

Article

Not peer-reviewed version

Deep Learning-Empowered RF Sensing in Outdoor Environments: Recent Advances, Challenges, and Future Directions

[Duc Minh Quang Nguyen](#) , William D. Lukito , Xuemeng Liu , [Chang Liu](#) *

Posted Date: 10 December 2024

doi: 10.20944/preprints202412.0809.v1

Keywords: RF signals; sensing; radio frequency; outdoor environments; deep learning; wireless technologies



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Deep Learning-Empowered RF Sensing in Outdoor Environments: Recent Advances, Challenges, and Future Directions

Duc Minh Quang Nguyen ¹, William D. Lukito ¹, Xuemeng Liu ² and Chang Liu ^{1,*}

¹ School of Computing, Engineering and Mathematical Sciences, La Trobe University, Melbourne, VIC 3086, Australia

² School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia

* Correspondence: c.liu6@latrobe.edu.au

Abstract: Recently, with advancements in Deep Learning (DL) technology, Radio Frequency (RF) sensing has seen substantial improvements, particularly in outdoor applications. Motivated by these developments, this survey presents a comprehensive review of state-of-the-art RF sensing techniques in challenging outdoor scenarios with practical issues such as fading, interference, and environmental dynamics. We first investigate the characteristics of outdoor environments and explore potential wireless technologies. Then, we study the current trends in applying DL to RF-based systems and highlight its advantages in dealing with large-scale and dynamic outdoor environments. Furthermore, this paper provides a detailed comparison between discriminative and generative DL models in support of RF sensing, offering insights into both the theoretical underpinnings and practical applications of these technologies. Finally, we discuss the research challenges and present future directions of leveraging DL in outdoor RF sensing.

Keywords: RF signals; sensing; radio frequency; outdoor environments; deep learning; wireless technologies

1. Introduction

RF sensing is a technology that utilizes wireless signals to detect and monitor objects, activities, or environmental changes. By interpreting changes in RF signal properties, such as amplitude, phase, and frequency, it can extract meaningful information about the surrounding environment [1]. This approach encompasses various platforms, including LoRa ("Long Range"), Wi-Fi, radar, Software-Defined Radio (SDR), and Radio Frequency Identification (RFID), each enabling specific methods for sensing applications [2]. RF sensing is known for being non-intrusive, cost-effective, and energy-efficient, making it a practical solution for monitoring multiple subjects simultaneously in large areas [3].

In recent years, RF sensing has expanded into outdoor environments, where it supports applications in various sectors [4–7]. RF sensing enables human activity detection for applications such as crowd management, movement tracking, and public safety, helping authorities monitor pedestrian flows during large events [1]. These systems rely on RF signals to detect the presence, movement patterns, and locations of individuals, enhancing urban traffic management [8] and security. Similarly, vehicles and unmanned aerial vehicles (UAVs) could benefit from RF sensing through technologies like millimeter wave (mmWave) radar, which is able to provide real-time localization, object detection, and tracking. Additionally, RF sensing aids environmental monitoring by providing insights into air quality, temperature, and pollution levels, contributing to sustainable urban development [9]. However, outdoor environments pose unique challenges for RF sensing, including multipath interference, signal attenuation, and environmental variability [10]. Obstacles like buildings, vehicles, and natural terrain disrupt RF signals, causing reflection, scattering, or attenuation. Additionally, harsh weather conditions, including rain and high humidity, can negatively impact signal quality. [11]. Furthermore, overlapping signals from other wireless systems operating in the same environment introduce significant interference, complicating the reliability of RF sensing [12,13]. To overcome those challenges, various signal processing techniques or machine learning algorithms have been applied, enhancing RF sensing capabilities in complex outdoor environments [13,14].

Deep Learning (DL), a subcategory of machine learning, has revolutionized the way data-driven solutions address these complex challenges. Over the past decade, advancements in DL have demonstrated exceptional performance across diverse domains [15–19], driven by innovations in model architectures such as Multilayer Perception (MLPs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and generative models like Autoencoders (AEs), Generative Adversarial Networks (GANs), Diffusion Models (DMs) and Large Language Models (LLMs). These models have shown remarkable adaptability and scalability, enabling them to process and analyze high-dimensional data with minimal manual intervention [20,21]. DL simplifies the process of traditional feature engineering by hierarchically representing raw data, resulting in more reliable solutions. It has emerged as a revolutionary approach in RF sensing, automating the extraction of complex features from raw signals and removing the reliance on manually crafted feature designs. Discriminative models like CNNs and transformers excel in tasks such as signal classification [22], anomaly detection [23], and behavioral prediction [24–27], while generative approaches, including GANs and DMs, enhance detection capabilities in noisy or low-data scenarios by modeling signal distributions. DL bridges the gap between traditional signal processing and intelligent automation, driving advancements in applications ranging from human activity recognition and health monitoring [1,6,28] to spectrum management and anomaly detection in communication systems [7,8,29–31]. These models adapt to dynamic environments, enabling continuous innovation and improving RF sensing performance in complex outdoor scenarios [1,4].

The main objective of this review is to systematically explore three key research questions that are crucial to advancing RF sensing in outdoor environments. Specifically, this work seeks to explore: (1) the recent advancements in outdoor RF sensing and their effectiveness in addressing unique challenges such as signal interference and environmental variability; (2) the wireless technologies commonly employed in outdoor applications and their suitability for various sensing requirements; and (3) the role of DL in enhancing RF sensing performance, particularly regarding accuracy and adaptability in complex outdoor settings. By thoroughly examining these questions, this review offers a structured framework for understanding the current landscape and future directions of DL-enabled RF sensing in outdoor applications. While recent surveys have extensively covered areas such as human sensing [1,6], smart homes [4], and indoor localization [32,33], their focus has primarily been on controlled indoor environments or specific applications. These reviews often overlook the unique and complex challenges of large-scale, dynamic outdoor environments. This gap emphasizes the necessity for a comprehensive review that examines existing efforts to address these challenges while exploring how DL can further enhance RF sensing performance. As far as we are aware, this is the first in-depth review to explore DL-empowered RF sensing in outdoor environments. The main contributions of this paper are summarized as follows:

1. Unlike existing surveys, this survey provides a comprehensive review of RF sensing in outdoor environments, identifying key challenges and examining wireless technologies best suited for these settings.
2. We offer a detailed analysis of DL approaches, focusing on both generative and discriminative models, and assess their effectiveness in enhancing RF sensing. We also review recent outdoor RF sensing studies utilizing various DL methods, categorizing them by approach, and underlining the specific benefits and limitations of each in distinct scenarios.
3. This survey paper explores the existing challenges of leveraging DL in outdoor RF sensing and presents insights and possible solutions for future tendencies.

The structure of this paper is shown in Figure 1. Following this introduction, Section 2 provides an overview of RF sensing in outdoor environments, with a focus on the unique challenges and the wireless technologies best suited to these settings. Building on this, Section 3 explores how DL models enhance RF sensing, categorizing prominent approaches and assessing their practical applications. Section 4 then addresses the core challenges in integrating DL with RF sensing, offering future research

directions to address these issues. Finally, Section 5 encapsulates the key insights, underscoring potential avenues for advancing RF sensing in outdoor environments.

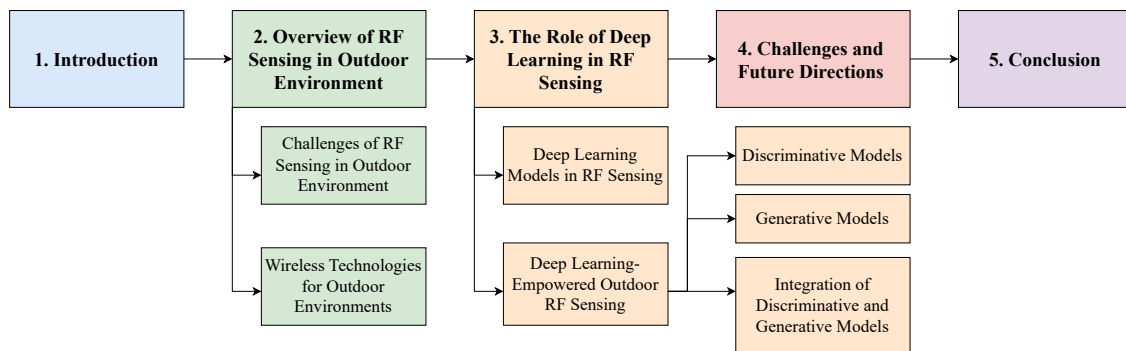


Figure 1. Organization of this survey.

2. Overview of RF Sensing in Outdoor Environments

RF sensing uses wireless signals to detect and monitor objects, activities, or environmental changes by analyzing variations in properties such as amplitude, phase, and frequency. However, outdoor environments introduce unique challenges like fading and interference, which complicate control and significantly affect accuracy and reliability. Therefore, it is essential to investigate these challenges to develop effective solutions. This section examines the unique challenges that RF sensing systems face in outdoor environments, as depicted in Figure 2. We begin by discussing specific issues such as multipath effects, interference by other wireless systems, signal attenuation, and environmental variability. Following this, we introduce wireless technologies best suited to address these challenges and ensure reliable performance.

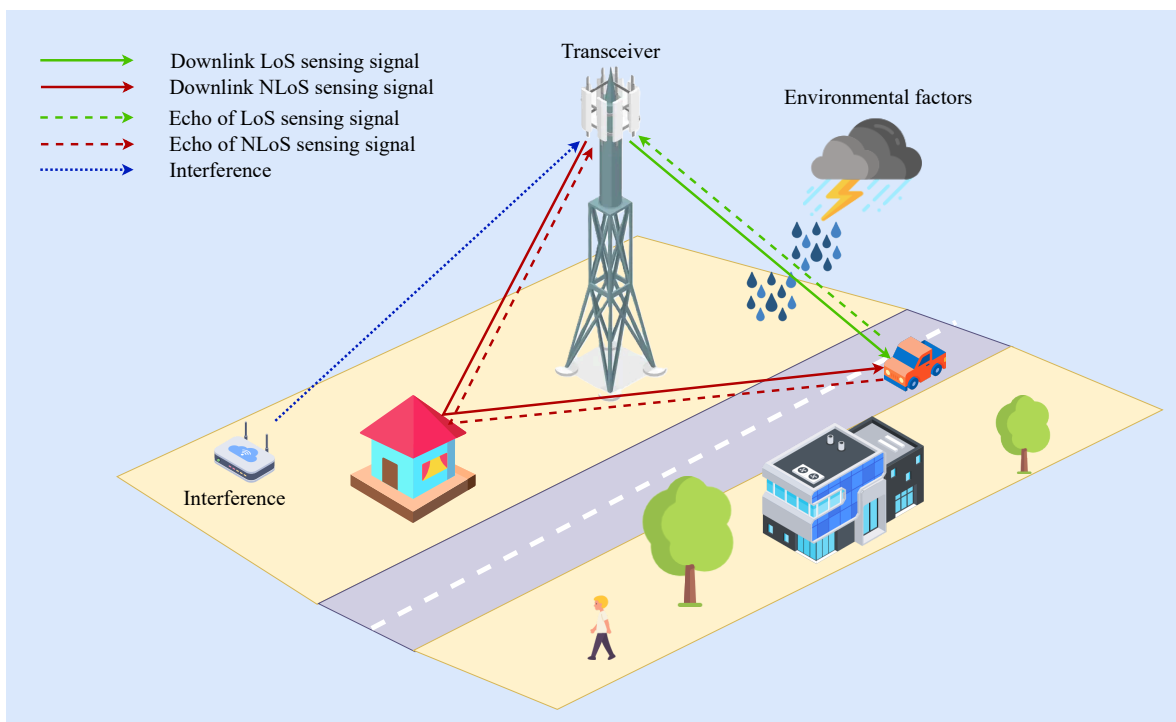


Figure 2. An illustration of the challenges of RF sensing in outdoor environments, including multipath interference, and environmental factors. It depicts how storm clouds obstruct line-of-sight (LoS) sensing signals, degrading signal quality. Additionally, it shows multipath effects from transceiver signals reflecting off objects and interference from other wireless systems.

2.1. Challenges of RF Sensing in Outdoor Environments

RF sensing systems deployed in outdoor environments face numerous challenges that are not present or easily managed in indoor settings. In outdoor scenarios, environmental conditions are dynamic and difficult to control, leading to fluctuations in signal quality and reliability [34,35]. These challenges arise from several key factors, including signal attenuation, multipath propagation, non-LoS (NLoS) conditions, interference by other wireless systems, and environmental variability.

