformer $\boldsymbol{v}$ and the modulated message $x(m)$, transmitting them together with the complex signal $\boldsymbol{y}$ across multiples antennas. At the co-located radar receiver, it processes the received signal $\boldsymbol{z}_r$, whereas the remote receiver observes the signal z_c via a single antenna. Then, the communication receiver directly recovers the transmitted message $\hat{\boldsymbol{m}}$ from z_c using available CSI κ, expressed as a product of a complex Gaussian channel β, angle information function $\boldsymbol{a}_{\text{tx}}^\top$ with the AoD range ϑ, and $\boldsymbol{v}$. Meanwhile, $\boldsymbol{z}_r$ is fed into three AE blocks at the radar receiver to determine target presence with probability q and estimate AoD $\hat{\theta}$ with an uncertainty $\sigma_{\hat{\theta}}$.

Additionally, [163] employs non-orthogonal multiple access (NOMA) protocols to achieve simultaneous communication and radar sensing functionality by scheduling target users with a set of communication users and exploiting the beamforming weight to fine-tune the energy concentration in the direction of NOMA users and the target. To quantify radar activity, a sensing mutual information neural estimation unit is introduced to estimate the mutual information from the sensing, followed by the proposal of an end-to-end learning architecture to optimize transmissions of transmitter-receivers. Numerical results validate the framework's effectiveness in enhancing both sensing and communication rates.

D. Antenna Allocation

Dynamic changes in road conditions pose safety challenges to autonomous vehicles. Adapting to continuous environmental transitions, such as those between urban and suburban areas raises concerns about allocating an optimal number of antennas for automotive MIMO radar systems. Inspired by this, an investigation on the state-action-reward-state-action RL (a.k.a., modified connectionist Q-learning) for dynamically distributing two transmit subarrays of communication and sensing has been put forward in [104], emphasizing on improving the accuracy of the DoA estimation for all targets while maximizing the SNR of communication data streams. Considerations of 10 transmitting antennas and 6 receiving antennas show that the proposed RL can converge within 100 trainings and achieve the optimal compromise between the CRB of DoA estimation and the SNR received by the communication users. Unlike [104], the work in [164] seeks to obtain joint/share resource spectrum access and reduce the hardware cost for ISAC systems. Specifically, the authors in [164] propose a new strategy to partition consecutively spanned antennas into smaller subarrays while introducing an alternating optimization approach to jointly select the best subarray and construct hybrid analog-digital beamformers. Experimental results show that the proposed method can provide 80% antenna selection accuracy and achieve slightly higher performance than the fully digital beamforming method.

On another front, the authors in [165] develop a DRL method for intelligent automotive DFRC systems to adaptively select a small subset of transmit antennas while configuring the phase shifter so that the beam energy is directed to the communicating user and target of interest with minimal interference on the other radars. However, the action space in this DRL design increases exponentially with the number of antennas and quantization bits. Therefore, they proposed to combine Wolpertinger and neighbourhood component strategies to project the large action space into a smaller space to maintain the desired performance in [166]. This combination reconciles the strengths of DQN and DDPG in agent-critic networks, i.e., allowing the agent to learn a more accurate representation of the optimal policy than traditional table methods and achieving a manageable training process over time. Compared to the traditional optimization-based approach, the proposed DRL achieves the same desirable beampattern for a uniformly linear antenna array and accurate alignment for the main lobes' direction toward the sensing target and communication receiver with the lowest transmission power.

E. Discussion

The review of state-of-the-art algorithms for spectrum, power, time, and antenna allocations shows that learning algorithms offer significant advantages in resource allocation, including reduced computational complexity, efficient resource balance, and real-time decision-making. However, they also require specific conditions and present unique challenges:

- *Spectrum allocation:* While LSTM networks are typically used to approximate reward functions with low error [143], RL/DRL stabilizes training and reduces learning dimensionality [141]. However, factors like limited CSI acquisition, interference, and Doppler shifts highly degrade these learning algorithms' quality. Thus, future RL/DRL research can combine the possible techniques, such as an iterative selecting PPO combined with model-agnostic meta-learning [148], water-filling inner bound [139], and MAB [145].
- *Power allocation:* RL/DRL techniques like MDP, DDPG, and DDQN are popular approaches in reducing interference impacts. For example, combining stochastic constraints and transfer learning improves the transmit beamforming [152]. Null-space-based waveform projection minimizes radar interference [151], while model-based online learning classifies interference from state transitions [107]. However, interference issues arise from both radar and communication, as well as

hardware design. Therefore, future RL/DRL research should combine multiple access techniques.

- *Time allocation:* There is a common intuition that ISAC-based learning frameworks primarily focus on DRL [158] and RL-based approaches like MDP with transfer learning [159], deep Q-learning [161], and model-free policies [162] to switch between radar sensing and communication. Some studies also use FNNs [109] and GNNs [160]. However, effective switching modes in ISAC are highly dominated by channel conditions [159], especially with hardware constraints. Accordingly, future research should further explore rate-splitting/non-orthogonal multiple access for optimal time and spectrum use [163].
- *Antenna allocation:* Only a few investigations are using RL techniques for antenna allocation [104], [164], [165], with limited breakthroughs. The main approach separates the transmit antenna array for communication and sensing, using strategies like sparse antenna decomposition [104] or neighborhood component combined Wolpertinger [164]. RL then allocates antennas based on dynamic wireless channels and environmental factors to address performance trade-offs in sensing and communication subarrays, but this is not a logical solution. For example, joint antenna allocation with digital beamforming [165] achieves only around 80% performance. Thus, future RL research should consider these resource factors.

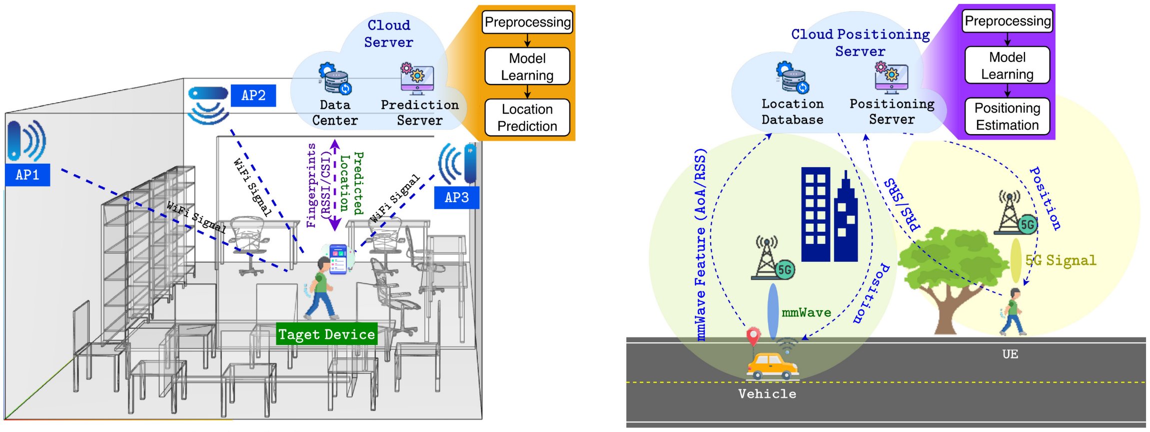
VII. LEARNING ALGORITHMS FOR WIRELESS SENSING

Wireless sensing for localization leverages communication signals, such as those from Wi-Fi and cellular networks, which are collected by access points and processed to determine position. Traditionally, wireless localization has relied on model-based techniques that estimate location based on predetermined geometric principles or signal propagation characteristics. Common traditional methods include range-based approaches that utilize direct measurements of signal properties to calculate distances or bearings relative to known anchor points. Examples include Time of Arrival (TOA) [167], Time Difference of Arrival (TDOA) [168], Angle of Arrival (AOA), and Received Signal Strength Indicator (RSSI) [169]. These methods often face challenges in real-world environments; for instance, TOA techniques may require sophisticated filtering algorithms to identify and mitigate non-line-of-sight (NLOS) errors in obstructed indoor spaces [167], while RSSI approaches often need enhancement algorithms, such as averaging selected peak values, to counteract multipath interference and improve accuracy [169]. Algorithmic refinements have also been developed, such as non-iterative algebraic methods for TDOA, which can estimate a moving source's position and velocity without requiring initial guesses, thereby avoiding convergence issues common in iterative solutions [168]. The estimated distances or angles from these range-based methods are typically fed into algorithms like trilateration or triangulation to compute a final position estimate. Alternatively, range-free approaches estimate position based on connectivity data or proximity to anchors, without requiring precise distance or angle measurements. Examples include centroid localization and hop-count based methods like DV-Hop. These methods also continue to be refined; for example, enhanced DV-Hop techniques for IoT sensor networks improve accuracy by incorporating multi-node distance estimation and hop loss metrics within multi-objective optimization frameworks [170]. While foundational and subject to ongoing research, these traditional methods generally depend on explicit environmental or signal models and known anchor locations, which can limit their accuracy and robustness in complex, dynamic scenarios subject to significant interference, blockages, and multipath effects. This contrasts with modern learning-based approaches, discussed next, which leverage data-driven techniques to adapt more effectively to such challenging conditions.

While foundational, traditional localization methods rely on predefined environmental and signal models, limiting their accuracy and robustness in complex, dynamic scenarios prone to multipath, blockage, and noise. In contrast, modern learning algorithms (including AI, ML, DL) offer powerful, data-driven alternatives. These methods automatically extract complex, non-linear patterns [171] from raw signal data without explicit models, thus leading to significant advantages: improved accuracy, enhanced resilience to environmental dynamics (like interference and obstacles), and continuous adaptation through ongoing learning [172]. For instance, ML-based Wi-Fi fingerprinting is common for indoor localization, while the capabilities of 5G and beyond networks (higher bandwidth, advanced beamforming) make them increasingly suitable for high-accuracy outdoor positioning (Fig. 15). This section reviews recent AI-driven advancements across key areas, including DL-based positioning, adaptive algorithms, ML integration with signal processing, sensor fusion, and robust ML models for challenging settings.

A. DL-Based Positioning and Localization

In recent years, DL has revolutionized positioning and localization techniques, particularly in the context of advanced wireless communication systems such as mmWave and massive MIMO networks. These systems, critical for next-generation networks like 5G and beyond, face challenges in positioning accuracy due to complex channel conditions, particularly in NLoS environments. Several innovative DL-based approaches have emerged, pushing the boundaries of what is achievable in accurate and reliable positioning in these challenging scenarios. A pioneering end-to-end (E2E) learning framework [173] integrates autoencoder (AE) architectures to jointly optimize transmitter precoding and receiver-side processing for mmWave downlink positioning. By addressing signal design and estimation together, it introduces specialized neural networks and loss functions to enhance robustness against hardware impairments. This approach outperforms conventional methods in RMSE for AoD and CRB across various SNRs, demonstrating how DL can improve joint beamforming and positioning accuracy, even under non-ideal conditions. It marks a significant advancement in leveraging DL for positioning in complex, hardware-limited scenarios.



Indoor Localization

Outdoor Positioning

Fig. 15. System models for indoor localization and outdoor positioning using learning algorithms. The left side illustrates a Wi-Fi-based indoor localization framework where access points communicate signal features to a cloud server for preprocessing, model training, and location prediction of the target device. The right side demonstrates a 5G-based outdoor positioning system leveraging signal features such as PRS and mmWave characteristics. The cloud positioning server preprocesses and processes data from user equipment to estimate precise locations using ML models.

The challenges of mmWave indoor localization are addressed in [174] through DL techniques applied to beamformed fingerprints to achieve state-of-the-art accuracy with an average error of 1.78 meters in NLoS conditions. By integrating CNNs, hierarchical models, and temporal convolutional networks into sequence-based architectures, the proposed method enables precise tracking besides enhancing localization prediction over time with historical data. This approach provides accurate real-time positioning, dynamic updates, and improved system adaptability. The availability of a public mmWave dataset for tracking further contributes to the field to support future research and development in this topic. In the context of outdoor localization, the work [175] studies a deep convolutional Gaussian process (DCGP) regression model to improve localization accuracy in mmWave outdoor environments. The convolutional structure of DCGP enables uncertainty estimation of location predictions in NLoS scenarios. By constructing mmWave beamforming images, the model captures location features effectively, and experimental results indicate that DCGP outperforms traditional CNNs in terms of the median location error and 95th percentile location error.

Building on similar themes, the work [171] introduces a DL-based approach for indoor fingerprinting-based localization using beam covariance matrix (BCM) images. This method uses a beam covariance learning (BCL) network, built upon the ResNet architecture, to capture location-specific patterns from multi-channel BCM images. To handle several challenging tasks (such as location classification and simultaneous location-and-orientation classification) and improve localization accuracy in NLoS environments, the BCL model leverages residual blocks. Through the superior performance attained by intensive simulations, this model demonstrates how DL techniques, such as residual networks, can significantly enhance the accuracy and reliability of positioning systems. In [176], DeepFi, an innovative indoor positioning system using CSI collected from access points (APs), is developed with a cutting-edge DNN architecture and trained with a greedy learning algorithm to optimize location accuracy while reducing computational complexity. The method involves training with a stack of restricted Boltzmann machines in a layer-by-layer form, followed by fine-tuning via backpropagation. Moreover, by exploiting a probabilistic data fusion mechanism with radial basis functions, the method performs more accurately than existing received signal strength indicator (RSSI) and CSI-based techniques. Another interesting approach is ConFi [177], a deep CNN-based indoor Wi-Fi localization model learning complex localization patterns from CSI data. The architecture of ConFi, which includes multiple convolutional layers and fully connected layers, captures correlations across time, frequency, and antenna domains, accordingly enhancing the depth of the deep network for learning underlying features more effectively while maintaining computational efficiency. In addition to improving localization accuracy compared to existing Wi-Fi-based methods, ConFi highlights the potential of CNNs for developing more accurate and scalable indoor localization systems than traditional ones.