One of the primary challenges in RF signal transmission is attenuation, where signal strength diminishes with distance. In free space, path loss can be represented by

$$L = \left(\frac{4\pi df}{c} \right)^2, \quad (1)$$

where d is the distance between transmitter and receiver, f is the frequency, and c is the speed of light. This equation quantifies the predictable loss in signal strength as a function of distance and frequency in a vacuum or unobstructed space. However, real-world outdoor environments further exacerbate attenuation due to natural obstacles such as trees, hills, and water bodies that can absorb or scatter signals, reducing effective transmission range. Urban areas introduce additional complexity with dense structures like buildings and vehicles that block or weaken signals.

In these environments, attenuation is affected by both large-scale and small-scale fading. Large-scale fading describes the path loss caused by terrain and significant obstacles over extended distances. On the other hand, small-scale fading involves quick variations in signal strength due to multipath propagation, which occurs as signals reflect off nearby objects like walls and vehicles [35]. These fading effects are essential to understanding RF performance across diverse environments.

A significant challenge in RF sensing is multipath propagation, which complicates reliable signal interpretation. In multipath environments, signals reflect off surfaces like buildings, vehicles, and the ground before reaching the receiver, creating NLoS conditions, as shown in Figure 2. As illustrated in Figure 3, each reflected signal can be modeled as a delayed, attenuated copy of the original, known as a delay tap. These overlapping signal paths interfere with each other at the receiver, causing distortions that obscure the true features of the original signal. This interference degrades the accuracy of the received signal, making it challenging to reliably interpret the original transmission or reconstruct the sensed environment [35]. Multipath propagation is particularly problematic in urban areas, where abundant reflective surfaces amplify this overlapping effect.

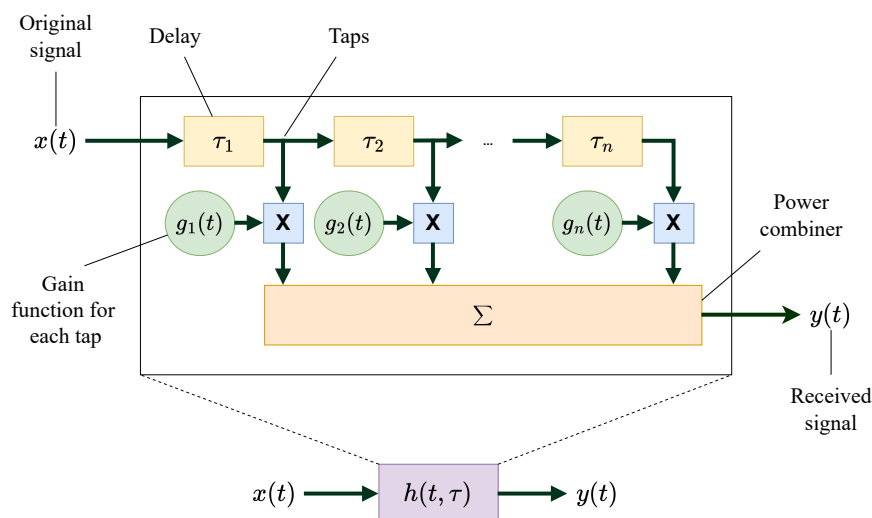


Figure 3. An illustration of multipath propagation.

For addressing multipath propagation challenges, reconfigurable intelligent surfaces (RIS) can indeed be beneficial, as they offer a way to create a controlled virtual LoS [36]. RIS technology can effectively mitigate multipath fading by redirecting RF signals in real-time and enhancing signal strength at the receiver [37]. However, optimizing RIS configurations for high accuracy of sensing tasks such as human posture recognition remains challenging [38]. Additionally, challenges such as the area and bandwidth of influence, as highlighted by Alexandropoulos et al. [39], require careful planning for RIS deployment in smart wireless environments.

Interference from other wireless systems is another challenge that can degrade the performance of outdoor RF sensing. Additionally, the accumulation of communication technologies such as Wi-Fi networks, cellular towers, and IoT devices, creates overlapping signals in the same frequency bands, leading to noise which causes signal degradation [40]. This interference reduces the reliability of measurements and increases the likelihood of errors, particularly in crowded urban areas where spectrum congestion is common.

In addition to interference, environmental factors such as weather conditions and moving objects introduce further variability. Temperature, humidity, and precipitation affect signal propagation by altering the path and strength of the signals [11,41]. For example, rain or fog can absorb or scatter RF signals, causing attenuation and reducing the effective range of the system. Furthermore, outdoor environments are dynamic, with constantly changing elements such as moving vehicles, and pedestrians. These temporary obstructions can significantly affect signal paths in RF sensing applications. For example, vehicle vibrations and nonlinear movements can degrade automotive radar sensor signals, impacting accuracy and detection probability [42].

2.2. Wireless Technologies for Outdoor Environments

To address those challenges mentioned above, the selection of appropriate wireless technologies is critical. Certain wireless technologies are less suitable for outdoor RF sensing applications due to limitations in range, susceptibility to interference, or insufficient robustness under varying conditions. Table 1 compares current wireless technologies, underlining the sensing range and power characteristics of each technology in experimental settings. For example, although Bluetooth is designed to operate effectively in noisy environments and can handle fading and interference, its relatively short range makes it less ideal for large-scale outdoor RF sensing [43]. ZigBee, while advantageous for low-power applications, faces challenges in outdoor settings due to interference, range limitations, and vulnerability to multipath effects and environmental fluctuations, limiting its effectiveness [44]. In the following section, we discuss some of the best wireless technology candidates for RF sensing in outdoor environments.

Table 1. Comparison of Wireless Technologies for Sensing Applications

Name	Sensing Range	Transmission Power	Operating Frequency
LoRa [45,46]	Up to 15 km	Up to 20 dBm	433 MHz, 868 MHz, 915 MHz
mmWave [4,47]	Up to 500 m	30–40 dBm	30–300 GHz
LTE	Up to 100 m	23–43 dBm	450 MHz–3.8 GHz
Wi-Fi [48]	Up to 100 m	Up to 30 dBm	2.4 GHz, 5 GHz, 6 GHz
RFID [49]	Up to 10 cm	N/A	125–134 kHz (Low Frequency)
	Up to 1 m	N/A	13.56 MHz (High Frequency)
	Up to 10 m	N/A	860–960 MHz (Ultra-High Frequency)
UWB [50]	Up to 200 m	–41.3 dBm	3.1–10.6 GHz
Terahertz [51]	Up to 10 m	N/A	0.3–3 THz
ZigBee [52]	Up to 100 m	Up to 20 dBm	2.4 GHz
Bluetooth [43]	Up to 100 m	0–20 dBm	2.4 GHz

Long Range (LoRa) [45] is a physical propriety technique based on spread spectrum modulation techniques derived from chirp spread spectrum (CSS) technology, offering distinct advantages for RF-based sensing, particularly in large-scale applications. With an extended range often spanning several kilometers, LoRa effectively addresses the range limitations common in expansive outdoor environments. Its CSS modulation and various spreading factors provide resilience to intra-technology interference and some tolerance to multipath effects [53]. However, despite this robustness, LoRa remains vulnerable to inter-technology interference from other devices sharing unlicensed frequency bands [54]. LoRa's minimal power consumption makes it especially well-suited for applications that require continuous battery operation in sensing devices. Moreover, its affordability and ease of implementation increase accessibility for researchers and practitioners [55]. Nevertheless, the small bandwidth of LoRa limits its capacity to capture and transmit detailed information. As a result, LoRa is commonly used in scenarios where low data rates are sufficient, such as object localization [56] and environmental monitoring [57].

Millimeter-Wave (mmWave) technology is an advanced wireless communication technology operating in the frequency range of 30-300 GHz. The directional character of mmWave signals, meaning they travel in focused, narrow beams rather than spreading broadly, helps to reduce interference from other wireless systems, enhancing reliability in densely populated areas. Additionally, mmWave radars are effective in various weather conditions and NLoS scenarios, making them a viable alternative to sensor-based methods like cameras and LiDAR in complex environments [1]. Its high bandwidth enables the capturing of detailed data, significantly enhancing the resolution and accuracy of sensing applications compared to technologies like LoRa, Wi-Fi, and traditional RF systems. This makes mmWave ideal for high-precision tasks that require detailed environmental information [4]. Furthermore, In localization, Hao et al. [58] proposed a mmWave-based multipath-assisted localization model that leverages multipath effects to improve indoor localization accuracy while traditional models often filter out these effects to minimize errors. This approach utilizes the high spatial resolution of mmWave to achieve precise position estimates. Despite these advantages, mmWave technology faces limitations in outdoor environments, particularly reduced range due to signal attenuation at high frequencies. This relationship is explained by Equation 1, where path loss increases with rising frequency. Thus, while mmWave excels in short-range, high-resolution sensing, its use in large-scale outdoor settings requires careful planning. Common applications include high-resolution localization [59], human sensing [1], and automotive radar for object detection in autonomous vehicles [4].

Long-Term Evolution (LTE) is a widely used wireless communication technology with emerging applications in RF sensing, particularly for outdoor environments. Originally designed for communication, LTE operates over a broad frequency range (450 MHz–3.8 GHz) with extensive coverage, which makes it adaptable for sensing tasks. Recent studies have utilized LTE to detect environmental changes, such as traffic and human activity, by analyzing signal reflections and interference patterns [60–62]. This approach allows the reuse of existing cellular infrastructure, potentially reducing additional sensor costs. However, LTE faces challenges in sensing specific targets due to interference from non-target objects, such as trees, vehicles, and pedestrians, over long distances. These factors, along with multipath effects, can degrade signal quality [60]. Additionally, integrating sensor networks into LTE may lead to network overload in dense areas [63]. Despite these issues, LTE's vast infrastructure and broad coverage make it a promising option for large-scale outdoor sensing where extended connectivity is essential.

Wireless Fidelity (Wi-Fi) operates in the 2.4 GHz, 5 GHz, and 6 GHz frequency bands, as defined by IEEE standards 802.11b/g/n (2.4 GHz), 802.11a/n/ac (5 GHz), and 802.11ax (6 GHz) [48]. Due to advantages such as affordability and widespread availability, Wi-Fi sensing has emerged as a versatile technology for applications like human activity recognition and movement tracking. Compared to radar, Wi-Fi-based sensing provides extensive coverage with fewer blind spots [6]. Additionally, Wi-Fi signals are able to penetrate walls and other obstacles, making them suitable for through-the-wall sensing tasks such as human presence detection. This capability, along with compatibility with

commodity devices like smartphones, enables activity sensing within rooms from outdoor locations without the need for specialized or invasive equipment [64]. However, Wi-Fi signals are highly susceptible to various factors, including environmental variability and multipath effects, particularly in multi-object scenarios [6]. Ensuring reliable Wi-Fi sensing across diverse real-world settings is challenging, as it involves complex optimization problems. Moreover, Wi-Fi was not originally intended for sensing; thus, using it for such applications can degrade communication performance due to network interference and limited resources [65].

Radio Frequency Identification (RFID) technology, encompassing both passive and active systems, is widely utilized in RF sensing applications for object tracking and environmental monitoring [66]. An RFID system comprises tags, readers, and antennas. Passive tags, which rely on energy from the RFID reader to operate, are advantageous for long-term monitoring applications, such as environmental sensing or asset tracking, due to their low power requirements [67]. However, passive tags are more susceptible to environmental factors like rain, humidity, and temperature variations, which can cause phase drift and signal attenuation, reducing accuracy in outdoor location tracking [68]. In contrast, active RFID tags have an internal power source, enabling them to broadcast signals continuously or at specified intervals. This results in a greater range and improved signal strength, making them especially suitable for outdoor applications with significant distances between the tag and reader. Due to their ability to emit signals and operate at higher frequencies than passive RFID, active RFID tags can also be effectively used in localization systems [69].

Ultra-Wideband (UWB) radar has been employed in a variety of military and civilian contexts for high-resolution sensing and imaging [70]. Recently, it has gained recognition as an effective solution for accurate localization, particularly in scenarios where global navigation satellite systems (GNSS) are unavailable, due to its exceptional precision in range estimation [71,72]. Compared to other common localization technologies like Bluetooth, Zigbee, and radio frequency identification (RFID), UWB offers improved ranging accuracy, effective multipath resolution, and strong resistance to interference, attributed to its wide bandwidth of 800 MHz [50]. This capability is particularly valuable in scenarios requiring high-precision localization [73]. Despite these advantages, UWB's range is limited compared to other RF sensing technologies like LoRa or LTE, making it more suitable for localized sensing tasks rather than long-range applications.

Terahertz (THz) radiation occupies the electromagnetic spectrum between the microwave and far-infrared ranges, serving as a link between electronics and optics within the so-called "terahertz gap" [74]. THz wavelengths are generally defined as ranging from 1.0 to 0.1 mm, corresponding to frequencies of 300 GHz to 3 THz [75]. The short wavelengths of THz radiation enable high-resolution imaging, which is beneficial for detailed inspections and for identifying fine structures in materials. Additionally, THz photon energies are far lower than those of X-rays, making them non-ionizing and safe for human-involved applications, such as imaging and security screening [75]. However, THz radiation has a significant drawback: high signal loss. THz waves experience greater free-space path loss than lower frequency bands, as path loss increases with the square of the frequency as illustrated in Equation 1. As a consequence, due to its high-resolution sensing capabilities on a small scale, THz radiation is applied in imaging [76], food quality inspection [77] and various pharmaceutical industry processes [78]. Finally, Naftaly et al. [79] reviewed both current and emerging industrial applications, highlighting the significant market potential of THz technology.

In summary, each wireless technology presents distinct characteristics, strengths, and limitations. Selecting an appropriate wireless technology based on the specific requirements of outdoor sensing tasks is essential for addressing challenges in various scenarios, such as the need for long-distance sensing or the capture of detailed environmental information. The selected technology should achieve an optimal balance of cost-effectiveness, durability, and performance to support effective and reliable outdoor RF sensing. Wireless technology serves as the foundation for these systems, but the implementation approach is equally critical. Traditional methods often face performance limitations in complex outdoor environments. In contrast, DL exploits a data-driven approach to offer significant potential to

enhance accuracy and robustness, enabling more reliable and adaptive RF sensing solutions, which will be discussed in the subsequent section.

3. The Role of Deep Learning in RF Sensing

Over the past decade, DL has made significant strides [15–19], influencing a wide range of fields, including RF sensing [6,13,14,28]. While wireless technologies provide the foundational infrastructure for outdoor RF sensing, the performance of traditional approaches remains limited in addressing the complexities of outdoor environments. DL addresses these limitations by automatically identifying patterns and extracting essential features, enabling improved accuracy and adaptability in outdoor scenarios. The following section explores the role of DL in RF sensing, focusing on common DL models and analyzing recent studies that leverage these techniques for outdoor applications.

3.1. Deep Learning Models in RF Sensing

Before discussing model architectures, it is essential to differentiate between two key approaches to learning patterns in data: discriminative and generative models [80]. Discriminative models aim to model the connections between input and output variables by directly modeling the conditional probability $P(Y|X)$, where X describes the input and Y the output. This method allows them to predict the outcomes corresponding to observed data, making them effective for classification tasks. In contrast, generative models seek to understand the data generation process by modeling the joint probability $P(X, Y)$. They model the underlying data distribution, allowing them to produce new samples that closely mirror the input data [80]. This distinction is crucial in RF sensing: while discriminative models are effective for various prediction tasks, such as detecting human presence or specific movements from RF signals, generative models excel in simulating or reconstructing RF environments. They allow for more accurate predictions, particularly in scenarios with limited data or unseen environments [81]. In the following sections, we will briefly explain the architecture and mechanisms of popular DL models used in RF Sensing, as illustrated by Figure 4 and Figure 5. Table 2 presents the type of approach, objectives, advantages, and disadvantages of each model.

Table 2. Comparison of Model Types in RF Sensing

Type of approach	Model	Objectives	Advantages	Disadvantages
Discriminative	MLPs	Classification, regression	Simple architecture, easy to implement, efficient for small datasets	Limited capacity for spatial/temporal information, not scalable for complex tasks
	CNNs	Signal representation, feature extraction	Good at extracting spatial features	Limited for temporal information without additional structures
	RNNs	Sequential signal analysis, time-series prediction	Handles sequential and temporal dependencies well	Prone to vanishing/exploding gradient problems, less efficient for long sequences
Generative	AEs	Dimensionality reduction, anomaly detection	Good for feature extraction, data compression	Poor reconstruction with complex signals, requires tuning of latent space size
	GANs	RF signal generation, data augmentation, anomaly detection	Capable of generating realistic data	Difficult to train, sensitive to hyperparameters
	DMs	Signal denoising, enhancement, and generative modeling	High quality in denoising and generating diverse data, robust training	Computationally intensive, slow to generate outputs compared to GANs
	LLMs	Cross-modal RF sensing, sequence modeling	Excellent for capturing long-range dependencies, scalable, adaptable to different tasks (e.g., classification, localization)	Requires large datasets or well-pre-trained models, computationally expensive

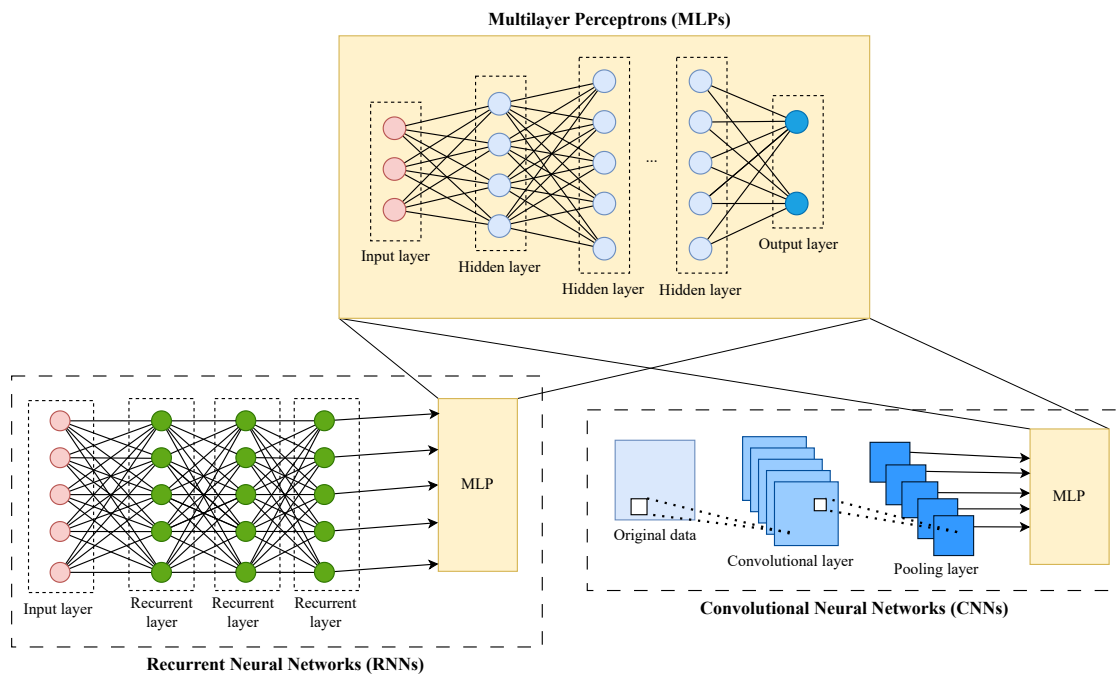


Figure 4. Representative architectures of discriminative DL models commonly applied in RF sensing.

Multilayer Perceptrons (MLPs) form the core of neural network architectures, comprising an input layer, multiple hidden layers, and an output layer, all interconnected as depicted in Figure 4. Each neuron in a layer connects to every neuron in the subsequent layer. To capture non-linear relationships, MLPs utilize activation functions such as ReLU or sigmoid, enabling the learning of complex patterns. Data flows through the network in a feedforward fashion, starting from the input layer and passing through hidden layers to the output layer without any loops or recurrences. Training involves backpropagation, which updates weights and biases by minimizing the loss function based on the error between actual and predicted outputs. MLPs are well-suited for structured data and simple tasks like classification. In RF sensing, they can utilize channel state information (CSI) for downstream tasks such as localization [82] and activity recognition [83].

Convolutional Neural Networks (CNNs) are specifically designed to analyze and extract spatial features from high-resolution 2-dimensional matrices. Building upon the MLP architecture [84], CNNs incorporate convolutional and pooling layers, which are particularly effective in reducing the output size of each layer, especially for inputs with high dimensions, as illustrated in Figure 4. Convolutional layers utilize filters (kernels) to scan the input image and identify patterns such as edges, textures, or shapes, generating feature maps that emphasize various characteristics of the input. Pooling layers follow, reducing the spatial dimensions of these feature maps while retaining critical features and lowering computational demands. A fully connected layer is then used to link every neuron from the previous layer to those in the next, producing the final classification output. This hierarchical feature extraction enables CNNs to excel in tasks like image classification, object detection, and segmentation by progressively capturing more abstract features at deeper layers [85]. Moreover, when RF data is structured appropriately, it can display spatial patterns similar to those in large-scale 2-dimensional matrices, making CNNs highly effective for RF representation learning [24,86].

Recurrent Neural Networks (RNNs) are designed to address sequence prediction tasks by incorporating recurrent layers, where the output of a neuron at one-time step is used as input for the same neuron at the next step. This structure creates a hidden state that acts as a memory, carrying information across the sequence. Despite their utility, RNNs face challenges such as vanishing and exploding gradients, which can lead to diminishing or excessively growing weights, complicating

training, and reducing their overall effectiveness [87]. To overcome these issues, advanced variants like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks were developed. These models introduce mechanisms to handle long-term dependencies in sequential data by replacing the traditional single non-linear activation function with multiple specialized functions and incorporating copying and concatenation mechanisms, enabling the retention of crucial information over extended sequences. Despite these advancements, RNNs still face challenges, such as sequential processing and limited memory capacity, which restrict parallel computation and complicate the modeling of very long-term dependencies, even with LSTM and GRU networks [88]. Nonetheless, RNNs remain valuable for applications involving RF time-series data, such as activity recognition [89,90].

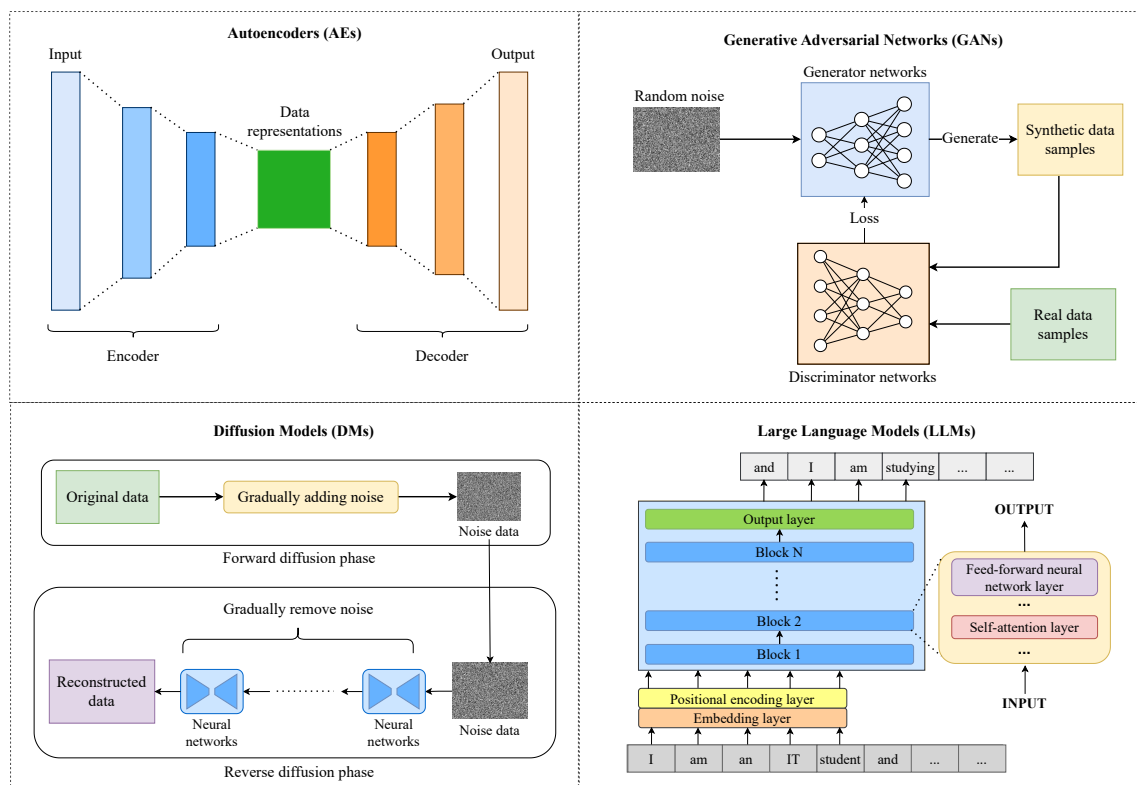


Figure 5. Representative architectures of generative DL models commonly applied in RF sensing.

Autoencoders (AEs) [91] are unsupervised learning neural networks designed to learn compact data representations through compression and reconstruction. Their architecture is composed of two main parts: an encoder and a decoder, as illustrated in Figure 5. The encoder compresses the input data into a lower-dimensional latent space, capturing its key features, while the decoder reconstructs the original data from this reduced representation. The network is trained by minimizing reconstruction loss, which quantifies the difference between the input and its reconstruction, enabling Autoencoders to effectively capture meaningful and concise representations of the data. This capability makes AEs well-suited for tasks like dimensionality reduction and anomaly detection [92]. Advanced versions, such as convolutional and denoising autoencoders, enhance the basic design for specialized applications, including filling gaps in radio maps [93] or removing noise from RF signal data [94].