Following similar advancements in leveraging learning-based methods for enhanced localization, some recent works [178], [179] offer refinement mechanisms in fingerprint-based and DL-enhanced localization techniques. In [179], FCLoc leverages robust principal component analysis (RPCA) to mitigate noise and handle environmental variations in Wi-Fi RSSI data, a common challenge in traditional K-nearest neighbors (KNN)-based systems. By integrating RPCA with a stability assessment mechanism for APs, FCLoc dynamically weights APs based on their reliability to maintain high accuracy, even when certain APs become unavailable. By ensuring noise resilience and adaptability, this approach makes FCLoc robust in highly variable environments. The iPos-5G system [178] leverages 5G New Radio (NR) CSI to enhance indoor positioning accuracy using unsupervised learning. At its core, an amplitude-phase deep AE processes CSI data to extract meaningful features without manual intervention as illustrated in Fig. 16. By integrating a modified radial basis function for similarity calculation and a unique probability fusion function, iPos-5G improves precision and robustness

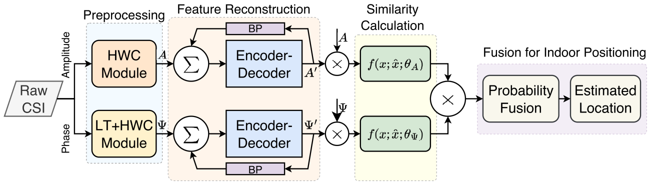


Fig. 16. An overall architecture of iPos-5G [178] with four parts: quality control module for raw CSI data preprocessing, amplitude-phase deep autoencoder network for training and feature reconstruction, radial basis function for similarity calculation, and probability fusion module for positioning, where LT: liear transformation and HWC: a combination module with Hampel identifier, wavelet filter, and cross-correlation detector.

across diverse environments. These advancements enable stable, high-accuracy positioning via significant reductions of mean absolute error and standard deviation. This work highlights how deep AE and optimized similarity functions can refine localization performance under real-world conditions, thus exemplifying the evolution of ML-based localization for resilient and high-performance indoor navigation.

For 3D user positioning, the work [180] utilizes angle-delay channel power matrix (ADCPM) fingerprints processed and learned by a 3D CNN. Advanced techniques like convolution refinement and extended Inception modules enhance multi-scale feature extraction to increase localization accuracy and reduce computational complexity. A regression module with global average pooling further optimizes performance, thus showcasing the capability of 3D CNNs in handling complex data for precise localization in massive MIMO-OFDM systems. In [181], a multi-carrier cumulative DL model enhances CSI-based positioning by treating each CSI of subcarriers as an independent fingerprint to reduce dataset complexity. An LSTM-based architecture accumulates features across subcarriers, thus yielding accurate positioning with low average error assessed in both LoS and NLoS conditions, even with limited training data. Recently, the work [182] introduces Monte Carlo dropout to estimate uncertainty in CNN-based localization. As a Gaussian process (GP) approximation, it provides confidence intervals efficiently, offering a robust, low-complexity alternative to Bayesian networks, accordingly improving real-time localization in variable wireless environments.

Further advancing ISAC capabilities in vehicular contexts, Cai et al. [183] propose a novel uplink sensing scheme utilizing OTFS modulation and a LSTM model. Their approach focuses on initial parameter estimation via OTFS followed by LSTM-based trajectory prediction from historical channel data, showcasing a novel application of LSTM for robust vehicle tracking in high-mobility ISAC scenarios.

B. Adaptive Algorithms and Feature Learning for Localization

Recent advancements in adaptive algorithms and feature learning for indoor localization have focused on enhancing positioning accuracy through novel approaches to signal adaptiveness and feature extraction within DL frameworks. The self-adaptive weighted K-nearest neighbor algorithm [184] exemplifies adaptiveness by dynamically adjusting the K-value in KNN calculations based on signal strength, unlike traditional fixed K-value models. This real-time flexibility improves accuracy by 21% compared to conventional methods, especially important in dynamic indoor environments with variable signal strengths. Complementing this, the work [185] addresses challenges in CSI-based fingerprinting using an attention-augmented residual CNN. By combining residual blocks and attention mechanisms, this model captures local and global CSI features simultaneously to improve learning efficiency and generalize across diverse indoor environments without retraining. Reformulating trajectory tracking as a denoising problem enhances robustness and accuracy, consequently making it suitable for dynamic conditions and variations in inertial measurement units (IMUs). To tackle environmental and channel imperfections, an adaptive deep model using a Transformer architecture with attention mechanisms is deployed in [186]. Incorporating absolute positional encoding and a global context token enables this model to capture structural information within CSI to better adapt to spatial and temporal changes without relying on convolution or recurrence. Validated through extensive simulations, this approach demonstrates scalability and robustness in centralized and distributed antenna systems, accordingly highlighting the importance of adaptive feature learning for localization. Addressing the labor-intensive nature of Wi-Fi RSS fingerprint localization, the work [187] introduces a crowdsourcing-based solution leveraging a Gaussian Process (GP) model with an adaptive mean function. This approach interpolates positions based on timestamp and motion data to reduce localization errors by 72% in sparse data scenarios while simplifying the site survey process, consequently enhancing scalability and practicality. Collectively, these innovations advance localization accuracy and adaptability, accordingly establishing a foundation for real-world applications that require continuous adjustment to dynamic and unpredictable indoor environments.

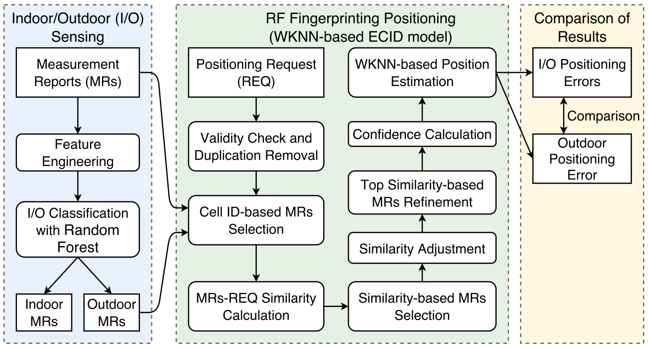


Fig. 17. The processing flowchart of integrated wireless sensing and positioning for cellular network in [188], in which random forest is deployed for indoor/outdoor sensing and WKNN is leveraged for RF fingerprinting positioning.

C. Integration of ML and Hybrid Techniques for Enhanced Localization

Hybrid approaches that combine model-driven methods with ML are advancing localization accuracy and robustness in complex environments. For example, the work [189] addresses mmWave-based 3D localization and channel estimation in vehicular networks using a hybrid strategy tailored for diverse LoS and NLoS conditions. A modified multidimensional orthogonal matching pursuit algorithm enables high-resolution channel estimation with reduced complexity by extracting essential parameters such as direction-of-departure and direction-of-arrival. Building on this, PathNet, a cutting-edge-designed deep network, classifies channel paths, isolating LoS and first-order NLoS paths critical for accurate localization. This classification improves data reliability, thus reducing localization errors and enhancing classification accuracy compared to conventional deep models. To further refine these estimates, ChanFormer [190], a Transformer-inspired model, applies attention mechanisms to fine-tune position predictions. This attention-based refinement achieves sub-meter accuracy in LoS conditions and significantly boosts performance in NLoS scenarios. Combining ML components with traditional signal processing also provides advantages like denoising, outlier removal, and environmental bias mitigation. Furthermore, the use of realistic vehicular channel data reinforces the practical applicability of framework to accordingly offer a robust, adaptable solution for vehicular networks. Similarly, [188] explores an ML-based framework to optimize integrated indoor and outdoor positioning in cellular networks, enhancing context-aware services. A random forest classifier is employed for indoor/outdoor scenario classification, paired with a weighted K-nearest neighbor model for precise positioning. By utilizing contextual data and denoising techniques, the framework reduces RF fingerprinting errors by 4% compared to conventional methods. Its scenario-aware design allows dynamic fine-tuning for increased localization accuracy, especially in mild indoor settings. These advancements underscore how ML-based hybrid methods can enable responsive, context-aware positioning systems, thus enriching user experiences through improved localization and service responsiveness.

D. Integration of Wi-Fi and Sensor Fusion for Localization and Sensing

The integration of Wi-Fi and sensor fusion technologies has advanced localization and sensing, with applications in human activity recognition and autonomous navigation. A recent study [191] introduces a Wi-Fi sensing framework combining learning-based and hybrid methodologies to address challenges in human activity recognition and object tracking. Leveraging DNNs with residual connections and attention mechanisms, the framework processes complex signal data with high accuracy and adaptability in diverse environments. Additionally, it optimizes data processing and communication parameters, such as bandwidth and sampling rates, to establish effective and replicable benchmarks for real-world Wi-Fi sensing applications. Complementing this, WiSion [192] focuses on state estimation for micro aerial vehicles (MAVs) by fusing Wi-Fi and IMU data to estimate six degrees of freedom (6-DoF) without relying on fixed Wi-Fi access point positions. This infrastructure independence enhances flexibility in dynamic environments. Advanced processing modules in WiSion extract essential features like range, angle-of-arrival, Doppler shift, and displacement from Wi-Fi signals, which, combined with nonlinear optimization, correct IMU drifts and significantly reduce positioning and orientation errors. This approach outperforms recent models like WINS [193], particularly in indoor environments with obstructions and limited access points. In summary, these studies highlight the potential of combining Wi-Fi with sensor fusion to enhance localization and sensing. The learning-based Wi-Fi framework demonstrates excellent performance in activity recognition and object tracking by tackling real-world challenges in data processing and adaptability. Meanwhile, WiSion demonstrates how Wi-Fi and IMU data fusion enables precise, real-time state estimation for MAVs, to offer robust navigation in environments. Addressing both algorithmic and practical issues, these works contribute to the development of reliable, versatile applications in human-centered sensing and autonomous navigation.

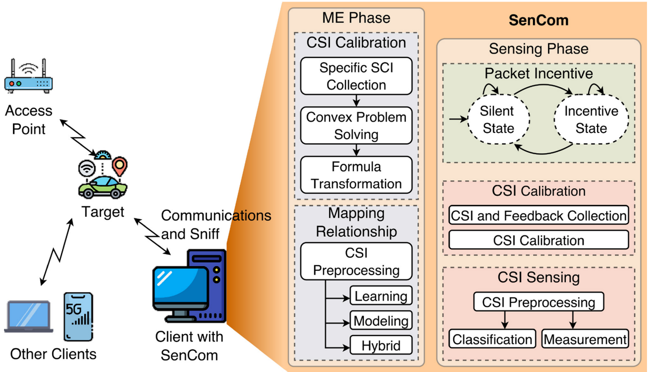


Fig. 18. An overview of SenCom's architecture in [194] includes two main phases: mapping establishment (ME) phase and sensing phase, wherein CSI of the wireless channel from AP is collected by sniffing transmissions to various clients.

In the context of data fusion to learn complex patterns, recent works [194], [195] demonstrate the potential of Wi-Fi and multi-sensor fusion to enhance the robustness and adaptability of localization and sensing systems. In [194], SenCom manipulates the existing Wi-Fi infrastructures to enable high-accuracy sensing tasks, such as human activity recognition and respiration monitoring without requiring hardware modifications. By employing innovative CSI calibration techniques, as illustrated in Fig. 18, including a transformation formula for MIMO compatibility and a compensation formula to manage beamforming effects, SenCom achieves consistent data dimensionality through a fitting-resampling scheme. When integrated into ISAC systems, SenCom attains a robust activity recognition rate of 94.4%, a precise person identification rate of 97.6%, and a highly accurate respiration monitoring capability with an error margin of less than 2 beats per minute. In contrast, the DFF-Loc framework [195] combines 5G NR synchronization signal block and Wi-Fi data to tackle challenges in indoor positioning. A multi-module approach in DFF-Loc integrates an extended Kalman filter for signal dynamics, a backpropagation NN for prepositioning, and an improved particle filter with dynamic weighting for precise final positioning. This layered methodology improves accuracy over conventional 5G and Wi-Fi-based positioning by handling both non-linear signal relationships and outlier data in DFF-Loc. Interestingly, the prompt adaptability of DFF-Loc to varying environments, combined with its integration of diverse signal sources, presents a promising solution for real-time smartphone navigation in challenging indoor settings, thus illustrating a pioneering step in sensor fusion applications for next-generation navigation systems.

E. Advanced Learning Techniques for Robust Localization in Complex Environments

Achieving accurate sensing and localization in dynamic environments requires advanced techniques that adapt to changing conditions while maintaining efficiency. One innovative approach involves integrating quantum machine learning (QML) with commercial Wi-Fi devices for indoor human gesture recognition [196]. This work introduces a quantum neural network (QNN) optimized through Bayesian optimization for automating the typically complex design and tuning processes of QNNs. The simplified development process enables the sensing model to achieve over 80% accuracy in human pose recognition with minimal trainable parameters, thus emphasizing the huge potential of QML for practical applications. Validating QML within ISAC systems advances quantum-enhanced localization and tracking, thereby forming a new standard for gesture recognition in indoor environments. To overcome less ambiguous location fingerprints from Wi-Fi CSI data, a sophisticated device-free localization framework [197], depicted in Fig. 19, is developed by exploiting a multilayer extreme learning machine (ML-ELM). By transforming raw CSI data into a unique feature space, ML-ELM fine-tunes hidden layer parameters as effective fingerprints to embed location information that enhances distinctiveness and reduces ambiguity in fingerprint matching. The deployment of non-iterative, multi-scaling representations in DFL enhances feature extraction and pattern learning, accordingly contributing to more reliable and efficient indoor positioning. Similarly, [198] proposes an innovative indoor localization method using multi-beam 5G NR signals in the sub-6 GHz band. A CatBoost-based algorithm, optimized with tree-structured Parzen estimators, achieves 1.06-meter localization accuracy, a 48% improvement over single-beam methods. This work highlights the potential of standard signal strength metrics and advanced learning frameworks to enhance precision while maintaining simplicity, thus offering scalable and efficient solutions for dynamic environments.

Generative AI (GenAI) is revolutionizing wireless sensing in ISAC systems by employing advanced models such as GANs and VAEs to synthesize realistic sensor data, mitigate data scarcity, and improve system performance [199]. By modeling underlying data distributions, GenAI enhances environmental perception and enables efficient data fusion, fostering smarter, context-aware communication. For instance, a GenAI-assisted human flow detection system has been proposed to leverage CSI for accurately estimating human movement parameters like velocity, acceleration, time of flight, and DoA [200]. This system integrates a unified weighted conditional diffusion model to denoise estimation results and resolve ambiguities in DoA spectrum, achieving a remarkable 91% accuracy in detecting subflow sizes [200]. A different study introduces a novel generative model based on a conditional variational autoencoder (cVAE) with a sequential gate self-attention mechanism, which synthesizes high-quality virtual RSSI maps for indoor localization [201]. This approach enhances positioning accuracy by reducing Euclidean distance errors by over 40% while significantly decreasing the need for extensive field measurements [201]. These advancements clearly demonstrate the transformative power of GenAI to revolutionize wireless sensing and ISAC applications.

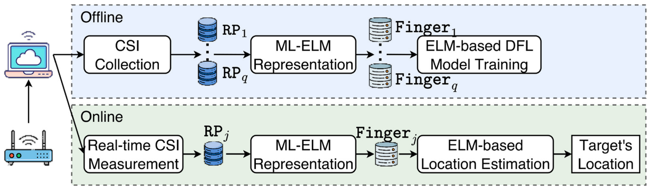


Fig. 19. An overall framework of ML-ELM-based DFL in [197] consists of two phases: offline phase is specially designed with ML-ELM to generate weight-based fingerprints and reduce fingerprint similarity's effect on the localization performance, and online phase estimates target's location using the trained DFL model.