Generative Adversarial Networks (GANs) [95] are a category of generative DL models that utilize a generator-discriminator architecture to produce realistic data samples. GANs are composed of two neural networks: a generator, which creates synthetic data samples (e.g., RF signals, images) from random noise, and a discriminator, which attempts to distinguish between real and synthetic samples. As depicted in Figure 5, these networks are trained simultaneously in an adversarial setup, where the generator strives to generate increasingly convincing samples, and the discriminator works

to accurately differentiate between real and generated data. This adversarial training continues until the generator becomes proficient enough that the discriminator can no longer reliably differentiate between the two. GANs are particularly valuable for tasks such as data generation, augmentation, and anomaly detection. However, training GANs can be challenging due to issues like mode collapse and unstable convergence [96]. In RF sensing, GANs find applications in areas like radio map construction [24,97].

Diffusion Models (DMs) [98] are a type of generative models that use a two-phase process involving forward and reverse diffusion. In the forward phase, noise is gradually added to the data over multiple steps, transforming it into a noise distribution and obscuring the original structure. The reverse phase, handled by a neural network, learns to remove this noise step by step to reconstruct the original data. This training process enables the network to generate new samples from noise by reversing the diffusion. DMs are highly effective for creating high-quality images, often surpassing traditional methods like GANs in stability and output quality. Their probabilistic framework offers controlled data generation, making them appropriate for synthetic RF signal generation [25,99], or RF data augmentation [100].

Large Language Models (LLMs), such as GPT (Generative Pre-trained Transformer) [101] and LLaMA [102], are built upon transformer architecture [103]. As displayed in Figure 5, the input text is tokenized and then encoded with positional information to capture word order before being passed through multiple blocks comprising self-attention and feed-forward neural network layers. This structure concludes with a softmax output layer for predicting the next sequence. The transformer architecture enables parallelized training, significantly accelerating the training process, particularly when handling vast datasets. Moreover, the self-attention mechanism determines the contextual relevance of each word in relation to others, allowing the model to capture long-range dependencies and context, a capability that addresses the limitations of RNNs in maintaining long-term memory. During training, LLMs are exposed to extensive text corpora and are typically optimized with self-supervised objectives, such as masked language modeling and next-word prediction. These approaches result in models proficient at diverse natural language processing tasks such as text generation, translation, and summarization. Although limited research has explored the application of LLMs to RF sensing, their emerging capabilities, such as in-context learning, instruction following, and adaptability to varying conditions, suggest the potential for RF applications, especially in scenarios characterized by variability in devices and environments.

3.2. Deep Learning-Empowered Outdoor RF Sensing

Having explored a range of DL architectures commonly employed in RF sensing, including foundational models for feature extraction and advanced architectures for generating and interpreting complex data patterns, we now turn to recent studies that leverage these techniques specifically for outdoor RF sensing. These studies demonstrate how various DL approaches address outdoor challenges such as signal interference, multipath effects, and environmental variability, showcasing the practical applications and effectiveness of these models in real-world scenarios.

3.2.1. Generative Models

Applications of generative models in recent RF sensing research have focused on enhancing data reconstruction, augmentation, and synthesis to address challenges such as sparse datasets, missing signal information, and low signal-to-noise ratios (SNR). Leveraging their powerful ability to generate synthetic RF data, these models can fill in missing points in radio maps and create diverse scenarios for model training. For example, techniques like GANs and DMs are used to reconstruct and predict RF signals by learning from visual or prior RF data. Conditional GANs [97] support accurate localization and traffic monitoring, while DMs, such as RF-Diffusion [25], achieve high accuracy in reconstructing RF signals across various domains. Additionally, models like RFGen [99] synthesize realistic RF data using simulation-based approaches, which significantly enhance applications like gesture recognition

and posture estimation in previously unseen environments. By generating and expanding data, generative DL models improve the performance and generalization capabilities of RF sensing systems, enabling more accurate detection, classification, and localization tasks.

Moreover, generative models are also highly effective in multimodality tasks within RF sensing, where data from different sources, or modalities, are combined to improve learning and inference. In multimodality, models can use information from one type of data, such as RF signals, visual or audio, to enhance understanding of another type. This approach is particularly beneficial when one modality has sparse or incomplete data, as models can leverage information from another modality to fill in gaps and improve accuracy [104]. With recent advances in Multimodal Large Language Models (MLLMs), DL now enables the integration of multiple modalities, such as LiDAR, images, and RF signals, creating unified representations of surrounding environments and providing comprehensive datasets to train other models more effectively.

Radio map construction is a critical application in outdoor RF sensing, allowing for the creation of spatial maps that capture signal characteristics across different areas. This process is essential for tasks like localization and environmental monitoring, where accurately mapping RF signals enhances situational awareness. Generative models, particularly GANs, are well-suited for this task due to their architecture, which can learn complex data distributions and recover missing signal information. For example, CoSense [24] utilizes conditional GANs to reconstruct incomplete mmWave signals by learning patterns from visual data, addressing the challenge of weak reflectivity in mmWave signals and filling in gaps within the radio map. CoSense combines networking and sensing tasks through time-sharing and residual networks to dynamically predict heatmaps, which improves traffic monitoring and pedestrian safety at corners, even in adverse weather conditions. Using picocells, CoSense improves the detection and localization of pedestrians and vehicles. In terms of performance, CoSense acquires a median Intersection over Union (IoU) of 0.55 for pedestrians and 0.63 for vehicles. IoU is a metric used to evaluate the overlap between the predicted area (e.g., of an object) and the actual area, with values closer to 1 indicating higher accuracy. Therefore, these IoU scores demonstrate a reasonable level of accuracy for detecting and localizing pedestrians and vehicles while maintaining network throughput and reducing sensing overhead by 70%.

In real-time RF sensing scenarios, transmitter information, such as the location and power of base stations or mobile devices, is often unavailable due to practical limitations. Zhang et al. [97] develop a Cooperative Radio Map Estimation (CRME) approach for 6G networks, which estimates spatial radio signal distribution, or radio maps, without relying on transmitter information. The authors implement a conditional GAN framework, called GAN-CRME, which leverages distributed received signal strength (RSS) samples from mobile users. The GAN-CRME method achieves a root mean squared error (RMSE) of approximately 0.009 when using over 1000 RSS samples, outperforming the state-of-the-art RadioUNet model which scores 0.021 in RMSE.

Teganya and Romero [93] develop a DL approach using AE for accurate radio map estimation by filling the missing points in the radio map, focusing on power spectral density (PSD) estimation in wireless communication networks. The authors propose a deep completion autoencoder architecture that leverages spatial and frequency information. They trained the model using both synthetic and real-world data, applying techniques such as denoising and transfer learning to enhance the network's performance. The proposed method achieves a 30% reduction in root mean square error (RMSE) compared to existing techniques when using synthetic data, and up to 100% improvement in RMSE for real-world data in urban environments, demonstrating its superior performance in both scenarios.

The generative DL model also has the important ability to generate synthetic data, which can help address issues with sparse datasets when training models for downstream tasks. RFGen [99] enhances the generalization capabilities of mmWave sensing systems by synthesizing RF data using cross-model diffusion and ray tracing simulation. The authors integrate a physics-based ray tracing simulator with a DM to generate diverse 3D scenes and corresponding RF data. They utilize RF sensing prompts to create specific applications and environments, combining structured multipath noise and

environmental context through a path-based intermediate representation to simulate realistic RF signals. RFGen shows a significant improvement in unseen scenarios, increasing gesture recognition accuracy to above 90% across various orientations and reducing average posture estimation errors by up to 90% compared to baseline models.

Among significant works using DMs, Chi et al. [25] introduced RF-Diffusion, a versatile generative DL model for RF signals based on DMs. This approach leverages diffusion processes across time and frequency domains to reconstruct RF signals accurately. The model is trained using a complex-valued neural network to handle RF signal characteristics and incorporates attention-based modules for effective feature extraction. RF-Diffusion achieves an average Structural Similarity Index Measure (SSIM) of 0.81 for Wi-Fi signals and 0.75 for Frequency Modulated Continuous Wave (FMCW) signals, surpassing baseline models like Denoising Diffusion Probabilistic Model (DDPM), Deep Convolutional GAN (DCGAN), and Conditional Variational Autoencoder (CVAE) by up to 71.3%.

Due to the limitations of traditional augmentation methods, such as rotation and flipping, and the modest improvements achieved by advanced techniques like GANs in automatic modulation classification (AMC), Xu et al [100] introduce a Diffusion-based Radio Signal Augmentation (DiRSA) algorithm, enhancing dataset diversity and reducing overfitting for DL models. DiRSA reconstructs radio signals from noise using modulation category prompts to expand and diversify the training dataset, specifically targeting scenarios with limited data availability. DiRSA improves classification accuracy by up to 6% at SNR levels above 0 dB when applied with the LSTM model, compared to traditional methods like rotation and flipping.

Chang et al. [105] improve the accuracy of radio frequency fingerprint (RFF) recognition by developing a method that integrates prior information of wireless signals to combine the characteristics of transient and steady-state signals for enhanced classification, particularly under low SNR conditions. The proposed model segments incoming signals into transient and steady-state components and uses an autoencoder for denoising transient features. The model fuses these features to form a comprehensive RFF, leveraging high SNR prior information to enhance accuracy at 100% recognition accuracy on the LFM6 dataset and a 99.91% accuracy on the LFM15 dataset at 5 dB SNR. It also maintains 90.64% accuracy even under challenging conditions with -5 dB SNR, demonstrating its robustness and effectiveness in various environments.

3.2.2. Discriminative Models

Discriminative DL models in RF sensing are typically applied for tasks like signal detection, classification, and localization [7,13,28,106]. These models, such as CNNs, LSTMs, and advanced hybrid networks, excel in extracting and classifying features from RF signals to identify patterns or specific entities (e.g., drones, human activities, and vehicles) and manage RF spectrum efficiently. For instance, models like WRIST [107] and DeepFeat [108] use spectro-temporal and LTE-specific features, respectively, to detect and localize RF emissions or devices, achieving high accuracy levels. These applications extend to human and drone recognition tasks, where CNNs, LSTMs, and attention-based models process RF data like spectrograms and CSI for precise identification, even in dynamic or congested environments [24,86]. Discriminative models are also used in health monitoring systems like HealthDAR [109], where they identify vital signs and human activities through RF signals, offering non-intrusive monitoring solutions. Despite their effectiveness, these models often face challenges in maintaining performance under varying environmental conditions and depending on high-quality training data.

To improve real-time outdoor localization accuracy in dense urban environments where GNSS-based methods fail due to LoS issues, Yapar et al. introduced LocUNet [10], an end-to-end CNN for localization using path loss radio maps. The authors compare four different DL models: RadioUNet [110], fingerprint-based kNN [111], Adaptive KNN [112], and LocUNet, trained using path loss radio maps to improve accuracy despite signal variability and interference. LocUNet outperforms other

models with an average localization error of 5 meters, showing 11 and 14-meter improvements over kNN-based methods [111,112].

Nguyen et al. [107] present a wideband, real-time, spectro-temporal (WRIST) RF identification system aimed at addressing the critical challenge of spectrum scarcity and dynamic RF spectrum management. The authors employ a DL framework inspired by the YOLO [113] object detection model, leveraging transfer learning and an iterative training approach with synthesized and over-the-air RF data to develop an efficient identification system. WRIST supports the detection, classification, and precise localization of RF emissions in the 2.4 GHz industrial, scientific, and medical (ISM) band, achieving high performance with a class detection accuracy of over 99% and precision and recall values reaching 94% in controlled environments. However, challenges persist in maintaining the system's accuracy in highly congested, real-world scenarios. The authors suggest future work should focus on expanding the dataset and enhancing the system's capabilities to improve robustness and adaptability under diverse and dense spectrum conditions.

DeepFeat [108] is a deep-learning-based framework optimized for large-scale outdoor localization in LTE networks. The approach integrates a feature selection module utilizing Chi-squared and correlation techniques, effectively reducing computational load by 20.6% while enhancing accuracy through a refined selection of 12 LTE-specific features. By employing a deep feed-forward neural network, the system reaches a median localization accuracy of 13.179 m for a 6.27 km² area and 13.7 m for a 45 km² region. Despite these positive outcomes, maintaining consistent accuracy in varied environments remains challenging, with future efforts focused on improving model adaptability and performance in larger urban areas.

UAVs have become a popular research focus due to their extensive applications in areas like public safety, agriculture, and communications [114,115]. However, with this growth comes the need for advanced detection and classification techniques to ensure safe and secure UAV operation. Xue et al. [116] propose a DL approach for UAV identification using RF signals, specifically addressing challenges in scenarios involving nonstandard waveforms, such as unknown operating channels and environmental variations. To handle these, the system includes a carrier frequency offset (CFO) estimation and compensation method based on morphological filtering, which enhances signal alignment in varied conditions. The authors explore different signal representations, including in-phase and quadrature (IQ) samples, amplitude envelopes, and spectrograms, in combination with DL models like CNNs and LSTMs. IQ samples refer to the complex-valued representations of RF signals that capture the real (in-phase) and imaginary (quadrature) components, which are critical for preserving the signal's phase information. Their experiments demonstrate that a real-valued CNN using spectrogram inputs achieves optimal results, with 97% classification accuracy and efficient processing. However, further enhancements are necessary to improve the system's resilience to dynamic wireless conditions, particularly in terms of maintaining accuracy across diverse, real-world electromagnetic environments.

Another approach transforms RF signals into spectrograms, which are then processed using deep neural networks (DNNs) as proposed by Podder et al. [86]. The ResNet-50V2 model initially applied in noise-free indoor conditions, achieved an 85.39% accuracy. However, in outdoor environments at 50 m and 100 m distances, the accuracy decreased to 68.90% and 56.88%, respectively. To address this, the authors developed a CNN model optimized for outdoor settings, which improved classification accuracy to 78.12%. They further enhanced their system using a binary classification task, achieving a 95.08% accuracy on a mixed dataset of UAV and non-UAV images. Despite these achievements, challenges remain, especially in maintaining high accuracy at longer distances and under varied noise conditions, where different levels of additive white Gaussian noise (AWGN), distort signal quality by lowering the SNR.

An end-to-end DL model for detecting and identifying drones by using RF signals, addressing challenges posed by interference from other signals like Bluetooth and Wi-Fi operating in the same 2.4 GHz band by Alam et al. [5]. Their approach employs a multiscale feature extraction technique using CNNs to extract enriched features without manual intervention, reducing computational overhead. The

model achieves 97.53% accuracy for overall detection, with precision, sensitivity, and F1-score values reaching over 98% across varying SNR. However, challenges persist, particularly in maintaining high accuracy under low SNR conditions, where signal clarity is compromised by increased noise, and in complex, real-world environments, which introduce additional signal interference and environmental variability.

In other cases, human subjects are also utilized for various tasks, including activity recognition and presence detection. HealthDAR [109] is a contactless health monitoring system designed for vital sign monitoring, human activity recognition, and tracking. The system leverages a compact, low-energy radar integrated with a DL network to detect coughs and monitor vital signs like heart rate and breath rate. HealthDAR achieves a high precision in uncontrolled environments, showing a Pearson correlation coefficient of 0.99 for heart rate and 0.98 for breath rate when compared to ground truth data, indicating a strong similarity between HealthDAR's estimations and the actual measured values. The mean error for heart rate is 0.3 beats per minute with a standard deviation of 0.04, while for breath rate, it consistently falls within the expected range for adult respiratory rates. Despite its promising results, the system faces challenges in accurately recognizing activities for new subjects and under diverse scenarios, indicating the need for further enhancements to improve robustness and generalization.

Wang et al. [89] propose mmParse, a novel system for human parsing using mmWave radar point clouds, designed to overcome challenges like sparsity and specular reflection inherent in such data. The system employs a multi-task learning framework, combining human parsing with auxiliary tasks like pose estimation to enhance structural feature extraction. In particular, they employ LSTM as part of the feature extraction process for pose estimation. Evaluations demonstrate that mmParse achieves around 92% overall accuracy and an 84% mean IoU across various environments. Nonetheless, challenges remain, especially regarding signal deflection, which can lead to missing body parts in the data. To overcome these challenges, future research could explore data augmentation techniques or additional sensor fusion approaches to compensate for incomplete data and validate the system's performance in a wider range of real-world conditions.

To accelerate other research in the area of human sensing, Yang et al. [90] presented SenseFi, a comprehensive DL framework and benchmark designed for Wi-Fi-based human sensing. The system evaluates various DL models, such as CNNs, LSTMs, and transformers, using CSI data for tasks like human activity recognition, gesture recognition, and human identification. Extensive experiments demonstrate that shallow models, such as CNN-5, often outperform deeper architectures like ResNet in diverse Wi-Fi environments, achieving an accuracy of 98.11% on UT-HAR datasets. Despite the success, challenges include optimizing models for cross-domain adaptation and ensuring efficiency in real-time applications.

Human-vehicle recognition (HVR) has achieved significant interest due to its prospect of enabling non-intrusive, efficient detection of traffic participants, which is essential for enhancing intelligent transportation systems. Song et al. [117] introduced wireless-based lightweight attention deep learning (Wi-LADL), a lightweight DL model for HVR that leverages attention mechanisms in wireless sensing to improve feature discrimination. Wi-LADL uses RSS data processed with convolutional block attention modules (CBAM) to capture detailed features, achieving an impressive 98.8% average accuracy at a 2.4 GHz frequency with an antenna height of 0.8 meters. This model effectively distinguishes five categories: one-pedestrian, two-pedestrian, one-bicycle, two-bicycles, and one-car. Although Wi-LADL demonstrates high accuracy, challenges persist, particularly in adapting the model to varying antenna heights and frequencies.

3.2.3. Integrating Discriminative and Generative Models

Many recent works integrate generative and discriminative models to leverage their complementary strengths. Wang et al. [118] introduce UAV-CTNet, a hybrid DL network designed to improve UAV detection and identification in response to growing security concerns. UAVs often operate in

the 2.4 GHz ISM band, where their signals are difficult to detect reliably. UAV-CTNet integrates CNN and Transformer architectures: the Transformer captures global features of the RF signals, while the CNN focuses on extracting local features from minimum variance distortionless response (MVDR) spectral vectors. This approach achieves over 90% detection accuracy under various SNR values, outperforming conventional methods (FBREWT + CNNNet) [119].

Another significant work, Babel [120], aims to enhance multimodal sensing through a technique known as expandable modality alignment. Babel incorporates a pre-trained modality tower (BERT, STGCN, ResNet3D, ViT, and Point Transformer) to encode multiple sensing modalities (IMU, Video, Wi-Fi signals, mmWave signals, and LiDAR data) into a unified representation for downstream sensing tasks. The results demonstrate this framework improved human activity recognition accuracy by up to 22% compared to state-of-the-art methods, showing an average accuracy improvement of 12.02% across six modalities. The authors of the paper suggest that future work should focus on expanding Babel's capability to align additional sensing modalities involving inviting contributions from the broader community to further align additional modalities, as Babel is designed with a scalable architecture.

AirECG [121], a contactless Electrocardiogram monitoring system using mmWave sensing, employs a custom cross-domain DM that translates mmWave signals into ECG data through multiple denoising iterations and calibration guidance. AirECG uses a CNN-based patchification approach to process the multichannel mmWave data. The CNN encodes the data into patches (tokens) that contain cardiac features. These tokens serve as the input for the diffusion process, allowing the system to integrate and enhance cardiac information from multiple reflection points on the chest. The core of the denoising process is a hierarchical Transformer, which takes the CNN-encoded mmWave tokens and performs denoising in multiple iterations. To further enhance the fidelity of the generated electrocardiogram (ECG), a calibration guidance mechanism is integrated. This module uses historical ECG data (from reference devices) to guide the denoising steps. AirECG achieved a Pearson correlation coefficient (PCC) of 0.955 and 0.860 for normal heartbeat abnormal beats, respectively, showing 15.0% to 21.1% improvements over existing methods.

Despite the differences in characteristics and preferences between discriminative and generative models, they can share the same tasks in RF sensing. For example, Nie et al. [55] employs four DL models: CNN-LSTM, Swin Transformer, ConvNext [122], and Vision TF [123], to enhance human activity recognition using LoRa wireless RF signals. The models process LoRa spectrograms generated using short-time Fourier transform (STFT) and differential signal processing (DSP) for better feature extraction. Each model is trained and evaluated across tasks i.e., activity classification, identity recognition, room identification, and presence detection. The result shows that ConvNext achieved the highest performance, with a 96.7% accuracy in activity classification and 97.9% in identity recognition. Vision TF excelled in presence detection with 98.5% accuracy. CNN-LSTM and Swin Transformer showed moderate performance, highlighting ConvNext's superiority in spatial feature extraction and Vision TF's effectiveness in global context understanding, however, the experiment was undertaken in controlled indoor environments.

While generative models have shown remarkable performance in tasks such as classification and regression, discriminative models often excel in specific scenarios where their efficiency and precision are advantageous. For example, in cases where the dataset is limited, discriminative models can perform better due to their ability to directly optimize class boundaries without the need for extensive data generation, which generative models typically rely on. Additionally, when the system requires minimal latency, such as in real-time RF sensing applications, discriminative models are often preferred since they generally involve less computational overhead and can provide faster outputs compared to generative models, which can be more time-consuming. The lower computational cost of simple DL models like MLPs or CNNs also makes discriminative models more suitable when resource efficiency is critical. Additionally, when robust data processing pipelines are already established, discriminative models can achieve high accuracy without necessitating the synthetic data generation or enhancement capabilities offered by generative models, thereby streamlining the system's overall design.

While DL models have various strengths, their flexibility allows them to be used for a wide range of tasks in RF sensing that are not limited to a specific type of model. Some DL systems can handle the entire RF sensing process, from signal preprocessing to outcome prediction, based solely on the input and desired output [24]. It is essential to understand the strengths and weaknesses of both generative and discriminative models to optimize these systems, enabling the development of more effective and adaptable RF sensing solutions. Integrating both types of models can enhance RF sensing capabilities, especially in challenging environments [100,118]. For instance, combining generative models for data augmentation or synthetic signal generation with discriminative models for accurate classification can improve performance in situations with low signal-to-noise ratios or variable signal conditions, ensuring robust and reliable operation as demonstrated in Figure 6.

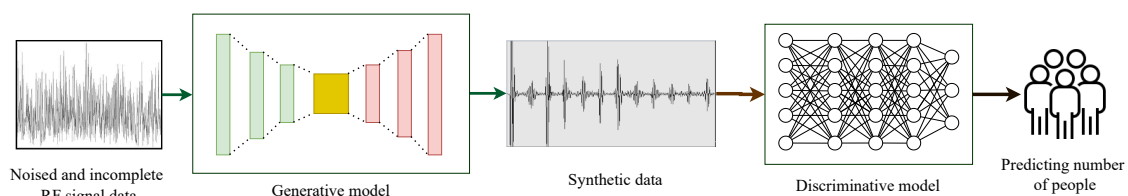


Figure 6. A proposal for an RF sensing system that operates under low SNR conditions, predicting the number of people by integrating both generative and discriminative models to enhance accuracy.

4. Challenges and Future Directions

Despite the optimistic progress highlighted in the previous section, several challenges remain in applying DL to outdoor RF sensing. These challenges arise from the unique features of RF data and the diverse conditions of outdoor applications, ranging from data scarcity and processing demands to the complexities of integrating multimodal data. Handling these issues is essential for the continued advancement of RF sensing technology and its applicability across domains like traffic management, environmental monitoring, and urban infrastructure. In this section, we investigate key challenges in RF sensing, including data limitations, the gap between synthetic and real-world data, and the need for integrated sensing and communication systems, as well as emerging approaches such as federated learning that hold promise for overcoming these obstacles.

4.1. The Scarcity of Training Data

Acquiring labeled data for applications like human tracking, activity recognition, or vehicle monitoring poses substantial privacy and ethical concerns, restricting the availability of large-scale datasets for these tasks [120,124,125]. Even when sufficient training data is available, manual annotation of specific events is often required, which is labor-intensive, time-consuming, and costly [126]. This issue is particularly severe for RF data labeling, as unlike visual data that can be easily reviewed offline through recordings, RF data is not intuitively interpretable by humans without specialized tools [28]. One potential solution is to combine RF data with more easily labeled modalities, such as vision or audio, to build more robust multi-modal models. This approach allows DL models to leverage information from more interpretable data types, reducing dependence on large volumes of purely RF-labeled data. However, collecting synchronized multi-modal data in outdoor environments remains technically challenging due to the complexity of integrating and aligning data from different sensors, especially when deployed over large areas.

4.2. The Gap Between Synthetic and Real-World Data

A principal challenge in training with synthetic data is bridging the gap between controlled simulations and the variability of real-world scenarios [10,110]. While simulated environments allow precise control, real-world data introduces variations in material properties, signal conditions, and dynamic outdoor factors that are difficult to replicate accurately. This difference makes it challenging to verify that models trained on synthetic data function effectively in real-world scenarios. Elements

such as changing weather, diverse terrains, and sensor noise add further complexity to generating synthetic data that captures the nuances of actual environments. Addressing this gap remains critical for deploying reliable models in practical settings. Current methods mitigate this issue by balancing synthetic and real-world data in model training. Additionally, cross-modality data generation offers promise; for example, Li et al. [127] proposed SBRF, a model to generate RF signals from video data by integrating ray tracing with electromagnetic computation. This physics-based approach overcomes some limitations of purely data-driven and model-driven techniques, which may lack precision or be costly and labor-intensive. However, challenges persist in adapting this method to complex, real-world conditions.