F. Discussion

The application of learning algorithms in wireless sensing presents distinct advantages, faces notable shortcomings, and points towards specific future directions:

- Learning algorithms significantly enhance wireless sensing by improving accuracy (the precision of sensing tasks like localization [175], [192] or activity recognition [191], [194]) and robustness (the ability to maintain performance despite challenges like NLoS, noise, or interference [190]). This is achieved through advanced feature extraction from signals like CSI [176] or mmWave beams [189]. These algorithms enable adaptability to dynamic conditions, facilitate effective sensor fusion for richer environmental perception, and allow exploration of cutting-edge techniques like GenAI for tasks such as synthetic data creation, and QML for potentially new, highly efficient sensing capabilities [199], [200].
- Key challenges of existing wireless sensing methods include a heavy reliance on large, high-quality datasets which are difficult to obtain [188], the high computational complexity of many advanced models limiting real-time deployment on constrained devices [189], sensitivity to environmental variations requiring robust adaptive mechanisms [202], and inherent privacy concerns associated with collecting granular sensing data. Some techniques also face limitations in specific scenarios (e.g., 3D, dynamic environments [197]) or depend on immature hardware [196].
- Future work for learning-based wireless sensing should focus on developing data-efficient methods, including

data augmentation and leveraging GenAI [201], to overcome data scarcity. Creating lightweight, computationally efficient algorithms via techniques like model compression is crucial for practical deployment [195]. Enhancing model robustness and adaptability through continuous learning and advanced adaptive algorithms is needed. Further exploration of novel paradigms such as QML [196] and systematically addressing privacy concerns through techniques like federated learning are also essential for progress.

VIII. LEARNING ALGORITHMS FOR OTHER EMERGING ISSUES

The ISAC systems sense targets and communicate with users, and thus they are vulnerable to security issues. Due to the uncertainty of the locations of the attackers and the dynamics of the targets, learning algorithms like DRL can be used to effectively address the attackers. Second, ISAC systems like autonomous vehicle systems are highly dynamic, which makes channel estimation challenging. Learning algorithms like unfolding learning can be used to estimate the channels for data communication. Third, recent works consider joint computation offloading and sensing, i.e., sharing the waveform signal, power resource, and bandwidth, to improve resource utilization. However, there may be a trade-off between the computation offloading performance, e.g., latency, and the sensing performance, e.g., radar detection range and velocity resolution. Under the dynamics and stochastics of the fading wireless channels, arrivals of computing tasks, and target mobility, learning algorithms have been used to effectively address the trade-off.

A. ISAC Security

To perform the sensing and communication functions, the ISAC systems may need to transmit more types of signals at the same time. Thus, attackers like eavesdroppers and jammers are easier to detect and attack the system. Few intelligent approaches have been recently proposed to address the attackers. The first work can be found in [203] which consists of an ISAC-equipped node and a jammer. The jammer tries to lower the probability of detection by generating jamming signals. As such, the radar receiver of the ISAC node cannot separate the target echo signal and the jamming signal. The ISAC-equipped node aims to select the stay idle, communication function, sensing function, or deception, i.e., transmitting fake signals to mislead the jammer, to maximize the SINR of the communication and radar functions. A DDQN algorithm with prioritized experience replay (PER) is used to solve the problem that enables the ISAC-equipped node to learn the decisions properly. The PER technique allows the DDQN algorithm to take important and frequent experiences, thus helping to learn the decisions more efficiently. However, the work is limited to a single ISAC-equipped node, and a general case with multiple nodes needs to be further investigated.

Different from [203], the authors in [204], [205] address the physical layer security in an ISAC system. For instance, the system [205] consists of a BS that serves multiple legitimate users and senses a sensing target with the presence of an eavesdropper. Both the sensing target and the eavesdropper eavesdrop on the communication signals transmitted from the BS to the users. An STAR-RIS is deployed to improve the total secrecy rate of the users corresponding to the sensing target and the eavesdropper. An optimization problem is formulated that aims to maximize the beamforming and the STAR-RIS transmitting and reflecting coefficients to maximize the total secrecy rate. To handle the dynamics of the system of the problem, two DRL learning algorithms, i.e., SAC and DDPG, are developed to solve the optimization problem. Simulation results show that the SAC outperforms the DDPG in terms of reward, hence the total secrecy rate. However, the proposed approach requires perfect knowledge of the channel state information, which is practically difficult to obtain, especially with the channel associated with the eavesdropper. To further improve the secrecy rate, artificial noise can be used at the BS as presented in [206]. In this case, the BS optimizes the transmit beamforming, AN power, and the RIS phase shifts to maximize the achievable secrecy rate subject to the beam pattern error requirement of the radar function. Then, the SAC algorithm is used to solve the optimization problem. Compared with the DQN or DDQN, the SAC algorithm can improve the robustness of training and address the overestimation of Q-value since it explores more strategies randomly due to the addition of entropy to the reward function. Indeed, simulation results show that the proposed SAC algorithm can increase the secrecy rate up to 4.4% compared with the baseline schemes.

Different from the aforementioned works, e.g., [206], the work in [207] considers a coexistence of a radar system and communication with the presence of an eavesdropper. Thus, the optimization problem needs to account for the minimum interference requirement between the radar and communication systems. The channel state information between the radar system and communication system may not be known due to the incorporation of the two systems, an autoencoder technique is deployed to determine the desired signals and cross-interference signal. Then, a SCA algorithm is developed to optimize the beamforming of the two systems to maximize the secrecy rate. Simulation results show the feasibility of the proposed algorithm. However, it is not clear how the output of the autoencoder is used as the input of the SCA algorithm. Also, where the autoencoder is deployed is not mentioned.

B. ISAC Channel Estimation

In scenarios with high mobility such as V2X communications, radar can be used to predict the channels supporting the communications as presented in [208]. In this work, an RSU is equipped with a radar unit and a communication unit serving the vehicles. The radar module uses the root MUSIC algorithm [209] to determine the angle information of the vehicles. The non-zero position information of the sparse angle-frequency channel is obtained based on the estimated angles. Based on the information, the learned iterative shrinkage thresholding algorithm for complex group row-sparse matrix signals [210] is adopted to estimate the channel. Simulation results show that the proposed algorithm improves the normalized mean square error up to 5 dB compared with the learned iterative shrinkage thresholding algorithm for complex group row-sparse matrix signals [210].

To predict the locations of the targets, the BS can leverage the MUSIC algorithm [211] to estimate the delay and the Doppler frequency values based on the reflected signals. To accurately predict the locations of the next time slot, the LSTM network is used where the input layer consists of locations estimated in the past time slots, and the output predicts the locations of the targets in the next time slot. Simulation results show that the proposed LSTM can achieve the RMSE up to 10^{-4} as the number of historical time steps is larger. However, the increase in historical time steps may not further improve the accuracy of the trajectory due to the redundant historical information.

C. Joint Sensing and Federated Learning

Federated learning has recently emerged as a decentralized machine learning model that allows a deep learning model owned by, e.g., a BS, to be trained by multiple clients. Particularly, the BS as model owner transmits a global model to the clients. Then, the clients use their local dataset to train the model. After that, the clients transmit the trained models to the BS for model aggregation. During the transmission of the trained models, the clients can leverage the signals to sense or track a common target. As such, client selection and resource allocation may impact both the federated learning performance and sensing performance.

The authors in [212] addressed the client selection and power control in federated learning (FL) with sensing, computing and communication. The system model is illustrated in Fig. 20 which consists of ISAC-equipped clients for local training and a BS for model aggregation. After locally training the FL models, the clients send their gradients to the BS for the FL AirComp aggregation. The clients also receive signals reflected from a common target, which are then transmitted to the BS to detect the presence of the target using the Neyman-Pearson detector. Since the sensing and communication may use different modulation schemes with different waveforms, only clients close to the target are selected for the target detection. Thus, the optimization problem aims to select the clients for the sensing and FL aggregation as well as the transmit power to minimize the MSE of the AirComp subject to the detection probability requirement. The Lagrangian relaxation method is proposed to solve the NP-hard mixed-integer nonlinear problem. This can be the first work that considers the sensing in the FL. However, since the clients need to transmit the reflected signals to the BS for detection, the FL latency can increase. Moreover, the use of different modulations and waveforms for the sensing and communication functions may result in a lower number of clients selected for the FL, this decreases the MSE of the AirComp and the performance of the FL.

The model similar to [212] can be found in [213] where multiple ISAC-equipped clients perform the target sensing while uploading their local models to the BS. However, different from [212], the work in [213] aims to jointly optimize the beamformers associated with the communication signals for uploading the local models and associated with the sensing signals of the clients to minimize the total energy consumption and CRB of the target sensing of the clients. The non-convex problem is reformulated by using the Lagrangian dual problem, which is then solved by multi-task learning based on multiple gradient descent algorithm (MGDA) [214]. Simulation results show the convergence (thus the feasibility) of the algorithm to solve the problem. However, the proposed algorithms require the proper selection of the parameters of the learning algorithm, which is practically challenging.

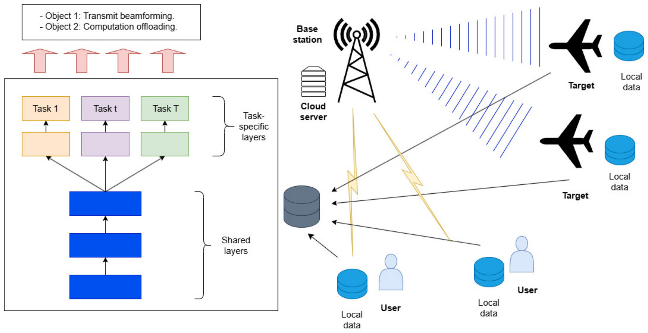


Fig. 20. Illustration of the system model proposed in [212], [213]. Here, the clients transmit the gradients of their local models back to the BS so that they can be aggregated. The BS also receives and processes the signals of a common target detected and tracked by the clients.

D. Joint Sensing and Computation Offloading

Joint sensing and computation offloading approaches have been recently proposed for the ISAC systems. The key idea is to leverage the offloading signals, i.e., transmitted from task users, for the radar function for sensing or tracking targets. As such, the radar performance can be influenced by the transmit power of the offloading signal, the length of the computation task, and the size of part of the offloaded task.

The work in [215] presents an approach to joint computation offloading and sensing. Particularly, the IoT devices equipped with ISAC units offload their tasks to an edge server via the communication functions. The radar receivers of the IoT devices leverage the offloading signals reflected from the targets to estimate them. Given the high mobility of the targets, the advantage actor critic (A2C) [216] algorithm (rather than DQN or DDQN) is used due to its faster convergence. Therein, the action space consists of the number of offloaded bits, the number of subcarriers assigned to the IoT devices and the corresponding transmit power. The state space consists of the energy of the IoT devices and wireless channels involved in the system. The reward is defined as a function of the computation efficiency and conditional MI of the IoT devices. Simulation results show that the A2C algorithm improves the performance up to 78% compared with the DDQN. However, it faces a large space of the problem as the number of IoT devices increases.

Different from [215], the work in [217] allows an ISAC-equipped vehicle to share the power budget between the sensing and communication functions rather than the offloading signal. Particularly, considering VEC scenarios, the ISAC-equipped vehicle divides its tasks into multiple subtasks, encodes the subtasks with the (m, k)-maximum distance separable (MDS) code, and transmits the coded subtasks to other vehicles the task execution. The (m, k)-maximum distance separable (MDS) code aims to address the straggler issue that often occurs in vehicle edge computing scenarios. Given the high mobility of the vehicles as well as the uncertainty of computing resources and stochastic networking resources, DDQN combined with transfer learning is used to optimize the fractions of power allocated to the radar and communication functions and the MDS parameters to minimize the overall computing latency and maximize the radar detection range. An extension of this work can consider the use of the FWCM technique with mmWave for the ISAC to improve the radar and communication performance. In this case, multi-antenna should be used, and beamforming prediction needs to be implemented. DRL algorithms based on LSTM can be used for the beamforming prediction.

Different from [217], the authors of [218] allow the task vehicles to offload their tasks to RSUs. Also, each RSU is equipped with an ISAC unit that allows it to predict the locations of the vehicles and to forward the received tasks to other RSUs. This helps to improve the transmission rate between the task vehicle and the RSUs and to reduce the overall computation latency. The work aims to optimize the vehicle-RSU association and computation resource allocation to minimize the system cost, i.e., the latency and energy consumption for the vehicles, subject to the queue constraints at the vehicles and RSUs. The Lyapunov optimization method is first used for the queue constraint transformation. Under the dynamics and uncertainty of the system, a DRL based on SAC [69] is proposed to solve the optimization problem. The SAC introduces an entropy term to encourage exploration, which better strikes the balance between exploitation and exploration and leads to the development of a more robust and explorative policy compared with traditional DRL algorithms like DDPG. Experimental results illustrate that the proposed SAC approach outperforms the baseline schemes in terms of system cost and queue length. However, how the sensing performance, e.g., sensing error, impacts the system performance is not discussed.

E. Interference Management

Interference management is another primary challenge that occurs in the coexistence of communications systems and radar systems, e.g., FMCW-based radar. Traditional techniques such as [219] are proposed to detect and remove the interference. However, the interference mitigation performance heavily depends on the interference detection. Interference mitigation approaches based on DL have been proposed as in [220]. Particularly, when receiving the signal reflected from the target, the processing of the range and velocity information results in a two-dimensional image, i.e., the Range-Doppler (RD) map. The RD maps in the current time slot and the previous time slots are fed into a CNN autoencoder with spatial attention. The spatial attention helps to focus on some important regions of the image and ignores unimportant regions of the image. The output is the interfered RD map. Simulation results implemented with real dataset RaDICaL [221] show that compared with the existing interference mitigation approaches, i.e., the RD RIS Model D [222], the proposed CNN approach improves the SINR up to 4% while reducing the EVM, i.e., the magnitude between the clean RD map and the interfered RD map, up to 34%.

The authors of [223] consider the extended vehicular A (EVA) multipath channel, which is widely used in 4G/5G wireless network systems, and propose a modified autoencoder framework for learning MIMO communication systems. Specifically, the authors aim to track the inter-symbol interference (ISI) problem caused by multipath environment with an autoencoder-fueled framework [224]. This model combines feedforward neural layers and convolutional layers to track ISI problems, achieving better bit-error-rate (BER) performance than traditional MIMO systems without the requirement of CSI. Furthermore, the authors show that the proposed framework can be applied to single-input single-output (SISO) and multiple-input single-output (MISO) systems with even better BER performances compared to current deep learning-based methods. It is worth mentioning that this framework can be trained and tested under additive white Gaussian noise (AWGN), flat Rayleigh fading and multipath channels, demonstrating generality. However, more efficient choices than autoencoder can be considered for interference reconstruction.

Similar to [223], the work in [225] proposes another approach to signal interference elimination in ISAC systems using autoencoder [224]. The authors argue that removing interference can be transformed into a type of denoising, which is one of the more well-known problems. An autoencoder-based scheme is then devised to jointly optimize communication and remote sensing, whose information is unified through compression (encoding), polluted with noise and reconstructed using the decoder. Furthermore, an inception model is integrated into the convolution of the decoder to maximize the efficiency of information extraction. Simulations show the success of this model in balancing the sensing and communication performance and eliminating interference in continuous transmission procedures. A depiction of the framework can be found in Fig. 21.

The authors of [226] aim to remove communication interference using a complex-valued neural network named ADMM-Net. This neural network, designed using deep unfolding on the alternating direction method of multipliers (ADMM) [227] optimization algorithm, can simultaneously retrieve a super-solution angle-Doppler-range image [228] and eliminate communication interference. The model addresses an uncooperative spectrum-sharing scenario where the radar serves as the primary function and the communications exploits parts of the shared spectrum. In this case, the return of a given pulse shall only be interfered by signals that spectrally overlap with that pulse. The signals also tend to be sparse in the frequency domain. As a result, interference

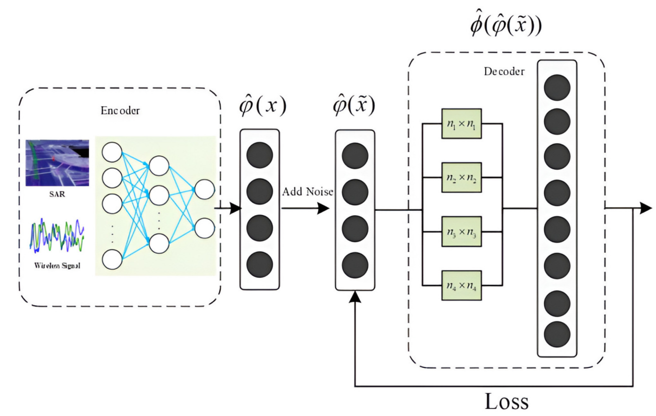


Fig. 21. Autoencoder-based interference elimination scheme for ISAC systems, as introduced by the authors of [225]. Here, the original input x consists of different types of signals, namely synthetic aperture radar (SAR)-based information and wireless signals. After x is transformed to a latent representation using the encoder $\hat{\varphi}$, it is mixed with noise features. The decoder $\hat{\phi}$ receives the noisy encoding $\hat{\varphi}(\tilde{x})$ and reconstructs it in reference to x.

appears in the form of sparse noise in radar measurements, inspiring an optimization problem for joint image recovery and interference removal. The underlying structure of the aforementioned ADMM-Net is then fueled by the ADMM equations of this problem. To the authors' best knowledge, this is the only significant work that addresses communication interference removal using deep unfolding techniques. It also demonstrates the potential of deep unfolding in terms of enhancing optimization algorithms for interference removal problems. The biggest issue that this method encounters is the off-grid assumption, which may not be true in all cases.

The work in [229] introduces a challenging Integrated Sensing, Communication and Computation (ISCC) system in which the BS serves both the FL and radar sensing task at the same time. In this framework, analog-integrated downlink signals are employed for global model dissemination, as well as radar sensing. Consequently, the model received at each edge device will be distorted by the sensing signal, channel fading and receiver noise. For uplink transmission, the Aircomp-enabled analog uplink transmissions are utilized for model aggregation. To mitigate the interference between the model signal and radar echo signal during uplink transmission, a receive beamforming approach is proposed. Moreover, the authors conduct a theoretical analysis to illustrate how the performance of FL is affected by device selection downlink and uplink transmission. FL enhancement considering the constraints of radar sensing is then formulated as an optimization problem, which the authors design a greedy iterative algorithm to solve. This work showcases the capability of beamforming methods in interference removal, further demonstrating its versatility in the communication system design field. Nevertheless, the proposed solution might struggle to efficiently handle interference when the number of devices becomes larger.

F. Discussions

Apart from the beamforming design, waveform design, and resource management, learning algorithms can be used to address other emerging issues in ISAC systems, such as security, interference management, and channel estimation. Some ISAC frameworks including joint sensing and federated learning and joint sensing and computation offloading are considered to be research directions. Major observations from this section are as follows:

- Regarding ISAC security, existing learning approaches are mostly proposed to address the jamming attack and eavesdropping attack that are two common and serious security issues in ISAC systems. In these works, DRL-based learning algorithms are used due to the fact that the locations of the attackers, e.g., eavesdropper, are usually hard to be detected, and that the DRL are able to learn the consequences of its actions via trial and error. However, the learning algorithms for ISAC security are mostly standard DRL algorithms, which face large state and action spaces. In general, there are still few learning approaches for the ISAC security, and thus there is plenty of room for improvement on this direction. For example, both jamming attacks and eavesdroppers can co-exist in the ISAC systems. In this case, covert communication technique can be used, and learning algorithms are then applied to minimize the false detection probability and missed detection probability.
- Learning algorithms can be utilized for ISAC channel estimation more effectively due to their ability to model nonlinearities in channels (e.g., multipath fading or Doppler). RNN-based models like LSTMs are also capable of leveraging short-term and long-term dependencies to provide sensing predictions. However, they require extensive memory to store the historical time steps, which may not be feasible in resource-constrained situations.
- Joint sensing and federated learning or computation offloading are considered novel ISAC systems. These schemes allow ISAC clients or users to sense targets at the time they upload local models or task request to the server. Leveraging the signals of local models or offloading for the sensing function results in an improvement in resource allocation. Nevertheless, Leveraging the signals of the local models or offloading for the sensing function helps to improve the resource utilization. However, the sensing performance and latency may be directly influenced by the training process or task execution. Existing works assume that the latency of the training process or task execution is insignificant. However, in reality, the global neural network or task size is very large with high computation latency. As a result, the target update may not be updated on time. Moreover, the existing works do not account for the dynamics of federated learning or computation offloading systems such as wireless fading, task size, computation resource. Adaptive learning algorithms need to be further investigated.
- Interference management is a popular issue in the ISAC due to the coexisting of sensing and communication functions. Most of the existing works use autoencoders that are able to jointly learn how encode (transmit) and decode (receive) signals to minimize interference. This enables

the receiver to distinguish the desired signal from the inference signals. Nevertheless, to generalize well across different interference scenarios, the autoencoders require large and diverse datasets. Moreover, autoencoders are computationally intensive and requires extensive hyperparameter tuning to maintain stability, which can be challenging for resource-constrained ISAC devices such as ISAC IoT devices. These issue can be tackled by using generative AI to synthesize signal data, enriching the data resource for the models. In addition, lightweight learning models need to be further investigated.

IX. CONCLUSION, CRITICAL TECHNIQUES, AND FUTURE WORKS

This paper has presented a comprehensive survey of the applications of learning algorithms for ISAC systems. Particularly, we have provided detailed reviews, analyses, and comparisons of the learning approaches to emerging issues in ISAC systems. The emerging issues include beamforming designing/tracking, waveform design, resource allocation, AoA/AoD estimation, signal classification, and wireless sensing. Despite several advantages, learning algorithms have critical techniques. Particularly, learning algorithms such as DRL usually face large spaces that make them fail to converge. Multi-agent DRL algorithms can be used, but they require a significant amount of overhead among the agents. Some advanced learning algorithms like CNN have a large number of hyperparameters, which makes it difficult to find a proper set of parameters for the algorithms to guarantee the optimal solution with an acceptable convergence speed. Requirements of high computing resources, time consumption, and large datasets are other critical issues of learning algorithms when being developed in ISAC systems. Even with those, with the super-performance of the learning algorithms compared with the traditional algorithms, the following research directions can be considered to be future works:

- *Incentive mechanisms for ISAC:* ISAC devices perform dual communication and radar functions to serve the communication users and sensing targets. Thus, the ISAC devices have a high demand for resources, i.e., spectrum, to guarantee their QoSs. As a result, the spectrum trading between the ISAC devices and spectrum owners, e.g., network operators, needs to be further considered. To motivate the network operators to provide spectrum resources, incentive mechanisms based on DL [230], [231], [232] can be used to maximize the network operators' revenue while guaranteeing desired economic properties.
- *Integration of emerging techniques with ISAC:* Rate-splitting Multiple Access (RSMA) [233], [234] and simultaneously transmitting and reflecting (STAR) reconfigurable intelligent surface (RIS) (STAR-RIS) [235], [236] as emerging techniques have been recently deployed to improve the communication performance. Therefore, they can be used to effectively enhance the performance of the ISAC systems. However, the combination of these techniques will make the beamforming design more complicated and challenging. Investigation of robust learning algorithms to address the challenges is a promising research direction. Moreover, incorporating advanced signal processing techniques like Improper Gaussian Signaling (IGS) [237], [238] to enhance channel estimation is also another exciting research direction.
- *Near-field communications with ISAC:* Future BSs can be equipped with large-scale antenna arrays, which leads to near-field communications for both target sensing and data communication, namely near-field ISAC [239]. Particularly, near-field sensing helps to gain the sensing performance, e.g., in terms of angular and range estimation. However, clutter interference may be one of the key issues in near-field ISAC. Channel knowledge maps can be designed to eliminate interference through learning algorithms [239].
- *Data fusion with learning algorithms for ISAC:* Data fusion in combination with advanced learning algorithms can revolutionize future ISAC systems. By integrating diverse data sources like sensor readings, communication feedback, and contextual information, ISACs become more robust and adaptable. In this context, ML, especially DL and RL, can extract complex patterns from fused data, thus resulting in more accurate predictions, resource allocation, and decision-making. Furthermore, FL comprehensively addresses privacy concerns by combining data from distributed nodes, while multi-modal learning should leverage heterogeneous data for improved situational awareness. Some existing challenges like low-latency processing, energy efficiency, and model robustness in dynamic environments may require innovative solutions. Collaborative research focusing on these aspects will enable seamless ISAC deployment in various applications.
- *Diffusion-based generative AI for ISAC:* Channel estimation is of high importance that is used for the beamforming design and resource allocation in ISAC systems. However, this task is challenging in ISAC systems due to their complicated and dynamic environments. Recent works have shown that the generative AI tools are able to improve the channel estimation performance compared with traditional techniques like least square (LS) in communication systems. Therefore, using generative AI tools for channel estimation in ISAC systems is an emerging research topic. For example, the work in [240] modeled the channel matrix as an image, and then recast the channel estimation problem as a signal denoising task. Note that due to the ability of generating images, based on the channel estimation, the diffusion models can be further used to visualize the target in 3D for the ISAC system as presented in [241].
- *Learning algorithms for Space-Air Ground Integrated Network (SAGIN):* SAGIN as an emerging technology in the next generation networks [242] is able to provide wide coverage and high QoS to the ground users including ISAC IoTs [243] by leveraging benefits of different types of networks, i.e., satellites, UAVs, and cellular networks.

However, given the high heterogeneity of these networks and the mobility of the ground users, the user handover among the networks is a challenging problem. Given the dynamics of the SAGIN systems, learning algorithms such as multi-agent DRL can be adopted that allows the ground users locally decide which network to be served. Advanced learning techniques such as diffusion model-based genAI algorithms as proposed in [243] should be effective solutions. Another issue of the SAGIN systems is obtaining the relative locations of the ground users (e.g., serving for beamforming design). For this, ISAC-assisted schemes can be used for the location sensing [242]. Advanced learning algorithms like LSTM modules can be an effective solution for the ground user mobility tracking.

- *Learning algorithms for hybrid intelligent reflecting/refracting surfaces (HIRS)-enabled network systems:* HIRS allows an incoming signal to be reflective or refractive [244], which addresses the limitations of the traditional intelligent reflecting surface (IRS) [245]. With HIRS, an ISAC BS can communicate with users and sense targets at both sides of the IRS surface, which improves the versatility and coverage of the network compared with the traditional IRS. However, the mobility of users and targets may impose challenges to the phase shift optimization of the HIRS. Adaptive learning algorithms such as deep reinforcement learning algorithms can be used to learn and decide proper phase shifts given states of the users and targets. Even with those, training these algorithms with the continuous phase shifts may be hard due to the high complexity and large space of the system. Discretizing the phase shifts should address this issue, which however reduces the network performance.

ACKNOWLEDGMENT

Part of this work was done when he was with the Department of Electrical, Electronic, and Computer Engineering, University of Ulsan, South Korea. Part of this work was done when he was with the College of Computing and Data Science, Nanyang Technological University, Singapore.

REFERENCES

[1] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, and T. X. Han, “A survey on fundamental limits of integrated sensing and communication,” *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 994–1034, 2nd Quart., 2022.

[2] S. Lu, F. Liu, Y. Li, K. Zhang, H. Huang, and J. Zou, “Integrated sensing and communications: Recent advances and ten open challenges,” *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19094–19120, Jun. 2024.

[3] Z. Wang, X. Mu, and Y. Liu, “STARS enabled integrated sensing and communications,” *IEEE Trans. Wireless Commun.*, vol. 22, no. 10, pp. 6750–6765, Oct. 2023.

[4] Y. Cui, F. Liu, X. Jing, and J. Mu, “Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges,” *IEEE Netw.*, vol. 35, no. 5, pp. 158–167, Sep./Oct. 2021.

[5] A. Kaushik et al., “Integrated sensing and communications for IoT: Synergies with key 6G technology enablers,” *IEEE Internet Things Mag.*, vol. 7, no. 5, pp. 136–143, Sep. 2024.

[6] Y. Zhong et al., “Empowering the V2X network by integrated sensing and communications: Background, design, advances, and opportunities,” *IEEE Netw.*, vol. 36, no. 4, pp. 54–60, Jul. 2022.

[7] Y. Li, F. Liu, Z. Du, W. Yuan, Q. Shi, and C. Masouros, “Frame structure and protocol design for sensing-assisted NR-V2X communications,” *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 11045–11060, Dec. 2024.

[8] X. Cheng, D. Duan, S. Gao, and L. Yang, “Integrated sensing and communications (ISAC) for vehicular communication networks (VCN),” *IEEE Internet Things J.*, vol. 9, no. 23, pp. 23441–23451, Dec. 2022.

[9] Q. Zhang, H. Sun, X. Gao, X. Wang, and Z. Feng, “Time-division ISAC enabled connected automated vehicles cooperation algorithm design and performance evaluation,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2206–2218, Jul. 2022.