4.3. The Data Preprocessing Effort

Generative models are highly effective at processing raw signal data, particularly in outdoor environments where RF signals are often disrupted by ambient factors. However, generative AI models face challenges with latency and computational efficiency, making them less suitable for low-latency, resource-constrained scenarios such as autonomous vehicles or real-time surveillance [124]. For example, Zeng et al. [128] introduce a radio anomaly detection framework using denoising diffusion probabilistic models that address issues like unstable training and poor performance with low SNR signals. Despite its strong performance, the framework demands significant computational resources to achieve high anomaly detection accuracy. To address this, applying compression techniques or model distillation methods, as suggested by Menghani [129], could reduce computational demands while maintaining the model's desired performance, offering a practical solution for efficiency-sensitive applications. Another notable work by Liu et al. [130] introduced a predictive communication protocol utilizing a convolutional LSTM network based on historical channel data. This approach eliminates the need for explicit channel tracking, thereby significantly reducing signaling overhead.

4.4. Multimodal RF Sensing

Integrating multiple data modalities into a single RF sensing system can significantly reduce the cost and complexity of using separate systems for various urban applications while enhancing their connectivity. For instance, in traffic management, RF sensing can complement vision systems and sensor networks to enable vehicle detection, monitor pedestrian movements, and control smart traffic lights dynamically, leading to optimized traffic flow and improved safety [24]. In environmental monitoring, RF sensing combined with air quality sensors offers real-time tracking of pollution levels and environmental conditions [9]. However, achieving a functional multimodal system presents notable challenges, especially in integrating and synchronizing data from sensors with different sampling rates and formats, which requires sophisticated alignment methods. Furthermore, implementing real-time processing in large-scale urban settings necessitates robust data fusion algorithms and advanced infrastructure capable of managing high data volumes consistently and efficiently.

4.5. Integrated Sensing and Communication (ISAC)

RF sensing can be implemented within existing communication infrastructure, effectively utilizing resources while preserving communication capacity. This approach, known as Integrated Sensing and Communication (ISAC), has appeared as a compelling research area. ISAC refers to a design approach and set of enabling technologies that combine sensing and communication capabilities, aiming to optimize the utilization of wireless resources while providing mutual advantages [8]. This integration enables advancements in fields such as mobile crowd sensing, channel knowledge mapping, passive sensing networks, vehicular communications, satellite imaging, and broadcasting [8]. Additionally, ISAC leverages recent progress in machine learning and DL. Specifically, the functions of machine learning in ISAC systems are underlined by Demirhan and Alkhateeb [131], spanning joint sensing and communication (JSC), sensing-aided communication, and communication-aided sensing. These roles include optimizing waveform design, predicting beam patterns, enhancing system security,

and improving network-level operations, all contributing to significant performance gains based on real-world data.

4.6. Federated Learning

With the rapid advancement of RF sensing systems, scaling these systems to cover larger areas and track more objects introduces a significant challenge: the need for scalable, privacy-preserving solutions. These solutions must efficiently process vast amounts of data generated by multiple entities while protecting sensitive information collected in public spaces. One promising approach is federated learning [132], a machine learning framework that distributes model training across multiple devices or decentralized data sources. In this setup, each device retains its local data and trains the model locally, sharing only model updates with a central server instead of transmitting raw data. Federated learning provides several benefits for RF sensing, including enhanced data privacy through local data retention, reduced network latency by avoiding the transmission of large datasets, and improved training efficiency. By leveraging the computational power and diverse datasets across a network of IoT devices, federated learning can also accelerate model convergence and enhance overall learning accuracy [133].

5. Conclusion

In this comprehensive survey, we have examined recent applications of DL architectures and techniques in outdoor RF sensing. Our analysis indicates that the rapid advancement of generative DL models has significantly enhanced RF sensing systems, particularly in outdoor environments where RF signal quality and quantity are often compromised due to environmental challenges. The findings highlight the pivotal role of wireless technologies in determining RF sensing performance, highlighting the need for continuous development in this area. Moreover, the diversity of settings and devices presents challenges in establishing a universal framework for RF sensing tasks, which typically require high-quality, large-scale datasets for effective training. Looking ahead, we anticipate that improvements in infrastructure, leading to increased computational power and the ability to generate synthetic data, will facilitate the integration of more advanced DL techniques, such as MLLMs, thereby significantly enhancing the performance of RF sensing systems.

References

1. Zhang, J.; Xi, R.; He, Y.; Sun, Y.; Guo, X.; Wang, W.; Na, X.; Liu, Y.; Shi, Z.; Gu, T. A Survey of mmWave-Based Human Sensing: Technology, Platforms and Applications. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 2052–2087. doi:10.1109/COMST.2023.3298300.
2. Chen, Z.; Zheng, T.; Luo, J. Octopus: A Practical and Versatile Wideband MIMO Sensing Platform. 27th Annual International Conference on Mobile Computing and Networking (MobiCom '21); , 2021; pp. 601–614. doi:10.1145/3447993.3483267.
3. Lubna, L.; Hameed, H.; Ansari, S.; Zahid, A.; Sharif, A.; Abbas, H.T.; Alqahtani, F.; Mufti, N.; Ullah, S.; Imran, M.A.; Abbasi, Q.H. Radio frequency sensing and its innovative applications in diverse sectors: A comprehensive study. *Front. Commun. Netw.* **2022**, *3*. doi:10.3389/frcmn.2022.1010228.
4. Kong, H.; Huang, C.; Yu, J.; Shen, X. A Survey of mmWave Radar-Based Sensing in Autonomous Vehicles, Smart Homes and Industry. *IEEE Commun. Surv. Tutor.* **2024**. Early Access, doi:10.1109/COMST.2024.3409556.
5. Alam, S.S.; Chakma, A.; Rahman, M.H.; Bin Mofidul, R.; Alam, M.M.; Utama, I.B.K.Y.; Jang, Y.M. RF-Enabled Deep-Learning-Assisted Drone Detection and Identification: An End-to-End Approach. *Sensors* **2023**, *23*, 4202. doi:10.3390/s23094202.
6. Ahmad, I.; Ullah, A.; Choi, W. WiFi-Based Human Sensing With Deep Learning: Recent Advances, Challenges, and Opportunities. *IEEE Open J. Commun. Soc.* **2024**, *5*, 3595–3623. doi:10.1109/OJCOMS.2024.3411529.
7. Jagannath, A.; Jagannath, J.; Kumar, P.S.P.V. A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges. *Comput. Netw.* **2022**, *219*, 109455. doi:10.1016/j.comnet.2022.109455.

8. Liu, F.; Cui, Y.; Masouros, C.; Xu, J.; Han, T.X.; Eldar, Y.C.; Buzzi, S. Integrated Sensing and Communications: Toward Dual-Functional Wireless Networks for 6G and Beyond. *IEEE J. Select. Areas Commun.* **2022**, *40*, 1728–1767. doi:10.1109/JSAC.2022.3156632.
9. Van Truong, T.; Nayyar, A.; Masud, M. A novel air quality monitoring and improvement system based on wireless sensor and actuator networks using LoRa communication. *PeerJ Comput. Sci.* **2021**, *7*, e711. doi:10.7717/peerj-cs.711.
10. Yapar, C.; Levie, R.; Kutyniok, G.; Caire, G. Real-Time Outdoor Localization Using Radio Maps: A Deep Learning Approach. *IEEE Trans. Wirel. Commun.* **2023**, *22*, 9703–9717. doi:10.1109/TWC.2023.3273202.
11. Yi Lim, N.C.; Yong, L.; Su, H.T.; Yu Hao Chai, A.; Vithanawasam, C.K.; Then, Y.L.; Siang Tay, F. Review of Temperature and Humidity Impacts on RF Signals. 13th International UNIMAS Engineering Conference (EnCon 2020); , 2020; pp. 1–8. doi:10.1109/EnCon51501.2020.9299327.
12. Haenggi, M.; Ganti, R.K.; others. Interference in large wireless networks. *Found. Trends Netw.* **2009**, *3*, 127–248. doi:10.1561/1300000015.
13. Zheng, T.; Chen, Z.; Ding, S.; Luo, J. Enhancing RF Sensing with Deep Learning: A Layered Approach. *IEEE Commun. Mag.* **2021**, *59*, 70–76. doi:10.1109/MCOM.001.2000288.
14. Wang, X.; Wang, X.; Mao, S. RF Sensing in the Internet of Things: A General Deep Learning Framework. *IEEE Commun. Mag.*, *56*, 62–67. doi:10.1109/MCOM.2018.1701277.
15. Wang, X.; Zhao, Y.; Pourpanah, F. Recent advances in deep learning. *Int. J. Mach. Learn. Cybern.* **2020**, *11*, 747–750.
16. Bayoudh, K.; Knani, R.; Hamdaoui, F.; Mtibaa, A. A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets. *Vis. Comput.* **2022**, *38*, 2939–2970.
17. Ni, J.; Young, T.; Pandelea, V.; Xue, F.; Cambria, E. Recent advances in deep learning based dialogue systems: A systematic survey. *Artif. Intell. Rev.* **2023**, *56*, 3055–3155.
18. Iman, M.; Arabnia, H.R.; Rasheed, K. A review of deep transfer learning and recent advancements. *Technol.* **2023**, *11*, 40.
19. Choudhary, K.; DeCost, B.; Chen, C.; Jain, A.; Tavazza, F.; Cohn, R.; Park, C.W.; Choudhary, A.; Agrawal, A.; Billinge, S.J.; others. Recent advances and applications of deep learning methods in materials science. *npj Comput. Mater.* **2022**, *8*, 59.
20. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444.
21. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
22. Elbir, A.M. DeepMUSIC: Multiple signal classification via deep learning. *IEEE Sens. Lett.* **2020**, *4*, 1–4. doi:10.1109/LSSENS.2020.2980384.
23. Conn, M.A.; Josyula, D. Radio Frequency Classification and Anomaly Detection Using Convolutional Neural Networks. 2019 IEEE Radar Conference (RadarConf); , 2019; pp. 1–6. doi:10.1109/RADAR.2019.8835709.
24. Regmi, H.; Sur, S. CoSense: Deep Learning Augmented Sensing for Coexistence with Networking in Millimeter-Wave Picocells. *ACM Trans. Internet Things* **2024**, *5*, 17:1–17:35. doi:10.1145/3670415.
25. Chi, G.; Yang, Z.; Wu, C.; Xu, J.; Gao, Y.; Liu, Y.; Han, T.X. RF-Diffusion: Radio Signal Generation via Time-Frequency Diffusion. 30th Annual International Conference on Mobile Computing and Networking (MobiCom '24); , 2024; pp. 77–92. doi:10.1145/3636534.3649348.
26. Liu, C.; Liu, X.; Li, S.; Yuan, W.; Ng, D.W.K. Deep CLSTM for Predictive Beamforming in Integrated Sensing and Communication-Enabled Vehicular Networks. *Journal of Communications and Information Networks* **2022**, *7*, 269–277.
27. Liu, C.; Liu, X.; Wei, Z.; Ng, D.W.K.; Schober, R. Scalable Predictive Beamforming for IRS-Assisted Multi-User Communications: A Deep Learning Approach. *arXiv preprint arXiv:2211.12644* **2022**.
28. Nirmal, I.; Khamis, A.; Hassan, M.; Hu, W.; Zhu, X. Deep Learning for Radio-Based Human Sensing: Recent Advances and Future Directions. *IEEE Commun. Surv. Tutorials* **2021**, *23*, 995–1019. doi:10.1109/COMST.2021.3058333.
29. Burghal, D.; Ravi, A.T.; Rao, V.; Alghafis, A.A.; Molisch, A.F. A Comprehensive Survey of Machine Learning Based Localization with Wireless Signals. *arXiv preprint arXiv:2012.11171* **2020**. doi:10.48550/arXiv.2012.11171.
30. Amjad, B.; Ahmed, Q.Z.; Lazaridis, P.I.; Hafeez, M.; Khan, F.A.; Zaharis, Z.D. Radio SLAM: A Review on Radio-Based Simultaneous Localization and Mapping. *IEEE Access* **2023**, *11*, 9260–9278. doi:10.1109/ACCESS.2023.3237330.

31. Soumya, A.; Krishna Mohan, C.; Cenkeramaddi, L.R. Recent Advances in mmWave-Radar-Based Sensing, Its Applications, and Machine Learning Techniques: A Review. *Sensors* **2023**, *23*, 8901. doi:10.3390/s23218901.
32. Zafari, F.; Gkelias, A.; Leung, K.K. A Survey of Indoor Localization Systems and Technologies. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2568–2599. doi:10.1109/COMST.2019.2911558.
33. Yassin, A.; Nasser, Y.; Awad, M.; Al-Dubai, A.; Liu, R.; Yuen, C.; Raulefs, R.; Aboutanios, E. Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 1327–1346. doi:10.1109/COMST.2016.2632427.
34. Dogan, D.; Dalveren, Y.; Kara, A. A Mini-Review on Radio Frequency Fingerprinting Localization in Outdoor Environments: Recent Advances and Challenges. 14th International Conference on Communications (COMM);, 2022; pp. 1–5. doi:10.1109/COMM54429.2022.9817189.
35. Budalal, A.A.; Islam, M.R. Path loss models for outdoor environment—with a focus on rain attenuation impact on short-range millimeter-wave links. *e-Prime - Adv. Electr. Eng. Electron. Energy* **2023**, *3*, 100106. doi:https://doi.org/10.1016/j.prime.2023.100106.
36. Liu, Y.; Liu, X.; Mu, X.; Hou, T.; Xu, J.; Di Renzo, M.; Al-Dhahir, N. Reconfigurable Intelligent Surfaces: Principles and Opportunities. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 1546–1577. doi:10.1109/COMST.2021.3077737.
37. Trichopoulos, G.C.; Theofanopoulos, P.; Kashyap, B.; Shekhawat, A.; Modi, A.; Osman, T.; Kumar, S.; Sengar, A.; Chang, A.; Alkhateeb, A. Design and Evaluation of Reconfigurable Intelligent Surfaces in Real-World Environment. *IEEE Open J. Commun. Soc.* **2022**, *3*, 462–474. doi:10.1109/OJCOMS.2022.3158310.
38. Hu, J.; Zhang, H.; Di, B.; Li, L.; Bian, K.; Song, L.; Li, Y.; Han, Z.; Poor, H.V. Reconfigurable Intelligent Surface Based RF Sensing: Design, Optimization, and Implementation. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2700–2716. doi:10.1109/JSAC.2020.3000805.
39. Alexandropoulos, G.C.; Crozzoli, M.; Phan-Huy, D.T.; Katsanos, K.D.; Wymeersch, H.; Popovski, P.; Ratajczak, P.; Bénédic, Y.; Hamon, M.H.; Gonzalez, S.H.; D’Errico, R.; Strinati, E.C. Smart Wireless Environments Enabled by RISs: Deployment Scenarios and Two Key Challenges. *arXiv preprint arXiv:2203.13478* **2022**. doi:10.48550/arXiv.2203.13478.
40. Huang, Y.; Chen, Z.; Wen, C.; Li, J.; Xia, X.G.; Hong, W. An Efficient Radio Frequency Interference Mitigation Algorithm in Real Synthetic Aperture Radar Data. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–12. doi:10.1109/TGRS.2022.3155068.
41. Ameloot, T.; Van Torre, P.; Rogier, H. Variable Link Performance Due to Weather Effects in a Long-Range, Low-Power LoRa Sensor Network. *Sensors* **2021**, *21*. doi:10.3390/s21093128.
42. Hau, F.; Baumgärtner, F.; Vossiek, M. The Degradation of Automotive Radar Sensor Signals Caused by Vehicle Vibrations and Other Nonlinear Movements. *Sensors* **2020**, *20*. doi:10.3390/s20216195.
43. Iannizzotto, G.; Milici, M.; Nucita, A.; Lo Bello, L. A Perspective on Passive Human Sensing with Bluetooth. *Sensors* **2022**, *22*. doi:10.3390/s22093523.
44. Tabassum, M.; Zen, K. Performance Evaluation of ZigBee in Indoor and Outdoor Environment. 9th International Conference on IT in Asia (CITA);, 2015; pp. 1–7. doi:10.1109/CITA.2015.7349825.
45. Semtech Semiconductor, IoT Systems and Cloud Connectivity | Semtech.
46. Augustin, A.; Yi, J.; Clausen, T.; Townsley, W.M. A Study of LoRa: Long Range I& Low Power Networks for the Internet of Things. *Sensors* **2016**, *16*, 1466. doi:10.3390/s16091466.
47. Chu, N.H.; Nguyen, D.N.; Hoang, D.T.; Pham, Q.V.; Phan, K.T.; Hwang, W.J.; Dutkiewicz, E. AI-Enabled mm-Waveform Configuration for Autonomous Vehicles With Integrated Communication and Sensing. *IEEE Internet Things J.* **2023**, *10*, 16727–16743. doi:10.1109/JIOT.2023.3270420.
48. Chen, C.; Chen, X.; Das, D.; Akhmetov, D.; Cordeiro, C. Overview and Performance Evaluation of Wi-Fi 7. *IEEE Commun. Stand. Mag.* **2022**, *6*, 12–18. doi:10.1109/MCOMSTD.0001.2100082.
49. Ahson, S.A.; Ilyas, M. *RFID Handbook: Applications, Technology, Security, and Privacy*; CRC press: Boca, FL, USA, 2017. ISBN 978-1-4200-5499-6.
50. Barbieri, L.; Brambilla, M.; Trabattoni, A.; Mervic, S.; Nicoli, M. UWB Localization in a Smart Factory: Augmentation Methods and Experimental Assessment. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–18. doi:10.1109/TIM.2021.3074403.
51. Hillger, P.; Grzyb, J.; Jain, R.; Pfeiffer, U.R. Terahertz Imaging and Sensing Applications With Silicon-Based Technologies. *IEEE Trans. Terahertz Sci. Technol.* **2019**, *9*, 1–19. doi:10.1109/TTHZ.2018.2884852.

52. Li, Y.; Chi, Z.; Liu, X.; Zhu, T. Passive-ZigBee: Enabling ZigBee Communication in IoT Networks with 1000X+ Less Power Consumption. 16th ACM Conference on Embedded Networked Sensor Systems (SenSys '18); , 2018; pp. 159–171. doi:10.1145/3274783.3274846.
53. Demeslay, C.; Rostaing, P.; Gautier, R. Theoretical Performance of LoRa System in Multipath and Interference Channels. *IEEE Internet Things J.* **2022**, *9*, 6830–6843. doi:10.1109/JIOT.2021.3108380.
54. Haxhibeqiri, J.; Van den Abeele, F.; Moerman, I.; Hoebeke, J. LoRa Scalability: A Simulation Model Based on Interference Measurements. *Sensors* **2017**, *17*, 1193. doi:10.3390/s17061193.
55. Nie, M.; Zou, L.; Cui, H.; Zhou, X.; Wan, Y. Enhancing Human Activity Recognition with LoRa Wireless RF Signal Preprocessing and Deep Learning. *Electron.* **2024**, *13*, 264. doi:10.3390/electronics13020264.
56. Islam, K.Z.; Murray, D.; Diepeveen, D.; Jones, M.G.K.; Sohel, F. LoRa-based outdoor localization and tracking using unsupervised symbolization. *Internet Things* **2024**, *25*, 101016. doi:10.1016/j.iot.2023.101016.
57. Wu, D.; Liebeherr, J. A Low-Cost Low-Power LoRa Mesh Network for Large-Scale Environmental Sensing. *IEEE Internet Things J.* **2023**, *10*, 16700–16714. doi:10.1109/JIOT.2023.3270237.
58. Hao, Z.; Yan, H.; Dang, X.; Ma, Z.; Jin, P.; Ke, W. Millimeter-Wave Radar Localization Using Indoor Multipath Effect. *Sensors* **2022**, *22*, 5671. doi:10.3390/s22155671.
59. Kwon, G.; Liu, Z.; Conti, A.; Park, H.; Win, M.Z. Integrated Localization and Communication for Efficient Millimeter Wave Networks. *IEEE J. Sel. Areas Commun.* **2023**, *41*, 3925–3941. doi:10.1109/JSAC.2023.3322760.
60. Feng, Y.; Xie, Y.; Ganesan, D.; Xiong, J. LTE-based Pervasive Sensing Across Indoor and Outdoor. 19th ACM Conference on Embedded Networked Sensor Systems; , 2021; SenSys '21, p. 138–151. doi:10.1145/3485730.3485943.
61. Sardar, S.; Mishra, A.K.; Khan, M.Z.A. Vehicle detection and classification using LTE-CommSense. *IET Radar Sonar Navig.* **2019**, *13*, 850–857. doi:10.1049/iet-rsn.2018.5231.
62. Sonny, A.; Rai, P.K.; Kumar, A.; Khan, M.Z.A. Deep learning-based smart parking solution using channel state information in LTE-based cellular networks. International Conference on COMMunication Systems & NETWORKS (COMSNETS); , 2020; pp. 642–645. doi:10.1109/COMSNETS48256.2020.9027447.
63. Jabbar, A.; Abdullah, F.Y. Long term evolution (LTE) scheduling algorithms in wireless sensor networks (WSN). *Int. J. Comput. Appl.* **2015**, *121*. doi:10.5120/21575-4626.
64. Gu, Y.; Chen, J.; He, K.; Wu, C.; Zhao, Z.; Du, R. WiFiLeaks: Exposing Stationary Human Presence Through a Wall With Commodity Mobile Devices. *IEEE Trans. Mob. Comput.* **2024**, *23*, 6997–7011. doi:10.1109/TMC.2023.3328349.
65. Ma, Y.; Zhou, G.; Wang, S. WiFi Sensing with Channel State Information. *ACM Comput. Surv.* **2019**, *52*, 1–36. doi:10.1145/3310194.
66. Landaluce, H.; Arjona, L.; Perallos, A.; Falcone, F.; Angulo, I.; Muralter, F. A Review of IoT Sensing Applications and Challenges Using RFID and Wireless Sensor Networks. *Sensors* **2020**, *20*, 2495. doi:10.3390/s20092495.
67. Shen, E.; Yang, W.; Wang, X.; Kang, B.; Mao, S. TagSense: Robust Wheat Moisture and Temperature Sensing Using RFID. *IEEE J. Radio Freq. Identif.* **2024**, *8*, 76–87. doi:10.1109/JRFID.2024.3389868.
68. Le Breton, M.; Baillet, L.; Larose, E.; Rey, E.; Benech, P.; Jongmans, D.; Guyoton, F. Outdoor UHF RFID: Phase Stabilization for Real-world Applications. *IEEE J. Radio Freq. Identif.* **2017**, *1*, 279–290. doi:10.1109/JRFID.2017.2786745.
69. Zhang, D.; Yang, L.T.; Chen, M.; Zhao, S.; Guo, M.; Zhang, Y. Real-time locating systems using active RFID for Internet of Things. *IEEE Syst. J.* **2014**, *10*, 1226–1235. doi:10.1109/JSYST.2014.2346625.
70. Florentin, I. Discussion on UWB Technology and Its Applicability in Different Fields. *J. Mil. Technol.* **2020**, *4*. doi:10.32754/jmt.2021.1.05.
71. Queralta, J.P.; Martínez Almansa, C.; Schiano, F.; Floreano, D.; Westerlund, T. UWB-based System for UAV Localization in GNSS-Denied Environments: Characterization and Dataset. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); , 2020; pp. 4521–4528. doi:10.1109/IROS45743.2020.9341042.
72. Gabela, J.; Retscher, G.; Goel, S.; Perakis, H.; Masiero, A.; Toth, C.; Gikas, V.; Kealy, A.; Koppányi, Z.; Błaszczak-Bak, W.; others. Experimental Evaluation of a UWB-Based Cooperative Positioning System for Pedestrians in GNSS-Denied Environment. *Sensors* **2019**, *19*, 5274. doi:10.3390/s19235274.
73. Yang, J.; Dong, B.; Wang, J. VULoc: Accurate UWB Localization for Countless Targets without Synchronization. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2022**, *6*. doi:10.1145/3550286.