[10] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, “Wireless sensing for human activity: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1629–1645, 3rd Quart., 2020.

[11] K. Meng et al., “UAV-enabled integrated sensing and communication: Opportunities and challenges,” *IEEE Wireless Commun.*, vol. 31, no. 2, pp. 97–104, Apr. 2024.

[12] X. Jing, F. Liu, C. Masouros, and Y. Zeng, “ISAC from the sky: UAV trajectory design for joint communication and target localization,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 12857–12872, Oct. 2024.

[13] J. Mu, R. Zhang, Y. Cui, N. Gao, and X. Jing, “UAV meets integrated sensing and communication: Challenges and future directions,” *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 62–67, May 2023.

[14] F. Deng, E. Jovanov, H. Song, W. Shi, Y. Zhang, and W. Xu, “WiLDAR: WiFi signal-based lightweight deep learning model for human activity recognition,” *IEEE Internet Things J.*, vol. 11, no. 2, pp. 2899–2908, Jan. 2023.

[15] U. M. Khan, Z. Kabir, S. A. Hassan, and S. H. Ahmed, “A deep learning framework using passive WiFi sensing for respiration monitoring,” in *Proc. GLOBECOM*, 2017, pp. 1–6.

[16] W. Saad, M. Bennis, and M. Chen, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May 2019.

[17] M. Razaviyayn, “Successive convex approximation: Analysis and applications,” Ph.D. dissertation, Dept. Comput. Sci., Univ. Minnesota, Minneapolis, MI, USA, May 2014.

[18] S. Desale, A. Rasool, S. Andhale, and P. Rane, “Heuristic and meta-heuristic algorithms and their relevance to the real world: A survey,” *Int. J. Comput. Eng. Res. Trends*, vol. 351, no. 5, pp. 2349–7084, 2015.

[19] N. C. Luong, X. Lu, D. T. Hoang, D. Niyato, and D. I. Kim, “Radio resource management in joint radar and communication: A comprehensive survey,” *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 780–814, 2nd Quart., 2021.

[20] R. Yang et al., “Deep learning-based joint transmit beamforming for dual-functional radar-communication system,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 15198–15211, Oct. 2024.

[21] N. C. Luong et al., “Applications of deep reinforcement learning in communications and networking: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3133–3174, 4th Quart., 2019.

[22] H. Xu and S. Mannor, “Robustness and generalization,” *Mach. Learn.*, vol. 86, pp. 391–423, Nov. 2012.

[23] U. Demirhan and A. Alkhateeb, “Integrated sensing and communication for 6G: Ten key machine learning roles,” *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 113–119, May 2023.

[24] H. Al Kassir, Z. D. Zaharis, P. I. Lazaridis, N. V. Kantartzis, T. V. Yioultsis, and T. D. Xenos, “A review of the state of the art and future challenges of deep learning-based beamforming,” *IEEE Access*, vol. 10, pp. 80869–80882, 2022.

[25] F. A. Aoudia and J. Hoydis, “Waveform learning for next-generation wireless communication systems,” *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 3804–3817, Jun. 2022.

[26] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, “Machine learning basics,” *Deep Learn.*, vol. 1, pp. 98–164, Jun. 2016.

[27] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[28] G. Zhong, L.-N. Wang, X. Ling, and J. Dong, “An overview on data representation learning: From traditional feature learning to recent deep learning,” *J. Finance Data Sci.*, vol. 2, no. 4, pp. 265–278, 2016.

[29] Z. Wei et al., “Integrated sensing and communication signals toward 5G-a and 6G: A survey,” *IEEE Internet Things J.*, vol. 10, no. 13, pp. 11068–11092, Jul. 2023.

[30] 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.

[31] W. Zhou, R. Zhang, G. Chen, and W. Wu, "Integrated sensing and communication waveform design: A survey," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1930–1949, 2022.

[32] J. A. Zhang et al., "Enabling joint communication and radar sensing in mobile networks—A survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 306–345, 1st Quart., 2022.

[33] Z. Wei et al., "Integrated sensing and communication channel modeling: A survey," *IEEE Internet Things J.*, vol. 12, no. 12, pp. 18850–18864, Jun. 2025.

[34] W. Jiang et al., "Terahertz communications and sensing for 6G and beyond: A comprehensive review," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 4, pp. 2326–2381, 4th Quart., 2024.

[35] A. M. Elbir, K. V. Mishra, S. Chatzinotas, and M. Bennis, "Terahertz-band integrated sensing and communications: Challenges and opportunities," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 39, no. 12, pp. 38–49, Dec. 2024.

[36] X. Fang, W. Feng, Y. Chen, N. Ge, and Y. Zhang, "Joint communication and sensing toward 6G: Models and potential of using MIMO," *IEEE Internet Things J.*, vol. 10, no. 5, pp. 4093–4116, Mar. 2022.

[37] E. F. Morales, R. Murrieta-Cid, I. Becerra, and M. A. Esquivel-Basaldua, "A survey on deep learning and deep reinforcement learning in robotics with a tutorial on deep reinforcement learning," *Intell. Service Robot.*, vol. 14, no. 5, pp. 773–805, 2021.

[38] N. Wu et al., "AI-enhanced integrated sensing and communications: Advancements, challenges, and prospects," *IEEE Commun. Mag.*, vol. 62, no. 9, pp. 144–150, Sep. 2024.

[39] M. A. K. Respati and B. M. Lee, "A survey on machine learning enhanced integrated sensing and communication systems: Architectures, algorithms, and applications," *IEEE Access*, vol. 12, pp. 170946–170964, 2024.

[40] H. Xing, G. Zhu, D. Liu, H. Wen, K. Huang, and K. Wu, "Task-oriented integrated sensing, computation and communication for wireless edge AI," *IEEE Netw.*, vol. 37, no. 4, pp. 135–144, Mar. 2024.

[41] D. Wen, Y. Zhou, X. Li, Y. Shi, K. Huang, and K. B. Letaief, "A survey on integrated sensing, communication, and computation," *IEEE Commun. Surveys Tuts.*, early access, Dec. 23, 2024, doi: 10.1109/COMST.2024.3521498.

[42] N. Gonzalez-Prelcic et al., "The integrated sensing and communication revolution for 6G: Vision, techniques, and applications," *Proc. IEEE*, vol. 112, no. 7, pp. 676–723, Jul. 2024.

[43] M. H. Khoshafa et al., "RIS-assisted physical layer security in emerging RF and optical wireless communication systems: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, early access, Oct. 28, 2024, doi: 10.1109/COMST.2024.3487112.

[44] X. Zhu et al., "Enabling intelligent connectivity: A survey of secure ISAC in 6G networks," *IEEE Commun. Surveys Tuts.*, vol. 27, no. 2, pp. 748–781, 1st Quart., 2025.

[45] J. A. Zhang et al., "An overview of signal processing techniques for joint communication and radar sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1295–1315, Nov. 2021.

[46] M. L. Rahman, J. A. Zhang, X. Huang, Y. J. Guo, and R. W. Heath, "Framework for a perceptive mobile network using joint communication and radar sensing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 3, pp. 1926–1941, Jun. 2019.

[47] Z. Feng, Z. Fang, Z. Wei, X. Chen, Z. Quan, and D. Ji, "Joint radar and communication: A survey," *China Commun.*, vol. 17, no. 1, pp. 1–27, Jan. 2020.

[48] X. Liu, T. Huang, N. Shlezinger, Y. Liu, J. Zhou, and Y. C. Eldar, "Joint transmit beamforming for multiuser MIMO communications and MIMO radar," *IEEE Trans. Signal Process.*, vol. 68, pp. 3929–3944, 2020.

[49] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, Nov. 2020.

[50] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, and A. Petropulu, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4264–4279, Aug. 2018.

[51] R. Liu, M. Li, Q. Liu, and A. L. Swindlehurst, "Dual-functional radar-communication waveform design: A symbol-level precoding approach," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1316–1331, Nov. 2021.

[52] X. Wang, Z. Fei, J. A. Zhang, J. Huang, and J. Yuan, "Constrained utility maximization in dual-functional radar-communication multi-UAV networks," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2660–2672, Apr. 2020.

[53] J. Choi, J. Park, N. Lee, and A. Alkhateeb, "Joint and robust beamforming framework for integrated sensing and communication systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 17602–17618, Nov. 2024.

[54] Z. Yu, X. Hu, C. Liu, and M. Peng, "IRS-aided non-orthogonal ISAC systems: Performance analysis and beamforming design," *IEEE Trans. Green Commun. Netw.*, vol. 8, no. 4, pp. 1930–1942, Dec. 2024.

[55] X. Li, S. Guo, T. Li, X. Zou, and D. Li, "On the performance trade-off of distributed integrated sensing and communication networks," *IEEE Wireless Commun. Lett.*, vol. 12, no. 12, pp. 2033–2037, Dec. 2023.

[56] R. Yang, C.-X. Wang, J. Huang, E.-H. M. Aggoune, and Y. Hao, "A novel 6G ISAC channel model combining forward and backward scattering," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 8050–8065, Nov. 2023.

[57] Y. Niu, Z. Wei, L. Wang, H. Wu, and Z. Feng, "Interference management for integrated sensing and communication systems: A survey," 2024, *arXiv:2403.16189*.

[58] N. Su, F. Liu, C. Masouros, and A. A. Hilli, "Security and privacy in ISAC systems," in *Integrated Sensing and Communications*. Heidelberg, Germany: Springer, 2023, pp. 477–506.

[59] K. Qu, J. Ye, X. Li, and S. Guo, "Privacy and security in ubiquitous integrated sensing and communication: Threats, challenges and future directions," *IEEE Internet Things Mag.*, vol. 7, no. 4, pp. 52–58, Jul. 2024.

[60] S. Kurma, K. Singh, S. Mumtaz, T. A. Tsiftsis, and C.-P. Li, "Resource optimization in active-STAR-RIS-aided THz ISAC systems with DDA modulation: A machine-learning approach," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 15291–15307, Oct. 2024.

[61] W. Chen et al., "MADDPG algorithm for coordinated welding of multiple robots," in *Proc. 6th Int. Conf. Autom. Control Robot. Eng. (CACRE)*, 2021, pp. 1–5.

[62] P.-C. Wu, L.-H. Shen, K.-T. Feng, and C.-Y. Chan, "Federated reinforcement learning for multi-dual-STAR-RIS assisted DFRC-enabled multi-BS in ISAC systems," in *Proc. IEEE Int. Conf. Commun.*, 2024, pp. 2986–2991.

[63] X. Long, Y. Zhao, H. Wu, and C.-Z. Xu, "Deep reinforcement learning for integrated sensing and communication in RIS-assisted 6G V2X system," *IEEE Internet Things J.*, vol. 11, no. 24, pp. 39834–39849, Dec. 2024.

[64] J. Chen and X. Wang, "Learning-based intermittent CSI estimation with adaptive intervals in integrated sensing and communication systems," 2024, *arXiv:2405.14724*.

[65] P. Saikia, K. Singh, O. Taghizadeh, W.-J. Huang, and S. Biswas, "Meta reinforcement learning-based spectrum sharing between RIS-assisted cellular communications and MIMO radar," *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 1, pp. 164–179, Feb. 2024.

[66] J. Beck et al. "A survey of meta-reinforcement learning." 2024. [Online]. Available: https://arxiv.org/abs/2301.08028

[67] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey. "Meta-learning in neural networks: A survey." 2020. [Online]. Available: https://arxiv.org/abs/2004.05439

[68] Y. Eghbali, S. Faramarzi, S. K. Taskou, M. R. Mili, M. Rasti, and E. Hossain, "Beamforming for STAR-RIS-aided integrated sensing and communication using meta DRL," *IEEE Wireless Commun. Lett.*, vol. 13, no. 4, pp. 919–923, Apr. 2024.

[69] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor–critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 1861–1870.

[70] X. Chen, X. Cao, L. Xie, and Y. He, "DRL-based joint trajectory planning and beamforming optimization in aerial RIS-assisted ISAC system," in *Proc. IEEE Int. Workshop Radio Freq. Antenna Technol.*, 2024, pp. 510–515.

[71] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. "Proximal policy optimization algorithms." 2017. [Online]. Available: https://arxiv.org/abs/1707.06347

[72] J. Peng, Y. Cai, S. Liu, G. Wang, C. Wu, and R. Yin, "Learning-enabled radar-assisted predictive beamforming for UAV-aided networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2024, pp. 1304–1309.

[73] J. R. Hershey, J. L. Roux, and F. Weninger, "Deep unfolding: Model-based inspiration of novel deep architectures," 2014, *arXiv:1409.2574*.

[74] J. Ye, L. Huang, Z. Chen, P. Zhang, and M. Rihan, "Unsupervised learning for joint beamforming design in RIS-aided ISAC systems," *IEEE Wireless Commun. Lett.*, vol. 13, no. 8, pp. 2100–2104, Aug. 2024.

[75] A. Adhikary, M. S. Munir, A. D. Raha, Y. Qiao, Z. Han, and C. S. Hong, "Integrated sensing, localization, and communication in holographic MIMO-enabled wireless network: A deep learning approach," *IEEE Trans. Netw. Service Manag.*, vol. 21, no. 1, pp. 789–809, Feb. 2024.

[76] J. M. Mateos-Ramos, B. Chatelier, C. Häger, M. F. Keskin, L. L. Magoarou, and H. Wymeersch, "Semi-supervised end-to-end learning for integrated sensing and communications," in *Proc. IEEE Int. Conf. Mach. Learn. Commun. Netw.*, 2024, pp. 132–138.

[77] A. G. Howard, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[78] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," 2015, *arxiv.abs/1509.02971*.

[79] Y. Zheng, R. Liu, M. Li, and Q. Liu, "End-to-end learning for SLP-based ISAC systems," 2024, *arXiv:2401.05663*.

[80] X. Zhang and Q. Zhu, "Federated learning based integrated sensing, communications, and powering over 6G massive-MIMO mobile networks," in *Proc. IEEE Conf. Comput. Commun. Workshops*, 2024, pp. 1–6.

[81] H. Bouzabia, G. Kaddoum, and T. N. Do, "Integrated sensing and communications using generative AI: Countering adversarial machine learning attacks," in *Proc. IEEE Int. Conf. Commun.*, 2024, pp. 2895–2900.

[82] X. Zhou et al., "Hierarchical adversarial attacks against graph-neural-network-based IoT network intrusion detection system," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9310–9319, Jun. 2021.

[83] Y. Cui et al., "Sensing-assisted high reliable communication: A transformer-based beamforming approach," *IEEE J. Sel. Topics Signal Process.*, vol. 18, no. 5, pp. 782–795, Jul. 2024.

[84] X. Zhang, W. Yuan, C. Liu, J. Wu, and D. W. K. Ng, "Predictive beamforming for vehicles with complex behaviors in ISAC systems: A deep learning approach," *IEEE J. Sel. Topics Signal Process.*, vol. 18, no. 5, pp. 828–841, Jul. 2024.

[85] L. Cazzella, M. Mizmizi, D. Tagliaferri, D. Badini, M. Matteucci, and U. Spagnolini, "Deep learning-based target-to-user association in integrated sensing and communication systems," 2024, *arXiv:2401.12801*.

[86] Z. Wang and V. W. Wong, "Deep learning for ISAC-enabled end-to-end predictive beamforming in vehicular networks," in *Proc. IEEE Int. Conf. Commun.*, 2023, pp. 5713–5718.

[87] Q. Zhou, Y. Gong, and A. Nallanathan, "Radar-aided beam selection in MIMO communication systems: A federated transfer learning approach," *IEEE Trans. Veh. Technol.*, vol. 73, no. 8, pp. 12172–12177, Aug. 2024.

[88] H. Zhang, S. Gao, X. Cheng, and L. Yang, "Integrated sensing and communications toward proactive beamforming in mmWave V2I via multi-modal feature fusion (MMFF)," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 15721–15735, Nov. 2024.

[89] S. Imran, G. Charan, and A. Alkhateeb, "Environment semantic communication: Enabling distributed sensing aided networks," 2024, *arXiv:2402.14766*.

[90] S. Tariq, B. E. Arfeto, U. Khalid, S. Kim, T. Q. Duong, and H. Shin, "Deep quantum-transformer networks for multimodal beam prediction in ISAC systems," *IEEE Internet Things J.*, vol. 11, no. 18, pp. 29387–29401, Sep. 2024.

[91] A. Graff, Y. Chen, N. González-Prelcic, and T. Shimizu, "Deep learning-based link configuration for radar-aided multiuser mmWave vehicle-to-infrastructure communication," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 7454–7468, Jun. 2023.

[92] K. Chen, D. Liu, and Z. Zhang, "Beam tracking based on multi-point radar sensing and DNN prediction," in *Proc. 8th IEEE Int. Conf. Netw. Intell. Digit. Content*, 2023, pp. 222–226.

[93] H. Liu, Y. Zhao, X. Liu, B. Wang, and K. Cheng, "Deep learning empowered IoV: ISAC RCG-net beam tracking for seamless road communication," in *Proc. 27th Int. Conf. Comput. Supported Cooperative Work Design*, 2024, pp. 711–716.

[94] Y. Du, Y. Liu, K. Han, J. Jiang, W. Wang, and L. Chen, "Multi-user and multi-target dual-function radar-communication waveform design: Multi-fold performance tradeoffs," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 1, pp. 483–496, Mar. 2023.

[95] K. Zhong, J. Hu, C. Pan, M. Deng, and J. Fang, "Joint waveform and beamforming design for RIS-aided ISAC systems," *IEEE Signal Process. Lett.*, vol. 30, pp. 165–169, 2023.

[96] Y. Ding, S. Yan, X. Zhou, F. Shu, and S. Feng, "Radar-communication waveform design with detection probability constraints," *IEEE Wireless Commun. Lett.*, vol. 12, no. 1, pp. 168–172, Jan. 2023.

[97] N. H. Chu et al., "AI-enabled mm-Waveform configuration for autonomous vehicles with integrated communication and sensing," *IEEE Internet Things J.*, vol. 10, no. 19, pp. 16727–16743, Oct. 2023.

[98] A. M. Ahmed, L. Gharsalli, S. Fortunati, and A. Sezgin, "Reinforcement learning for cognitive integrated communication and sensing systems," in *Proc. 20th Eur. Radar Conf.*, 2023, pp. 395–398.

[99] Z. Kang and Q. He, "Learning-based joint waveform optimization and receiver design for dual-functional MIMO radar and communications," in *Proc. 31st Eur. Signal Process. Conf.*, 2023, pp. 710–714.

[100] P. Saikia, K. Singh, O. Taghizadeh, W.-J. Huang, and S. Biswas, "DRL algorithms for efficient spectrum sharing in RIS-aided MIMO radar and cellular systems," in *Proc. IEEE Mil. Commun. Conf.*, 2022, pp. 55–60.

[101] V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[102] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double $Q-$learning," in *Proc. AAAI*, vol. 2, Feb. 2016, pp. 2094–2100.

[103] Z. Wang, T. Schaul, M. Hessel, H. Van Hasselt, M. Lanctot, and N. De Freitas, "Dueling network architectures for deep reinforcement learning," in *Proc. 33rd Int. Conf. Int. Conf. Mach. Learn. (ICML)*, 2016, pp. 1995–2003. [Online]. Available: https://arxiv.org/abs/1511.06581

[104] W. Zhai, X. Wang, M. S. Greco, and F. Gini, "Reinforcement learning based integrated sensing and communication for automotive MIMO radar," in *Proc. IEEE Radar Conf. (RadarConf)*, 2023, pp. 1–6.

[105] P. Pulkkinen and V. Koivunen, "Model-free online learning for waveform optimization in integrated sensing and communications," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2023, pp. 1–5.

[106] M. Bică and V. Koivunen, "Multicarrier radar-communications waveform design for RF convergence and coexistence," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2019, pp. 7780–7784.

[107] P. Pulkkinen and V. Koivunen, "Model-based online learning for joint radar-communication systems operating in dynamic interference," in *Proc. 30th Eur. Signal Process. Conf.*, 2022, pp. 992–996.

[108] Q. Qi, X. Chen, C. Zhong, C. Yuen, and Z. Zhang, "Deep learning-based design of uplink integrated sensing and communication," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 10639–10652, Sep. 2024.

[109] J. M. Mateos-Ramos et al., "End-to-end learning for integrated sensing and communication," in *Proc. IEEE Int. Conf. Commun.*, 2022, pp. 1942–1947.

[110] M. Kim, T. Jahani-Nezhad, S. Lit, R. F. Schaefer, and G. Caire, "Short-length code designs for integrated sensing and communications using deep learning," in *Proc. IEEE Int. Conf. Commun.*, 2024, pp. 3536–3541.

[111] X. Zhang et al., "NAS-AMR: Neural architecture search-based automatic modulation recognition for integrated sensing and communication systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 8, no. 3, pp. 1374–1386, Sep. 2022.

[112] X. Wang, Z. Fei, J. Huang, and H. Yu, "Joint waveform and discrete phase shift design for RIS-assisted integrated sensing and communication system under Cramer–Rao bound constraint," *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 1004–1009, Jan. 2022.

[113] C. Ouyang, Y. Liu, and H. Yang, "On the performance of uplink ISAC systems," *IEEE Commun. Lett.*, vol. 26, no. 8, pp. 1769–1773, Aug. 2022.

[114] N. Shazeer et al., "Outrageously large neural networks: The sparsely-gated mixture-of-experts layer," 2017, *arXiv:1701.06538*.

[115] H. Liu, K. Simonyan, and Y. Yang, "DARTs: Differentiable architecture search," 2018, *arXiv:1806.09055*.

[116] W. Jiang, D. Ma, Z. Wei, Z. Feng, P. Zhang, and J. Peng, "ISAC-NET: Model-driven deep learning for integrated passive sensing and communication," *IEEE Trans. Commun.*, vol. 72, no. 8, pp. 4692–4707, Aug. 2024.

[117] J. Zhang, M. Liu, J. Tang, N. Zhao, D. Niyato, and X. Wang, "Joint design for RIS-aided ISAC via deep unfolding learning," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 1, pp. 349–361, Feb. 2025.

[118] P. Krishnananthalingam, N. T. Nguyen, and M. Juntti, "Constant modulus waveform design for wideband multicarrier joint communications and sensing via deep unfolding," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2024, pp. 1–6.

[119] L. Yang et al., "Diffusion models: A comprehensive survey of methods and applications," *ACM Comput. Surveys*, vol. 56, no. 4, pp. 1–39, 2023.

[120] S. Huang, M. Zhang, Y. Gao, and Z. Feng, "MIMO radar aided mmWave time-varying channel estimation in MU-MIMO V2X communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7581–7594, Nov. 2021.

[121] Y. Ma, Y. Zeng, and S. Sun, "A deep learning based super resolution DoA estimator with single snapshot MIMO radar data," *IEEE Trans. Veh. Technol.*, vol. 71, no. 4, pp. 4142–4155, Apr. 2022.

[122] C. Muth and L. Schmalen, "Autoencoder-based joint communication and sensing of multiple targets," in *Proc. 26th Int. ITG Workshop Smart Antennas 13th Conf. Syst. Commun. Coding (WSA SCC)*, 2023, pp. 1–6.

[123] C. Muth, B. Geiger, D. G. Gaviria, and L. Schmalen, "Loss design for single-carrier joint communication and neural network-based sensing," in *Proc. IEEE 27th Int. Workshop Smart Antennas (WSA)*, 2024, pp. 131–137.

[124] Y. Wu, F. Lemic, C. Han, and Z. Chen, "Sensing integrated DFT-spread OFDM waveform and deep learning-powered receiver design for terahertz integrated sensing and communication systems," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 595–610, Jan. 2023.

[125] M. Delamou, A. Bazzi, M. Chafii, and E. M. Amhoud, "Deep learning-based estimation for multitarget radar detection," in *Proc. IEEE 97th Veh. Technol. Conf. (VTC-Spring)*, 2023, pp. 1–5.

[126] S. Naoumi, A. Bazzi, R. Bomfin, and M. Chafii, "Complex neural network based joint AoA and AoD estimation for bistatic ISAC," *IEEE J. Sel. Topics Signal Process.*, vol. 18, no. 5, pp. 842–856, Jul. 2024.

[127] L. M.-M. Suárez, K. Chen-Hu, M. J. F.-G. García, and A. G. Armada, "Deep learning-aided robust integrated sensing and communications with OTFS and superimposed training," in *Proc. IEEE Int. Mediterr. Conf. Commun. Netw. (MeditCom)*, 2023, pp. 1–6.

[128] W. Lin, H. Zheng, X. Feng, and Y. Chen, "Deep unfolding network for target parameter estimation in OTFS-based ISAC systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2024, pp. 1–6.

[129] M. A. Aref and S. K. Jayaweera, "Classification of radar and communications signals using wideband autonomous cognitive radios," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–6.

[130] B. Bhargava, A. Deshmukh, M. V. Rupa, R. P. Sirigina, S. K. Vankayala, and A. Narasimhadhan, "Deep learning approach for wireless signal and modulation classification," in *Proc. IEEE 94th Veh. Technol. Conf. (VTC-Fall)*, 2021, pp. 1–6.

[131] C. Liu, Y. Chen, and S.-H. Yang, "Deep learning based detection for communications systems with radar interference," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6245–6254, Jun. 2022.

[132] T. Huynh-The, N. C. Luong, H. Phan, D. B. D. Costa, and Q.-V. Pham, "Improved waveform classification for integrated radar-communication 6G systems via convolutional neural networks," *IEEE Trans. Veh. Technol.*, vol. 73, no. 9, pp. 13921–13925, Sep. 2024.

[133] A. H. Oveis, A. Capria, A. L. Saverino, and M. Martorella, "Convolutional neural network for joint communication and radar signals classification," in *Proc. IEEE 24th Int. Radar Symp. (IRS)*, 2023, pp. 1–10.

[134] T. Huynh-The, V.-P. Hoang, J.-W. Kim, M.-T. Le, and M. Zeng, "WaveNet: Toward waveform classification in integrated radar-communication systems with improved accuracy and reduced complexity," *IEEE Internet Things J.*, vol. 11, no. 14, pp. 25111–25123, Jul. 2024.

[135] Y. E. Sagduyu, T. Erpek, A. Yener, and S. Ulukus, "Joint sensing and task-oriented communications with image and wireless data modalities for dynamic spectrum access," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, 2024, pp. 57–62.

[136] Y. E. Sagduyu, T. Erpek, A. Yener, and S. Ulukus, "Joint sensing and semantic communications with multi-task deep learning," *IEEE Commun. Mag.*, vol. 62, no. 9, pp. 74–81, Sep. 2024.

[137] H. Wu, Z. Zhang, C. Jiao, C. Li, and T. Q. Quek, "Learn to sense: A meta-learning-based sensing and fusion framework for wireless sensor networks," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8215–8227, Oct. 2019.

[138] M. Andrychowicz et al., "Learning to learn by gradient descent by gradient descent," in *Proc. NIPS*, 2016, pp. 3988–3996.

[139] O. Ma, A. R. Chiriyath, A. Herschfelt, and D. W. Bliss, "Cooperative radar and communications coexistence using reinforcement learning," in *Proc. IEEE 52nd Asilomar Conf. Signals Syst. Comput.*, 2018, pp. 947–951.

[140] A. R. Chiriyath, B. Paul, G. M. Jacyna, and D. W. Bliss, "Inner bounds on performance of radar and communications co-existence," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 464–474, Jan. 2015.

[141] C. E. Thornton, M. A. Kozy, R. M. Buehrer, A. F. Martone, and K. D. Sherbondy, "Deep reinforcement learning control for radar detection and tracking in congested spectral environments," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 4, pp. 1335–1349, Dec. 2020.

[142] C. E. Thornton, R. M. Buehrer, and A. F. Martone, "Efficient online learning for cognitive radar-cellular coexistence via contextual Thompson sampling," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2020, pp. 1–6.

[143] P. Liu, Y. Liu, T. Huang, Y. Lu, and X. Wang, "Decentralized automotive radar spectrum allocation to avoid mutual interference using reinforcement learning," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 1, pp. 190–205, Feb. 2021.

[144] E. Selvi, R. M. Buehrer, A. Martone, and K. Sherbondy, "Reinforcement learning for adaptable bandwidth tracking radars," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 5, pp. 3904–3921, Oct. 2020.

[145] W. W. Howard, A. F. Martone, and R. M. Buehrer, "Distributed online learning for coexistence in cognitive radar networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 2, pp. 1202–1216, Apr. 2023.

[146] D. Salami, W. Ning, K. Ruttik, R. Jäntti, and S. Sigg, "A joint radar and communication approach for 5G NR using reinforcement learning," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 106–112, May 2023.

[147] B. K. Chalise, A. F. Martone, and B. H. Kirk, "Model- and deep learning-based bandwidth and carrier frequency allocation in distributed radar networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 8022–8036, Dec. 2023.

[148] Z. Shan, P. Liu, L. Wang, and Y. Liu, "A cognitive multi-carrier radar for communication interference avoidance via deep reinforcement learning," *IEEE Trans. Cogn. Commun. Netw.*, vol. 9, no. 6, pp. 1561–1578, Dec. 2023.