74. Siegel, P. Terahertz technology. *IEEE Trans. Microw. Theory Tech.* **2002**, *50*, 910–928. doi:10.1109/22.989974.
75. Bogue, R. Sensing with terahertz radiation: a review of recent progress. *Sens. Rev.* **2018**, *38*, 216–222. doi:10.1108/SR-10-2017-0221.
76. Jansen, C.; Wietzke, S.; Peters, O.; Scheller, M.; Vieweg, N.; Salhi, M.; Krumbholz, N.; Jördens, C.; Hochrein, T.; Koch, M. Terahertz imaging: applications and perspectives. *Appl. Opt.* **2010**, *49*, E48–E57. doi:10.1364/AO.49.000E48.
77. Ren, A.; Zahid, A.; Fan, D.; Yang, X.; Imran, M.A.; Alomainy, A.; Abbasi, Q.H. State-of-the-Art in Terahertz Sensing for Food and Water Security – A Comprehensive Review. *Trends Food Sci. Technol.* **2019**, *85*, 241–251. doi:10.1016/j.tifs.2019.01.019.
78. Pawar, A.Y.; Sonawane, D.D.; Erande, K.B.; Derle, D.V. Terahertz technology and its applications. *Drug Invent. Today* **2013**, *5*, 157–163. doi:10.1016/j.dit.2013.03.009.
79. Naftaly, M.; Vieweg, N.; Deninger, A. Industrial applications of terahertz sensing: State of play. *Sensors* **2019**, *19*, 4203.
80. Tomczak, J.M. Why Deep Generative Modeling? In *Deep Generative Modeling*; Springer Nature: Cham, Switzerland, 2021; pp. 1–12. doi:10.1007/978-3-030-93158-2_1.
81. Wang, L.; Zhang, C.; Zhao, Q.; Zou, H.; Lasaulce, S.; Valenzise, G.; He, Z.; Debbah, M. Generative AI for RF Sensing in IoT Systems. *arXiv preprint arXiv:2407.07506* **2024**.
82. Zhou, R.; Tang, M.; Gong, Z.; Hao, M. FreeTrack: Device-free human tracking with deep neural networks and particle filtering. *IEEE Syst. J.* **2019**, *14*, 2990–3000.
83. Wu, X.; Chu, Z.; Yang, P.; Xiang, C.; Zheng, X.; Huang, W. TW-See: Human activity recognition through the wall with commodity Wi-Fi devices. *IEEE Trans. Veh. Technol.* **2018**, *68*, 306–319.
84. Zhao, Y.; Wang, G.; Tang, C.; Luo, C.; Zeng, W.; Zha, Z.J. A Battle of Network Structures: An Empirical Study of CNN, Transformer, and MLP. *arXiv preprint arXiv:2108.13002* **2021**.
85. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 1–74.
86. Podder, P.; Zawodniok, M.; Madria, S. Deep Learning for UAV Detection and Classification via Radio Frequency Signal Analysis. 25th IEEE International Conference on Mobile Data Management (MDM); , 2024; pp. 165–174. doi:10.1109/MDM61037.2024.00040.
87. Shewalkar, A.; Nyavanandi, D.; Ludwig, S.A. Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU. *J. Artif. Intell. Soft Comput. Res.* **2019**, *9*, 235–245.
88. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D* **2020**, *404*, 132306.
89. Wang, S.; Cao, D.; Liu, R.; Jiang, W.; Yao, T.; Lu, C.X. Human Parsing with Joint Learning for Dynamic mmWave Radar Point Cloud. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2023**, *7*, 34:1–34:22. doi:10.1145/3580779.
90. Yang, J.; Chen, X.; Zou, H.; Lu, C.X.; Wang, D.; Sun, S.; Xie, L. SenseFi: A library and benchmark on deep-learning-empowered WiFi human sensing. *Patterns* **2023**, *4*. doi:10.1016/j.patter.2023.100703.
91. Kingma, D.P.; Welling, M. Auto-Encoding Variational Bayes. *arXiv preprint arXiv:1312.6114* **2013**.
92. Bank, D.; Koenigstein, N.; Giryas, R. Autoencoders. *Mach. Learn. Data Sci. Handb.* **2023**, pp. 353–374.
93. Teganya, Y.; Romero, D. Deep Completion Autoencoders for Radio Map Estimation. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 1710–1724. doi:10.1109/TWC.2021.3106154.
94. Almazrouei, E.; Gianini, G.; Almoosa, N.; Damiani, E. A Deep Learning Approach to Radio Signal Denoising. 2019 IEEE Wireless Communications and Networking Conference Workshops (WCNCW); , 2019; pp. 1–8. doi:10.1109/WCNCW.2019.8902856.
95. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144.
96. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65.
97. Zhang, Z.; Zhu, G.; Chen, J.; Cui, S. Fast and Accurate Cooperative Radio Map Estimation Enabled by GAN. *arXiv preprint arXiv:2402.02729* **2024**. doi:10.48550/arXiv.2402.02729.
98. Ho, J.; Jain, A.; Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 6840–6851.

99. Chen, X.; Zhang, X. RF Genesis: Zero-Shot Generalization of mmWave Sensing through Simulation-Based Data Synthesis and Generative Diffusion Models. 21st ACM Conference on Embedded Networked Sensor Systems (SenSys '23); , 2024; pp. 28–42. doi:10.1145/3625687.3625798.
100. Xu, Y.; Huang, L.; Zhang, L.; Qian, L.; Yang, X. Diffusion-Based Radio Signal Augmentation for Automatic Modulation Classification. *Electron.* **2024**, *13*, 2063. doi:10.3390/electronics13112063.
101. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving Language Understanding by Generative Pre-Training. *OpenAI* **2018**.
102. Touvron, H.; Lavril, T.; Izacard, G.; Martinet, X.; Lachaux, M.A.; Lacroix, T.; Rozière, B.; Goyal, N.; Hambro, E.; Azhar, F.; Rodriguez, A.; Joulin, A.; Grave, E.; Lample, G. LLaMA: Open and Efficient Foundation Language Models. *arXiv preprint arXiv:2302.13971* **2023**.
103. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*; , 2017; pp. 5998–6008.
104. Liu, G.; Van Huynh, N.; Du, H.; Hoang, D.T.; Niyato, D.; Zhu, K.; Kang, J.; Xiong, Z.; Jamalipour, A.; Kim, D.I. Generative AI for Unmanned Vehicle Swarms: Challenges, Applications and Opportunities. *arXiv preprint arXiv:2402.18062* **2024**.
105. Chang, J.; Zhou, Z.; Mi, S.; Zhang, Y. Radio frequency fingerprint recognition method based on prior information. *Comput. Electr. Eng.* **2024**, *120*, 109684. doi:10.1016/j.compeleceng.2024.109684.
106. Liu, C.; Wei, Z.; Ng, D.W.K.; Yuan, J.; Liang, Y.C. Deep Transfer Learning for Signal Detection in Ambient Backscatter Communications. *IEEE Trans. Wireless Commun.* **2021**, *20*, 1624–1638.
107. Nguyen, H.N.; Vomvas, M.; Vo-Huu, T.D.; Noubir, G. WRIST: Wideband, Real-Time, Spectro-Temporal RF Identification System Using Deep Learning. *IEEE Trans. Mob. Comput.* **2024**, *23*, 1550–1567. doi:10.1109/TMC.2023.3240971.
108. Mohamed, A.; Tharwat, M.; Magdy, M.; Abubakr, T.; Nasr, O.; Youssef, M. DeepFeat: Robust Large-Scale Multi-Features Outdoor Localization in LTE Networks Using Deep Learning. *IEEE Access* **2022**, *10*, 3400–3414. doi:10.1109/ACCESS.2022.3140292.
109. Li, A.; Bodanese, E.; Poslad, S.; Chen, P.; Wang, J.; Fan, Y.; Hou, T. A Contactless Health Monitoring System for Vital Signs Monitoring, Human Activity Recognition, and Tracking. *IEEE Internet Things J.* **2024**, *11*, 29275–29286. doi:10.1109/JIOT.2023.3336232.
110. Levie, R.; Yapar, C.; Kutyniok, G.; Caire, G. RadioUNet: Fast Radio Map Estimation With Convolutional Neural Networks. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 4001–4015. doi:10.1109/TWC.2021.3054977.
111. Bahl, P.; Padmanabhan, V.N. RADAR: An In-Building RF-Based User Location and Tracking System. 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000); , 2000; pp. 775–784. doi:10.1109/INFCOM.2000.832252.
112. Oh, J.; Kim, J. AdaptiveK-nearest neighbour algorithm for WiFi fingerprint positioning. *ICT Express* **2018**, *4*, 91–94. doi:https://doi.org/10.1016/j.ict.2018.04.004.
113. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; , 2016; pp. 779–788.
114. Yao, Y.; Lv, K.; Huang, S.; Li, X.; Xiang, W. UAV trajectory and energy efficiency optimization in RIS-assisted multi-user air-to-ground communications networks. *Drones* **2023**, *7*, 272.
115. Lukito, W.D.; Xiang, W.; Lai, P.; Cheng, P.; Liu, C.; Yu, K.; Zhu, X. Integrated STAR-RIS and UAV for Satellite IoT Communications: An Energy-Efficient Approach. *IEEE Internet of Things J.* **2024**. Early Access, doi:10.1109/JIOT.2024.10702554.
116. Xue, C.; Li, T.; Li, Y.; Ruan, Y.; Zhang, R.; Dobre, O.A. Radio-Frequency Identification for Drones With Nonstandard Waveforms Using Deep Learning. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–13. doi:10.1109/TIM.2023.3306822.
117. Song, M.; Lou, L.; Chen, X.; Zhao, X.; Hong, Y.; Zhang, S.; He, W. Wi-LADL: A Wireless-Based Lightweight Attention Deep Learning Method for Human–Vehicle Recognition. *IEEE Sens. J.* **2023**, *23*, 2803–2814. doi:10.1109/JSEN.2022.3232243.
118. Wang, Q.; Yang, P.; Yan, X.; Wu, H.C.; He, L. Radio Frequency-Based UAV Sensing Using Novel Hybrid Lightweight Learning Network. *IEEE Sens. J.* **2024**, *24*, 4841–4850. doi:10.1109/JSEN.2023.3346209.

119. Bremnes, K.; Moen, R.; Yeduri, S.R.; Yakkati, R.R.; Cenkeramaddi, L.R. Classification of UAVs utilizing fixed boundary empirical wavelet sub-bands of RF fingerprints and deep convolutional neural network. *IEEE Sens. J.* **2022**, *22*, 21248–21256.
120. Dai, S.; Jiang, S.; Yang, Y.; Cao, T.; Li, M.; Banerjee, S.; Qiu, L. Advancing Multi-Modal Sensing Through Expandable Modality Alignment. *arXiv preprint arXiv:2407.17777* **2024**.
121. Zhao, L.; Lyu, R.; Lei, H.; Lin, Q.; Zhou, A.; Ma, H.; Wang, J.; Meng, X.; Shao, C.; Tang, Y.; Chi, G.; Yang, Z. AirECG: Contactless Electrocardiogram for Cardiac Disease Monitoring via mmWave Sensing and Cross-domain Diffusion Model. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2024**, *8*, 144:1–144:27. doi:10.1145/3678550.
122. Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A ConvNet for the 2020s. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; , 2022; pp. 11976–11986. doi:10.1109/CVPR52688.2022.01167.
123. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; Uszkoreit, J.; Houlsby, N. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *9th International Conference on Learning Representations (ICLR)*; , 2021.
124. Lee, W.; Park, J. LLM-Empowered Resource Allocation in Wireless Communications Systems. *arXiv preprint arXiv:2408.02944* **2024**.
125. Khachatrian, H.; Mkrtchyan, R.; Raptis, T.P. Outdoor Environment Reconstruction with Deep Learning on Radio Propagation Paths. *arXiv preprint arXiv:2402.17336* **2024**. doi:10.48550/arXiv.2402.17336.
126. Chen, L.; Zheng, L.; Xia, D.; Sun, D.; Liu, W. STL-Detector: Detecting City-Wide Ride-Sharing Cars via Self-Taught Learning. *IEEE Internet Things J.* **2022**, *9*, 2346–2360. doi:10.1109/JIOT.2021.3094530.
127. Li, J.; Zhang, D.; Wu, Z.; Yu, C.; Li, Y.; Chen, Q.; Hu, Y.; Sun, Q.; Chen, Y. SBRF: A Fine-Grained Radar Signal Generator for Human Sensing. *IEEE Trans. Mob. Comput.* **2024**, pp. 1–17. Early Access, doi:10.1109/TMC.2024.3427406.
128. Zeng, J.; Liu, X.; Li, Z. Radio Anomaly Detection Based on Improved Denoising Diffusion Probabilistic Models. *IEEE Commun. Lett.* **2023**, *27*, 1979–1983. doi:10.1109/LCOMM.2023.3290813.
129. Menghani, G. Efficient deep learning: A survey on making deep learning models smaller, faster, and better. *ACM Comput. Surv.* **2023**, *55*, 1–37.
130. Liu, C.; Yuan, W.; Li, S.; Liu, X.; Li, H.; Ng, D.W.K.; Li, Y. Learning-based Predictive Beamforming for Integrated Sensing and Communication in Vehicular Networks. *IEEE J. Sel. Areas Commun.* **2022**, *40*, 2317–2334.
131. Demirhan, U.; Alkhateeb, A. Integrated Sensing and Communication for 6G: Ten Key Machine Learning Roles. *IEEE Commun. Mag.* **2023**, *61*, 113–119. doi:10.1109/MCOM.006.2200480.
132. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Aguera y Arcas, B. Communication-Efficient Learning of Deep Networks from Decentralized Data. *20th International Conference on Artificial Intelligence and Statistics (AISTATS)*; , 2017; pp. 1273–1282.
133. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A.; Li, J.; Vincent Poor, H. Federated Learning for Internet of Things: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 1622–1658. doi:10.1109/COMST.2021.3075439.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.