[149] H. Wang, Z. Wang, Y. Chen, Z. Lu, and X. Wen, "Power efficient resource allocation for ISAC: Combing Lyapunov optimization and DRL," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2023, pp. 1135–1140.

[150] Y. Wang, M. Lou, W. Qian, Y. Bai, L. Tang, and Y.-C. Liang, "Intelligent beam tracking in radar-assisted MIMO-OFDM communication systems," *IEEE Trans. Veh. Technol.*, vol. 73, no. 11, pp. 16774–16789, Nov. 2024.

[151] A. Mondal, A. Dutta, and S. Biswas, "Dueling-DQN based spectrum sharing between MIMO radar and cellular networks," in *Proc. Joint Eur. Conf. Netw. Commun. 6G Summit (EuCNC/6G Summit)*, 2023, pp. 1–6.

[152] X. Liu, H. Zhang, K. Long, A. Nallanathan, and V. C. Leung, "Distributed unsupervised learning for interference management in integrated sensing and communication systems," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9301–9312, Dec. 2023.

[153] X. Wang, H. Wu, Y. Xu, H. Cao, N. Kumar, and J. J. P. C. Rodrigues, "Resource allocation in multi-cell integrated sensing and communication systems: A DRL approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2023, pp. 3210–3215.

[154] T. Zhang, K. Zhu, S. Zheng, D. Niyato, and N. C. Luong, "Trajectory design and power control for joint radar and communication enabled multi-UAV cooperative detection systems," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 158–172, Jan. 2022.

[155] Y. Qin, Z. Zhang, X. Li, W. Huangfu, and H. Zhang, "Deep reinforcement learning based resource allocation and trajectory planning in integrated sensing and communications UAV network," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 8158–8169, Nov. 2023.

[156] Z. Xie, Z. Wang, Z. Zhang, J. Wang, Z. Jiang, and Z. Han, "Distributed UAV swarm for device-free integrated sensing and communication relying on multi-agent reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 73, no. 12, pp. 19925–19930, Dec. 2024.

[157] Y. Yuan, X. Liu, T. Zhang, G. Cui, and L. Kong, "Reinforcement-learning-enhanced adaption of signal power and modulation for LPI radar system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 6, pp. 8555–8568, Dec. 2024.

[158] N. Q. Hieu, D. T. Hoang, N. C. Luong, and D. Niyato, "irdrc: An intelligent real-time dual-functional radar-communication system for automotive vehicles," *IEEE Wireless Commun. Lett.*, vol. 9, no. 12, pp. 2140–2143, Dec. 2020.

[159] N. Q. Hieu, D. T. Hoang, D. Niyato, P. Wang, D. I. Kim, and C. Yuen, "Transferable deep reinforcement learning framework for autonomous vehicles with joint radar-data communications," *IEEE Trans. Commun.*, vol. 70, no. 8, pp. 5164–5180, Aug. 2022.

[160] J. Lee, Y. Cheng, D. Niyato, Y. L. Guan, and D. González, "Intelligent resource allocation in joint radar-communication with graph neural networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 10, pp. 11120–11135, Oct. 2022.

[161] Z. Lu, M. C. Gursoy, C. K. Mohan, and P. K. Varshney, "Learning-based resource management in integrated sensing and communication systems," in *Proc. IEEE Conf. Comput. Commun. Workshops*, 2024, pp. 1–6.

[162] J. Lee, D. Niyato, Y. L. Guan, and D. I. Kim, "Learning to schedule joint radar-communication with deep multi-agent reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 406–422, Jan. 2021.

[163] S. Sun, Z. Sun, J. Fan, M. Hao, and C. Zhao, "Incorporating sensing into MIMO-NOMA communications: Theoretical joint rates bound and learning-based optimal design," in *Proc. IEEE Int. Conf. Commun.*, 2024, pp. 2252–2257.

[164] A. M. Elbir, K. V. Mishra, and A. C. GürbüZ, "Deep sparse array design for integrated sensing and communications," in *Proc. IEEE Int. Radar Conf.*, 2023, pp. 1–5.

[165] L. Xu, S. Sun, Y. D. Zhang, and A. Petropulu, "Joint antenna selection and beamforming in integrated automotive radar sensing-communications with quantized double phase shifters," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2023, pp. 1–5.

[166] L. Xu, S. Sun, Y. D. Zhang, and A. P. Petropulu, "Reconfigurable beamforming for automotive radar sensing and communication: A deep reinforcement learning approach," *IEEE J. Sel. Areas Sensors*, vol. 1, pp. 124–138, 2024.

[167] J. M. Pak, C. K. Ahn, P. Shi, Y. S. Shmaliy, and M. T. Lim, "Distributed hybrid particle/FIR filtering for mitigating NLOS effects in TOA-based Localization using wireless sensor networks," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 5182–5191, Jun. 2017.

[168] K. Ho and W. Xu, "An accurate algebraic solution for moving source location using TDOA and FDOA measurements," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2453–2463, Sep. 2004.

[169] W. Xue, W. Qiu, X. Hua, and K. Yu, "Improved Wi-Fi RSSI measurement for indoor localization," *IEEE Sensors J.*, vol. 17, no. 7, pp. 2224–2230, Apr. 2017.

[170] P. Wang, X. Wang, W. Li, X. Fan, and D. Zhao, "DV-hop Localization based on distance estimation using multinode and hop loss in IoT," *IEEE Internet Things J.*, vol. 11, no. 16, pp. 27803–27815, Aug. 2024.

[171] S. He and S.-H. G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 466–490, 1st Quart., 2015.

[172] Y. Lin, K. Yu, F. Zhu, J. Bu, and X. Dua, "The state of the art of deep learning-based Wi-Fi indoor positioning: A review," *IEEE Sensors J.*, vol. 24, no. 17, pp. 27076–27098, Sep. 2024.

[173] S. Rivetti et al., "Spatial signal design for positioning via end-to-end learning," *IEEE Wireless Commun. Lett.*, vol. 12, no. 3, pp. 525–529, Mar. 2023.

[174] J. Gante, G. Falcao, and L. Sousa, "Deep learning architectures for accurate millimeter wave positioning in 5G," *Neural Process. Lett.*, vol. 51, no. 1, pp. 487–514, 2020.

[175] X. Wang, M. Patil, C. Yang, S. Mao, and P. A. Patel, "Deep convolutional Gaussian processes for Mmwave outdoor localization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2021, pp. 8323–8327.

[176] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.

[177] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information," *IEEE Access*, vol. 5, pp. 18066–18074, 2017.

[178] Y. Ruan et al., "iPos-5G: Indoor positioning via commercial 5G NR CSI," *IEEE Internet Things J.*, vol. 10, no. 10, pp. 8718–8733, May 2023.

[179] C. Hou, Y. Xie, and Z. Zhang, "FCLoc: A novel indoor Wi-Fi fingerprints Localization approach to enhance robustness and positioning accuracy," *IEEE Sensors J.*, vol. 23, no. 7, pp. 7153–7167, Apr. 2023.

[180] C. Wu et al., "Learning to localize: A 3-D CNN approach to user positioning in massive MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4556–4570, Jul. 2021.

[181] Z. Chen, Z. Zhang, Z. Xiao, C. Zhang, and Z. Yang, "CSI of each subcarrier is a fingerprint: Multi-carrier cumulative learning based positioning in massive MIMO systems," in *Proc. IEEE 34th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun.*, 2023, pp. 1–7.

[182] M. A. M. Sadr, J. Gante, B. Champagne, G. Falcao, and L. Sousa, "Uncertainty estimation via monte carlo dropout in CNN-based mmWave MIMO localization," *IEEE Signal Process. Lett.*, vol. 29, pp. 269–273, 2021.

[183] L. Cai, Z. Yan, X. Zhang, and S. Yang, "Uplink sensing in vehicular networks with OTFS signaling: A LSTM-based approach," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2024, pp. 1–6.

[184] J. Hu, D. Liu, Z. Yan, and H. Liu, "Experimental analysis on weight K-nearest neighbor indoor fingerprint positioning," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 891–897, Feb. 2019.

[185] B. Zhang, H. Sifaou, and G. Y. Li, "CSI-fingerprinting indoor localization via attention-augmented residual convolutional neural network," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5583–5597, Aug. 2023.

[186] A. Salihu, S. Schwarz, and M. Rupp, "Attention aided CSI wireless localization," in *Proc. IEEE 23rd Int. Workshop Signal Process. Adv. Wireless Commun.*, 2022, pp. 1–5.

[187] Y. Wei and R. Zheng, "Efficient Wi-Fi fingerprint crowdsourcing for indoor localization," *IEEE Sensors J.*, vol. 22, no. 6, pp. 5055–5062, Mar. 2022.

[188] L. Zhang, X. Chu, and M. Zhai, "Machine learning-based integrated wireless sensing and positioning for cellular network," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.

[189] Y. Chen, N. González-Prelcic, T. Shimizu, and H. Lu, "Learning to localize with attention: From sparse mmWave channel estimates from a single BS to high accuracy 3-D location," 2023, *arXiv:2307.00167*.

[190] M. M. Butt, A. Pantelidou, and I. Z. Kovács, "ML-assisted UE positioning: Performance analysis and 5G architecture enhancements," *IEEE Open J. Veh. Technol.*, vol. 2, pp. 377–388, 2021.

[191] F. Meneghello, C. Chen, C. Cordeiro, and F. Restuccia, "Toward integrated sensing and communications in IEEE 802.11 BF Wi-Fi networks," *IEEE Commun. Mag.*, vol. 61, no. 7, pp. 128–133, Jul. 2023.

[192] F. Xiao, S. Zhang, S. Tang, S. Shen, H. Dong, and Y. Zhong, "WISION: Bolstering MAV 3-D indoor state estimation by embracing multipath of WiFi," *IEEE Trans. Veh. Technol.*, vol. 72, no. 1, pp. 253–266, Jan. 2022.

[193] S. Zhang, W. Wang, and T. Jiang, "Wi-Fi-inertial indoor pose estimation for microaerial vehicles," *IEEE Trans. Ind. Electron.*, vol. 68, no. 5, pp. 4331–4340, May 2021.

[194] Y. He, J. Liu, M. Li, G. Yu, and J. Han, "Forward-compatible integrated sensing and communication for WiFi," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 9, pp. 2440–2456, Sep. 2024.

[195] H. Yang, L. Chen, H. Liu, and G. Zhu, "Dynamic feature-fused localization with smartphones exploiting 5G NR SSB and Wi-Fi for indoor environments," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–14, Jan. 2024.

[196] T. Koike-Akino, P. Wang, and Y. Wang, "AutoQML: Automated quantum machine learning for Wi-Fi integrated sensing and communications," in *Proc. IEEE 12th Sensor Array Multichannel Signal Process. Workshop*, 2022, pp. 360–364.

[197] J. Xue, J. Zhang, Z. Gao, and W. Xiao, "Enhanced WiFi CSI fingerprints for device-free localization with deep learning representations," *IEEE Sensors J.*, vol. 23, no. 3, pp. 2750–2759, Feb. 2023.

[198] X. Zhou, L. Chen, Y. Ruan, and R. Chen, "Indoor localization with multi-beam of 5G new radio signals," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11260–11275, Sep. 2024.

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

[200] J. Wang et al., "Generative artificial intelligence assisted wireless sensing: Human flow detection in practical communication environments," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 10, pp. 2737–2753, Oct. 2024.

[201] N. Yoon, W. Jung, and H. Kim, "DeepRSSI: Generative model for fingerprint-based Localization," *IEEE Access*, vol. 12, pp. 66196–66213, 2024.

[202] X. Rao, Z. Luo, Y. Luo, Y. Yi, G. Lei, and Y. Cao, "MFFALoc: CSI-based multifeatures fusion adaptive device-free passive indoor fingerprinting Localization," *IEEE Internet Things J.*, vol. 11, no. 8, pp. 14100–14114, Apr. 2024.

[203] L. Ismail, D. Niyato, S. Sun, D. T. Hoang, D. I. Kim, and Y.-C. Liang, "Jamming mitigation in JRC systems via deep reinforcement learning and backscatter-supported intelligent deception strategy," in *Proc. IEEE 6th Int. Conf. Comput. Commun. Syst.*, 2021, pp. 1053–1058.

[204] M. Huang, F. Gong, G. Li, N. Zhang, and Q.-V. Pham, "Secure precoding for satellite NOMA-aided integrated sensing and communication," *IEEE Internet Things J.*, vol. 11, no. 18, pp. 29533–29545, Sep. 2024.

[205] Z. Zhu, M. Gong, G. Sun, P. Liu, and D. Mi, "AI-enabled STAR-RIS aided MISO ISAC secure communications," *Tsinghua Sci. Technol.*, vol. 30, no. 3, pp. 998–1011, Jun. 2025.

[206] Q. Liu, Y. Zhu, M. Li, R. Liu, Y. Liu, and Z. Lu, "DRL-based secrecy rate optimization for RIS-assisted secure ISAC systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16871–16875, Dec. 2023.

[207] C. O. Nnamani and M. Sellathurai, "Interference and noise cancelation for joint communication radar (JCR) system based on contextual information," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1855–1865, 2023.

[208] J. Yang, X. Gong, B. Ai, and W. Chen, "A priori based deep unfolding method for mmWave channel estimation in MIMO radar aided V2X communications," in *Proc. IEEE Int. Conf. Commun.*, 2023, pp. 2946–2951.

[209] Q. Ren and A. Willis, "Extending MUSIC to single snapshot and on line direction finding applications," in *Proc. Radar*, Edinburgh, U.K., 1997, pp. 783–787.

[210] Y. Shi, S. Xia, Y. Zhou, and Y. Shi, "Sparse signal processing for massive device connectivity via deep learning," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2020, pp. 1–6.

[211] P. Gupta and S. Kar, "MUSIC and improved MUSIC algorithm to estimate direction of arrival," in *Proc. Int. Conf. Commun. Signal Process.*, 2015, pp. 757–761.

[212] P. Zheng, Y. Zhu, H. Yulin, and A. Schmeink, "Over-the-air federated learning client selection in integrated sensing, computing and communication," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2024, pp. 804–809.

[213] X. Liu, H. Zhang, C. Ren, H. Li, C. Sun, and V. C. M. Leung, "Multi-task learning resource allocation in federated integrated sensing and communication networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11612–11623, Sep. 2024.

[214] J.-A. Désidéri, "Multiple-gradient descent algorithm (MGDA) for multiobjective optimization," *Comptes Rendus Mathematique*, vol. 350, nos. 5–6, pp. 313–318, 2012.

[215] H. H. Le, N. H. K. Long, N. C. Luong, N. T. Hoa, T. X. Nam, and D. I. Kim, "Edge computing and wireless power transfer for integrated radar and communication-equipped IoT systems," *IEEE Wireless Commun. Lett.*, vol. 13, no. 9, pp. 2457–2461, Sep. 2024.

[216] V. Mnih et al., "Asynchronous methods for deep reinforcement learning," in *Proc. 33rd Int. Conf. Mach. Learn. (ICML)*, 2016, pp. 1928–1937.

[217] T. H. Nguyen et al., "Coded distributed computing for vehicular edge computing with dual-function radar communication," *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 15318–15331, Oct. 2024.

[218] Y. Liang, H. Tang, H. Wu, Y. Wang, and P. Jiao, "Lyapunov-guided offloading optimization based on soft actor-critic for ISAC-aided internet of vehicles," *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 14708–14721, Dec. 2024.

[219] S. Lee, J.-Y. Lee, and S.-C. Kim, "Mutual interference suppression using wavelet denoising in automotive FMCW radar systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 887–897, Feb. 2019.

[220] I. Ajayi and W. Njima, "Radar interference mitigation using CNN autoencoders with spatial attention," in *Proc. 19th Int. Symp. Wireless Commun. Syst.*, 2024, pp. 1–6.

[221] T.-Y. Lim, S. A. Markowitz, and M. N. Do, "Radical: A synchronized FMCW radar, depth, IMU and RGB camera data dataset with low-level FMCW radar signals," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 4, pp. 941–953, Jun. 2021.

[222] J. Rock, M. Toth, E. Messner, P. Meissner, and F. Pernkopf, "Complex signal denoising and interference mitigation for automotive radar using convolutional neural networks," in *Proc. 22th Int. Conf. Inf. Fusion*, 2019, pp. 1–8.

[223] C. Duan and J. Zhang, "A modified deep learning based MIMO communication for integrated sensing, communication and computing systems," *Digit. Signal Process.*, vol. 142, Oct. 2023, Art. no. 104196.

[224] M. A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks," *Aiche J.*, vol. 37, no. 2, pp. 233–243, 1991.

[225] Z. Jing, J. Mu, and X. Jing, "Signal interference elimination based on autoencoder for integrated sensing and communication," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast.*, 2022, pp. 1–5.

[226] J. Johnston, Y. Li, M. Lops, and X. Wang, "ADMM-net for communication interference removal in stepped-frequency radar," *IEEE Trans. Signal Process.*, vol. 69, pp. 2818–2832, 2021.

[227] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. London, U.K.: Now Found. Trends, 2011.

[228] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2275–2284, Jun. 2009.

[229] M. Du, H. Zheng, M. Gao, X. Feng, J. Hu, and Y. Chen, "Integrated sensing, communication, and computation for over-the-air federated learning in 6G wireless networks," *IEEE Internet Things J.*, vol. 11, no. 21, pp. 35551–35567, Nov. 2024.

[230] N. C. Luong, Z. Xiong, P. Wang, and D. Niyato, "Optimal auction for edge computing resource management in mobile blockchain networks: A deep learning approach," in *Proc. IEEE Int. Conf. Commun.*, 2018, pp. 1–6.

[231] P. Dütting, Z. Feng, H. Narasimhan, D. Parkes, and S. S. Ravindranath, "Optimal auctions through deep learning," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 1706–1715.

[232] N. C. Luong, N. H. Sang, N. D. D. Anh, F. Shaohan, D. Niyato, and D. I. Kim, "Edge computing for metaverse: Incentive mechanism versus semantic communication," *IEEE Trans. Mobile Comput.*, vol. 23, no. 5, pp. 6196–6211, May 2024.

[233] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-splitting multiple access: Fundamentals, survey, and future research trends," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 4, pp. 2073–2126, 3rd Quart., 2022.

[234] C. Xu, B. Clerckx, S. Chen, Y. Mao, and J. Zhang, "Rate-splitting multiple access for multi-antenna joint communication and radar transmissions," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2020, pp. 1–6.

[235] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3083–3098, May 2022.

[236] M. Ahmed et al., "A survey on STAR-RIS: Use cases, recent advances, and future research challenges," *IEEE Internet Things J.*, vol. 10, no. 16, pp. 14689–14711, Aug. 2023.

[237] H. T. Nguyen, H. D. Tuan, D. Niyato, D. I. Kim, and H. V. Poor, "Improper Gaussian signaling for D2D communication coexisting MISO cellular networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5186–5198, Aug. 2021.

[238] H. D. Tuan, A. A. Nasir, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Non-orthogonal multiple access with improper gaussian signaling," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 496–507, Jun. 2019.

[239] J. Cong et al., "Near-field integrated sensing and communication: Opportunities and challenges," *IEEE Wireless Commun.*, vol. 31, no. 6, pp. 162–169, Dec. 2024.

[240] S. Xu, J. Zhang, Y. Xu, C. Li, and L. Yang, "Brownian bridge-based diffusion channel denoising for ISAC massive MIMO systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 2, pp. 764–776, Apr. 2025.

[241] Y. Jiang, F. Gao, S. Jin, and T. J. Cui, "Electromagnetic property sensing based on diffusion model in ISAC system," *IEEE Trans. Wireless Commun.*, vol. 24, no. 3, pp. 2036–2051, Mar. 2025.

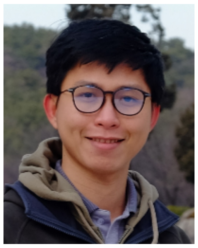
[242] W. Mao, Y. Lu, G. Pan, and B. Ai, "UAV-assisted communications in SAGIN-ISAC: Mobile user tracking and robust beamforming," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 1, pp. 186–200, Jan. 2025.

[243] S. Tao, M. Yuan, Q. Wu, R. Wang, and J. Hao, "Generative AI-aided vertical handover decision in SAGIN for IoT with integrated sensing and communication," *IEEE Internet Things J.*, vol. 12, no. 10, pp. 13297–13310, May 2025.

[244] S. Verma, Y. Kawamoto, N. Kato, T. Saiwai, and M. Yonehara, "An efficient beam searching in hybrid intelligent reflecting/refracting surfaces (IRS)-aided mmWave 6G network," *IEEE Trans. Veh. Technol.*, vol. 73, no. 12, pp. 19299–19312, Dec. 2024.

[245] Y. Kawamoto et al., "Simultaneous multiple connections and increased frequency efficiency using beam squint approach for IRS-based communication," *IEEE Trans. Veh. Technol.*, vol. 73, no. 11, pp. 17073–17082, Nov. 2024.

Nguyen Cong Luong received the B.S. and M.S. degrees from the School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Vietnam, and the Ph.D. degree from the Institut Galilée – Université Sorbonne Paris Nord, France. His research interest includes resource allocation and security management in the next generation networks.



Thien Huynh-The (Senior Member, IEEE) received the Ph.D. degree in computer science and engineering from Kyung Hee University (KHU), South Korea, in 2018, where he was a Postdoctoral Researcher with Ubiquitous Computing Laboratory from March 2018 to August 2018. From September 2018 to May 2022, he was a Postdoctoral Researcher with ICT Convergence Research Center, Kumoh National Institute of Technology, South Korea. He is currently a Lecturer with the Department of Computer and Communications Engineering, Ho Chi Minh City University of Technology and Education, Vietnam. His current research interests include digital image processing, radio signal processing, computer vision, wireless communications, IoT applications, machine learning, and deep learning. He was a recipient of the Superior Thesis Prize awarded by KHU. He was a recipient of the Golden Globe Award for Vietnamese Young Scientist in 2020 and the IEEE ATC Best Paper Award in 2023. He is currently serving as an Editor for IEEE COMMUNICATIONS LETTERS.

Thai-Hoc Vu (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Ulsan, South Korea, in 2025.

He is currently a Lecturer with the Institute of Information Technology, Digital Transformation, Thu Dau Mot University, Binh Duong, Vietnam. He specialises in the performance analysis of wireless communications systems, focusing on cutting-edge areas, such as non-orthogonal multiple access, reconfigurable intelligent surface, radio-frequency-based energy harvesting, physical-layer security, short-packet transmission, and applying convex optimization and machine learning to analyse, improve, and optimise advanced wireless communication networks and IoT systems. He has authored and co-authored more than 50 IEEE transactions/Journals and flagship IEEE conference papers. He was awarded the Vietnamese Young Scientists in Korea Award from the Vietnamese Students Association in South Korea and the Vietnam Embassy in the Republic of Korea in 2022. He completed his Ph.D. program with more than 20 Top IEEE publications as the first author and received the Best Research Award from BK21 from 2021 to 2025. He was a recipient of the Top Reviewer Award from IEEE COMMUNICATIONS LETTERS in 2023 and 2024, the IEEE ICCE Best Paper Student Awards in 2022 and 2024, and the IEEE ATC Best Paper Award in 2024.

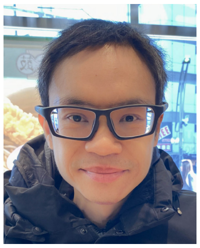
Duc Van Le (Senior Member, IEEE) received the B.Eng. degree (with Distinction) in electronics and telecommunications engineering from Le Quy Don Technical University, Vietnam, in 2011, and the Ph.D. degree in computer engineering from the University of Ulsan, South Korea, in 2016. He is currently a Postdoctoral Fellow with the School of Electrical Engineering and Telecommunications, University of New South Wales, Australia. From 2016 to 2024, he was a Senior Research Fellow with Nanyang Technological University, and a Research Fellow with the National University of Singapore. His research interests include sensor networks, Internet of Things, cyber-physical systems, and applied machine learning. He was a recipient of the ACM/IEEE ICCPS'23 Best Paper Award.

Huy T. Nguyen (Member, IEEE) received the B.Eng. degree (Hons.) in computer science and engineering from the HCMC University of Technology, Vietnam, in 2013, and the M.Eng. and Ph.D. degrees from the Department of Information and Communication System, Inje University, South Korea, in 2016 and 2019, respectively. He is currently a Lecturer with the Faculty of Information Technology, Van Lang University (VLU), where he also leads the Smart and Autonomous Systems Research Group. Before joining VLU, he served as a Chief Technology Officer-Asia Desk at Data Design Engineering, primarily developing data-driven AI solutions. His earlier career included significant contributions to the communication aspects of wireless smartgrid systems as a Scientist with the Institute for Infocomm Research (I2R-A*STAR) and Nanyang Technological University, Singapore. His research interests lie in applying AI/ML and optimization techniques to signal processing and wireless communications for 5G and beyond.

Nguyen Duc Hai (Student Member, IEEE) is currently pursuing the undergraduate degree with the Phenikaa School of Computing, Phenikaa University, Vietnam. His research interest includes deep learning, deep reinforcement learning, and game theory for computer networks and communication networks.

Gia-Vuong Nguyen received the B.S. degree in embedded systems and Internet of Things from the Ho Chi Minh City University of Technology and Education, Vietnam, in 2024. He is currently pursuing the graduate degree with the Department of Artificial Intelligence and Robotics, and the Department of Convergence Engineering for Intelligent Drone, Sejong University, South Korea. His current research interests include digital image processing, radio signal processing, computer vision, machine learning, and deep learning. He was a recipient of the REV-ECIT Best Paper Award in 2023.

Nguyen Duc Duy Anh is currently pursuing the undergraduate degree with the Phenikaa School of Computing, Phenikaa University, Vietnam. His research interest includes deep learning, deep reinforcement learning, and game theory for computer networks and communication networks.



Dusit Niyato (Fellow, IEEE) received the B.Eng. degree from the King Mongkuts Institute of Technology Ladkrabang, Thailand, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Canada. He is a Professor with the College of Computing and Data Science, Nanyang Technological University, Singapore. His research interests are in the areas of mobile generative AI, edge intelligence, decentralized machine learning, and incentive mechanism design.

Dong In Kim (Life Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1990. He was a Tenured Professor with the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada. He is currently a Distinguished Professor with the College of Information and Communication Engineering, Sungkyunkwan University, Suwon, South Korea. He was the first recipient of the NRF of Korea Engineering Research Center in Wireless Communications for RF Energy Harvesting from 2014 to 2021. He received several research awards, including the 2023 IEEE ComSoc Best Survey Paper Award and the 2022 IEEE Best Land Transportation Paper Award. He was selected the 2019 recipient of the IEEE ComSoc Joseph LoCicero Award for Exemplary Service to Publications. He has been listed as a 2020/2022 Highly Cited Researcher by Clarivate Analytics. He was the General Chair of the IEEE ICC 2022, Seoul. From 2001 to 2024, he serving as an Editor, an Editor at Large, and an Area Editor of Wireless Communications I for IEEE TRANSACTIONS ON COMMUNICATIONS. From 2002 to 2011, he served as an Editor and a Founding Area Editor of Cross-Layer Design and Optimization for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. From 2008 to 2011, he served as the Co-Editor-in-Chief for the *IEEE/KICS Journal of Communications and Networks*. He served as the Founding Editor-in-Chief for the IEEE WIRELESS COMMUNICATIONS LETTERS from 2012 to 2015. He is a Fellow of the Korean Academy of Science and Technology and a life member of the National Academy of Engineering of Korea.

Quoc-Viet Pham (Senior Member, IEEE) received the B.Sc. degree in telecommunications engineering from the Hanoi University of Science and Technology in 2013, and the Ph.D. degree (with Best Ph.D. Dissertation Award) in telecommunications engineering from Inje University in 2017.

He is currently an Assistant Professor of Networks and Distributed Systems with the School of Computer Science and Statistics, Trinity College Dublin, Ireland. He has special research interests in the areas of wireless AI, edge computing, Internet of Things, and distributed learning.

Dr. Pham was a recipient of the IEEE TVT Top Reviewer Award in 2020, the Golden Globe Award in Science and Technology for Vietnam's Young Researchers in 2021, the IEEE ATC Best Paper Award in 2022, the IEEE MCE Best Paper Award in 2023, the IEEE M-COMSTD Exemplary Editor Award in 2024, and a Clarivate Highly Cited Researcher Award in 2024. He was honoured with the IEEE ComSoc Best Young Researcher Award for EMEA 2023 in recognition of his research activities for the benefit of the Society. He was a Lead Guest Editor of the IEEE INTERNET OF THINGS JOURNAL special issue on Aerial Computing for the Internet of Things. He currently serves as an Editor for IEEE COMMUNICATIONS LETTERS, *IEEE Communications Standards Magazine*, IEEE COMMUNICATIONS SURVEYS AND TUTORIALS, *Journal of Network and Computer Applications*, IEEE TRANSACTIONS ON MOBILE COMPUTING, and IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING.