

# Wi-Fi Sensing Techniques for Human Activity Recognition: Brief Survey, Potential Challenges, and Research Directions

FUCHENG MIAO, AI Research Center, Nanjing Great Information Technology Co. Ltd, Nanjing, China

YOUXIANG HUANG, College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China

ZHIYI LU, AI Research Center, Nanjing Great Information Technology Co. Ltd, Nanjing, China

TOMOAKI OHTSUKI, Keio University, Minato-ku, Japan

GUAN GUI, College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China

HIKMET SARI, Nanjing University of Posts and Telecommunications, Nanjing, China

---

Recent advancements in wireless communication technologies have made Wi-Fi signals indispensable in both personal and professional settings. The utilization of these signals for Human Activity Recognition (HAR) has emerged as a cutting-edge technology. By leveraging the fluctuations in Wi-Fi signals for HAR, this approach offers enhanced privacy compared to traditional visual surveillance methods. The essence of this technique lies in detecting subtle changes when Wi-Fi signals interact with the human body, which are then captured and interpreted by advanced algorithms. This article initially provides an overview of the key methodologies in HAR and the evolution of non-contact sensing, introducing sensor-based recognition, computer vision, and Wi-Fi signal based approaches, respectively. It then explores tools for Wi-Fi-based HAR signal collection and lists several high-quality datasets. Subsequently, the article reviews various sensing tasks enabled by Wi-Fi signal recognition, highlighting the application of deep learning networks in Wi-Fi signal detection. Experimental results are then presented that assess the capabilities of different networks. The findings indicate significant variability in the generalization capacities of neural networks and notable differences in test accuracy for various motion analyses.

CCS Concepts: • **Networks** → **Wireless access networks**; • **Hardware** → *Wireless devices*; • **Human-centered computing** → *Ubiquitous and mobile computing systems and tools*; • **Computing methodologies** → **Activity recognition and understanding**; *Learning in probabilistic graphical models*; *Supervised learning by classification*;

---

This work was supported in part by the Natural Science Foundation of China under grants 62471247 and 62401281, in part by the Key Project of the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province under grant 22KJA510002, and in part by the Japan Science and Technology Agency (JST) Adopting Sustainable Partnerships for Innovative Research Ecosystem (ASPIRE) program under grant JPMJAP2326.

Authors' Contact Information: Fucheng Miao, AI Research Center, Nanjing Great Information Technology Co. Ltd., Nanjing, China; email: fuchengmiao@gmail.com; Youxiang Huang, College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, China; e-mail: 1322018705@njupt.edu.cn; Zhiyi Lu, AI Research Center, Nanjing Great Information Technology Co. Ltd., Nanjing, China; e-mail: luzhiyi6298@126.com; Tomoaki Ohtsuki, Keio University, Minato-ku, Tokyo, Japan; e-mail: ohtsuki@keio.jp; Guan Gui (Corresponding author), College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China; e-mail: guiguan@njupt.edu.cn; Hikmet Sari, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, China; e-mail: hsari@ieee.org.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://permissions.acm.org).

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 0360-0300/2025/01-ART107

<https://doi.org/10.1145/3705893>

Additional Key Words and Phrases: Non-contact sensing, Wi-Fi signals, intelligent sensing, human activity recognition, deep learning

#### ACM Reference Format:

Fucheng Miao, Youxiang Huang, Zhiyi Lu, Tomoaki Ohtsuki, Guan Gui, and Hikmet Sari. 2025. Wi-Fi Sensing Techniques for Human Activity Recognition: Brief Survey, Potential Challenges, and Research Directions. *ACM Comput. Surv.* 57, 5, Article 107 (January 2025), 30 pages. <https://doi.org/10.1145/3705893>

---

## 1 Introduction

In the wake of rapid advancements in wireless communication, the paradigm of information acquisition, storage, processing, and exchange has undergone a revolutionary transformation in our daily lives. Data and information, now predominantly encoded and disseminated in digital formats, offer unprecedented immediacy in access and distribution [1]. Central to this technological evolution is Wi-Fi, transcending its conventional role as a mere internet facilitator. Wi-Fi has emerged as a pivotal tool in advanced sensing and signal recognition applications [2, 3]. This article delves into the cutting-edge domain of Wi-Fi signal based non-contact sensing, spotlighting its application in **Human Activity Recognition (HAR)** [4].

HAR is a meaningful research direction. With emerging challenges such as severe population aging, the development of HAR has been increasingly driven by the needs in healthcare, patient monitoring, and safety protection for solitary elderly individuals [5]. Additionally, in fields like sports motion analysis, smart homes, and home security, HAR presents broad application prospects [6]. Existing HAR technologies can be primarily categorized into three aspects based on sensors, computer vision, and Wi-Fi signals.

As shown in Figure 1, sensor-based HAR employs diverse sensors to ascertain physiological data like body temperature, pulse rate, and electrocardiograms, facilitating the recognition of human states and activities. This approach leveraged concurrent multi-parameter measurements, providing an integrated dataset for enhanced predictive accuracy [7–17]. However, it confronts challenges including complex device wearability, potential inaccuracies from improper usage, and difficulties for certain users of self-equipment. Additionally, the prohibitive cost and limited convenience of sensor-based systems remain significant impediments [18–31].

In contrast, computer vision based HAR offers an intuitive methodology, capturing human activities through camera systems followed by computational analysis. Despite its straightforwardness and ease of use, this method is susceptible to varying environmental conditions like lighting and obstacle obstruction [32]. It also necessitates extensive labeled datasets for training, and the resource-intensive nature of deep learning based computer vision systems poses a barrier to lightweight application deployment [33].

Emerging as a novel frontier, Wi-Fi signal based HAR holds promising potential. With Wi-Fi technology evolving from its initial 802.11 standard to the latest 802.11ax (Wi-Fi 6) and the anticipated 802.11be (Wi-Fi 7) in 2024, and a surge in user adoption, Wi-Fi is increasingly integral in everyday environments, lending itself to innovative sensing applications. This non-contact sensing paradigm, aligning with current technological trends, seeks to infer the state of objects or individuals through signal analysis, eschewing direct physical interaction [34]. Key techniques in Wi-Fi-based recognition involve exploiting RSSI and CSI metrics. RSSI, indicative of signal strength, is represented in negative decibels (dBm) and is prone to environmental variances and multi-path distortions, limiting its reliability [35–37]. Conversely, CSI provides a comprehensive portrayal of the wireless channel's behavior, encompassing aspects like gain, phase shifts, and signal-to-noise ratio, offering a multi-dimensional dataset ideal for sophisticated Wi-Fi-based HAR tasks [38, 39].

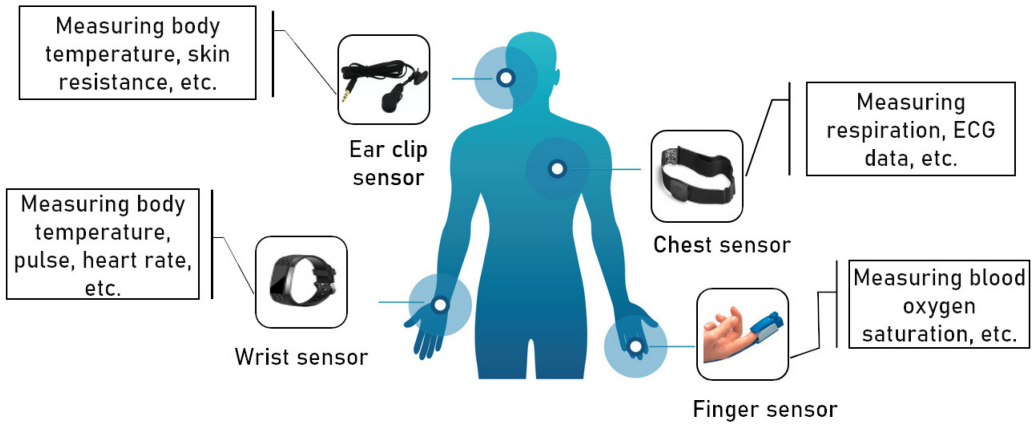


Fig. 1. An example illustration of sensor-based HAR methods.

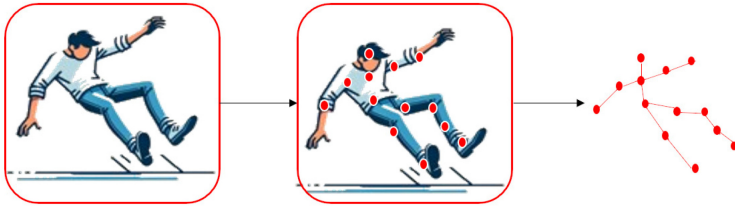


Fig. 2. Computer vision based HAR.

Wi-Fi-based non-contact sensing represents one of the most promising and valuable research directions in HAR. It holds advantages over other recognition methods in various aspects, and in the information age, tasks involving Wi-Fi-based non-contact sensing are set to increase [40]. Consequently, extensive research is required for further advancements in this field.

This article will comprehensively discuss the emergence of non-contact sensing, system models and problem descriptions of Wi-Fi-based HAR methods, and the deep learning models used in Wi-Fi-based sensing, addressing generalization issues.

## 2 The Rise of Non-Contact Sensing

Historically, HAR predominantly relied on sensor- and camera-based methods. Due to early technological constraints, researchers placed focus on enhancing sensor performance for improved monitoring [41]. The launch of Microsoft's Kinect sensor in 2010 marked a shift toward vision-based recognition, leveraging Kinect's ability to accurately infer skeletal information from depth images [42], thereby gaining traction in research applications [43–49].

Vision is a fundamental human interface for environment interaction and understanding, leading to the initial focus on vision-based methods in non-contact sensing [50, 51]. In computer vision, advancements in object detection and related networks have facilitated HAR through techniques like target and keypoint detection [52, 53]. This article also examines various image modalities used in computer vision, detailing their specific advantages and disadvantages [54–58]. Figure 2 presents an HAR detection method based on computer vision.

Table 1 reveals that despite the specific advantages of different image modalities for visual perception in various environments, they all share inherent limitations due to reliance on visual

Table 1. Brief Survey of Different Image Modalities

| Modality                   | Advantages                         | Disadvantages                        |
|----------------------------|------------------------------------|--------------------------------------|
| Infrared                   | Good for night and low light       | Limited use in normal light          |
| Depth                      | Provides depth data for 3D imaging | Needs extra sensors like TOF cameras |
| Magnetic resonance         | Excellent for medical imaging      | High cost, complex operation         |
| RGB (Red, Green, and Blue) | Versatile and widely used          | Limited depth and material detection |
| Panoramic                  | Ideal for wide-area VR             | High computational costs             |

imagery. This reliance can lead to issues such as obstructions, environmental variability, and privacy concerns.

The origins of non-contact HAR trace back to the edge detection and contour analysis of the 1970s to 1980s, later transitioning to computer vision technologies in the 1990s to 2000s [59]. Initial advancements were constrained by computational limitations and camera quality [60]. The introduction of depth cameras and infrared sensors in the early 2010s marked a significant leap in precision for human posture recognition [61]. However, in private domains like smart homes and elderly fall detection, traditional vision and sensor-based methods grapple with privacy and occlusion challenges [62–79].

The evolution of wireless communication has ushered in Wi-Fi-based systems for human activity detection, circumventing visual sensing by analyzing signal variations [80]. The emergence of CSI signal utilization in 2013 and subsequent integration of deep learning networks, including **Gated Recurrent Units (GRUs)** and **Long Short-Term Memory (LSTM)** models, have substantially enhanced the accuracy and complexity of HAR [81–83].

### 3 System Model and Problem Description

#### 3.1 System Model

In Wi-Fi environments, human activities modify signal reflections and diffractions, thus changing the signal's propagation path. Analyzing these changes in the distribution of Wi-Fi signal paths allows for the recognition of different human behaviors. A human posture recognition system based on a CSI signal includes key components such as wireless signal transmission and reception, CSI signal capture, data processing, and learning algorithms [84].

Figure 3 provides a simplified depiction of CSI signal collection for indoor human posture recognition. A router emits signals into the environment, which, upon encountering obstacles or human bodies, reflect and attenuate. The receiver then captures and records these reflections as CSI data. CSI data include details on the signal's amplitude and phase [85]. We can express the following formula:

$$H(t) = \sum_{l=0}^{L-1} a_l(t) e^{-j2\pi f \frac{d_l(t)}{c}}. \quad (1)$$

Here,  $L$  and  $d_l(t)$  respectively denote the number of paths for the  $L^{th}$  path and the path length. Meanwhile,  $f$  and  $c$  represent the carrier frequency and the speed of light.  $a_l(t)$  is a complex variable indicating phase and amplitude attenuation information.

In multi-path propagation, in addition to the dynamic transmission paths that change due to the movement of the sensing target, there are also static transmission paths that remain unaffected by this motion. The amplitude of these static paths does not vary with time, whereas the amplitude of the dynamic paths changes due to the reflection of the signal. Ultimately, the CSI matrix can be simplified into the following formula:

$$H(t) = H_s + H_d(t) + n(t) = H_s + \alpha e^{-j2\pi f \frac{d(t)}{c}} + n(t), \quad (2)$$

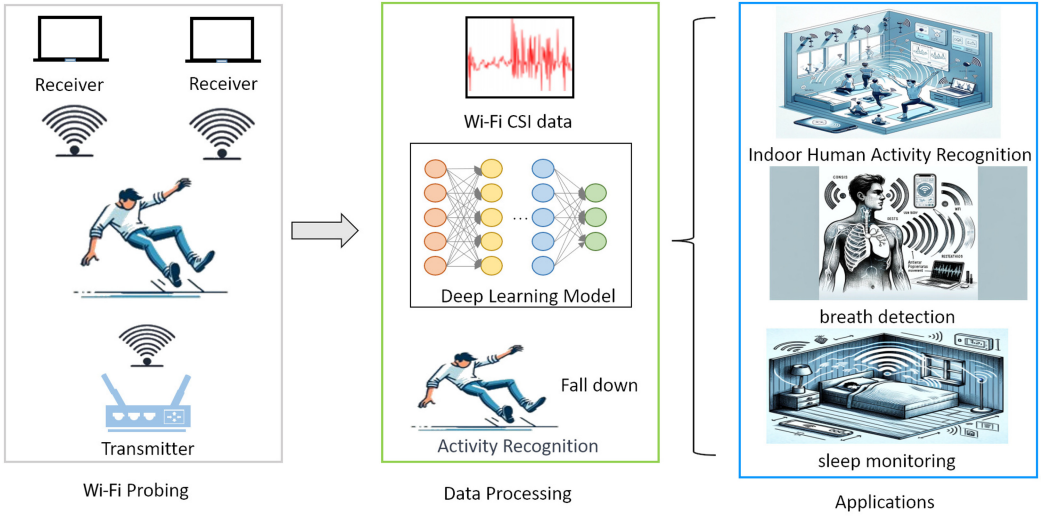


Fig. 3. Wi-Fi-based human posture detection system and its applications.

where  $H_s$  represents the static component of the static path, whose amplitude does not vary with the time index  $t$ , and  $H_d(t)$  represents the dynamic component of the dynamic path. The term  $\alpha$  denotes a complex coefficient representing both amplitude attenuation and the initial phase.  $n(t)$  is noise, which is typically approximated as additive white Gaussian noise.

From this formula, we observe that the three important factors affecting the CSI matrix are the static path component, the dynamic path component, and environmental noise. The static path component is independent of time and has little effect on the CSI. Changes in the reflected signal due to the movement of the sensing target cause variations in the dynamic path component. The stronger the reflected signal, the more pronounced the dynamic path component, leading to better CSI performance. Conversely, as noise increases, the perception performance of CSI worsens. To sum up, we can improve CSI perception either by adjusting the CSI path in the scene to increase the intensity of the reflected signal or by reducing environmental noise.

### 3.2 CSI Signal Collection Tools and Datasets

CSI signal collection tools are crucial in the field of wireless communication, used to capture subtle variations in wireless signal propagation [86]. These tools are generally utilized for Wi-Fi signal research, with applications including but not limited to wireless network optimization, indoor positioning, and human motion recognition. Next we discuss some common CSI signal collection tools. The Intel 5300 **Network Interface Card (NIC)** is a highly popular wireless network interface that provides detailed channel state information, suitable for professional research. The Intel 5300 NIC is designed for the 802.11n (Wi-Fi 4) standard. In 802.11n, up to 30 subcarriers can be used for CSI information collection. The Atheros CSI Tool is a software tool based on Atheros chips, used for collecting fine-grained wireless channel information under Linux systems. Theoretically, Atheros chips supporting 802.11n can provide up to 30 subcarriers for CSI information. However, the actual number of available subcarriers might vary depending on the specific device model and configuration. The Nexmon CSI Tool is an open source project that enables users to collect CSI data by modifying Broadcom wireless chip firmware. The number of carriers for Broadcom chips also depends on the supported wireless standard. For the 802.11n standard, it might offer up to 56 subcarriers, similar to the Intel 5300 NIC. The ESP32 CSI Tool is a

Table 2. Some Wi-Fi Collection Tools and Their Parameter Configurations

| Device Configuration  | Intel 5300 NIC  | Atheros CSI Tool | Nexmon CSI Tool  | ESP32 CSI Tool   |
|-----------------------|-----------------|------------------|------------------|------------------|
| Number of Subcarriers | 30              | 56 or 114        | 256              | 64               |
| Number of Antennas    | 3               | 3                | 1                | 1                |
| Size                  | > 30 cm × 20 cm | > 30 cm × 20 cm  | > 15 cm × 7.5 cm | > 5 cm × 3 cm    |
| Weight                | > 1 Kg          | > 1 Kg           | > 100 g          | > 10 g           |
| Deployment Scale      | Small scale     | Small scale      | Small scale      | Relatively large |
| Integration           | Low             | Low              | High             | High             |
| Memory                | 8 GB+           | 8 GB+            | 1 GB ~ 4 GB      | 500 KB ~ 4 MB    |
| Bandwidth             | 20 MHz          | 20/40 MHz        | 80 MHz           | 20/40 MHz        |
| Hardware Device       | Intel 5300 NIC  | Atheros NIC      | Raspberry Pi     | ESP32            |

data collection tool for the low-cost ESP32 chip, suitable for IoT projects. ESP32 is primarily used for 802.11b/g/n standard Wi-Fi networks. In 802.11n mode, ESP32 might provide a fewer number of subcarriers, as it is mainly designed for low-power and low-cost IoT applications. The specific number of subcarriers might need to be referred to in ESP32's technical specifications.

As shown in Table 2, these are some additional parameters of the Wi-Fi signal collection tools [87–104].

Ding et al. [105] mentioned that data was collected in various indoor environments using the Intel 5300 network card, with a focus on four typical activities at a sampling rate of 200 frames per second. The sampling duration for these activities was 5 seconds, resulting in a sample size of 650 frames. The study results indicated that activity recognition accuracy was higher when in a standing posture.

In contrast, Zhang et al. [106] employed two datasets: a gesture dataset encompassing 276 gestures across different environments, each with 20 samples, and a body activity dataset. The latter included 10 types of activities, each with 50 samples in lab and conference room environments, using CSI tools based on the 802.11 protocol.

Alazrai et al. [107] used a dataset containing 12 types of human-to-human interaction behaviors, with a total of 40 pairs of test subjects having performed 12 types of interactive actions, each conducted 10 times, resulting in a total of 4,800 samples. The dataset recorded two types of signals: the RSSI and CSI values. The dataset link was provided in their work [107]: <https://data.mendeley.com/datasets/3dhn4xnjxw>.

The work by Yang et al. [108] involved three subjects performing data collection in three different scenarios, including activities such as bending, making a phone call, clapping, drinking water, sitting down, squatting, walking, and waving hands, totaling approximately 600 data files. They mentioned that to simulate activities under real conditions, objects on the table were randomly moved during the data collection process, and the laboratory door was always kept closed.

Zhang et al. [109] collected 15 different gestures performed by 16 users at 15 locations in three environments.

The dataset discussed in the work by Yang et al. [110] included two main parts: CSI-based HAR and Human Identity Recognition. The collection utilized two TP-Link N750 access points as the transmitter and receiver. These devices were equipped with three pairs of antennas and operated at 5 GHz with a 40-MHz bandwidth, extracting complete CSI data of 114 subcarriers at each timestamp. The cloud server in their work was simulated through a powerful local server equipped with an NVIDIA RTX 2080Ti GPU. The data collection sampling rate was 500 Hz, meaning  $3 \times 114 \times 500$

Table 3. Summary of Select Wi-Fi Datasets for Reference

| Name                       | Collection Tool | Types of Labels | Total Samples | Availability |
|----------------------------|-----------------|-----------------|---------------|--------------|
| D. Xue et al. [105]        | Intel 5300 NIC  | 4               | 200~400       | Unknown      |
| Z. Yong et al. [106]       | Intel 5300 NIC  | 10              | 500           | Unknown      |
| R. Alazrai, et al. [107]   | Intel 5300 NIC  | 12              | 4,800         | Yes          |
| J. Yang et al. [108]       | Intel 5300 NIC  | 3               | 600           | Yes          |
| Y. Zhang et al. [109]      | Intel 5300 NIC  | 22              | 43,000        | Yes          |
| J. Yang et al. [110]       | Antheros NIC    | 6               | 1,200         | Yes          |
| D. Wang et al. [111]       | Antheros NIC    | 14              | 840           | Yes          |
| F. Glenn et al. [112]      | Nexmon          | 11              | 1,100         | Unknown      |
| F. Moshiri et al. [113]    | Nexmon          | 7               | 420           | Yes          |
| S. Xu et al. [114]         | ESP32           | 7               | 560           | Yes          |
| Ma.Yongsen, et al. [115]   | Intel 5300 NIC  | 5               | 14,555        | Yes          |
| Tan et al. [116]           | Intel 5300 NIC  | 6               | 300           | Unknown      |
| Damodaran et al. [117]     | Intel 5300 NIC  | 5               | 2,500         | Unknown      |
| B. A. Alsaify et al. [118] | Intel 5300 NIC  | 6               | 3,000         | Yes          |

data samples per second. For the HAR part, six human activities were collected, including running, walking, falling, boxing, turning arms, and cleaning the floor, performed by 20 subjects (13 males and 7 females), with each subject performing each activity 20 times, resulting in 400 samples per category. For the Human Identity Recognition part, subjects were asked to walk through the **Line of Sight (LoS)** of the Wi-Fi access point from different directions, recording the CSI profiles of 15 subjects, each with 60 samples. Forbes et al. [112] used a dataset capturing CSI data with Raspberry Pi 4, covering a series of everyday activities such as standing, sitting down, getting into bed, and cooking, more closely reflecting reality, with data collection conducted in a home environment.

The dataset used by Moshiri et al. [113] covered seven different daily activities: sitting down, standing, lying down, running, walking, falling, and bending. The dataset is available at <https://github.com/parisafm/CSI-HAR-Dataset> (accessed on October 27, 2021).

Xu et al. [114] collected CSI data contained 64 subcarriers, providing amplitude and phase information. After cleaning the raw data to remove ineffective non-activity data and labeling specific activities, the dataset includes activities such as waving hands, clapping, walking, lying down, sitting, falling, and standing up after falling, with a total of four test subjects, each activity with 20 samples, totaling 560 samples. This dataset can be used to test the generalization ability of neural networks in recognizing different subjects. The dataset is available at <https://github.com/NJUPTSivan/Cross-person-HAR>.

Damodaran and Schäfer [117] used a dataset that includes five types of human activities: sitting (SIT), standing (STAND), running (RUN), walking (WALK), and empty (EMPTY). The Linux 802.11n CSI Tool was used for data collection, which supports injection and monitor modes for collecting activity samples.

Table 3 presents a list of datasets and their respective collection tools, labels, and so on.

It is important to note that some papers marked as having public datasets did not provide links or download information for the datasets, which does not mean that the datasets are not publicly available. Sometimes researchers publish dataset information on other platforms (e.g., personal or laboratory websites). If you are interested in a specific dataset, it is recommended to contact the authors of the work or visit the official website of the journal for more information.

### 3.3 Converting CSI Data into Algorithmic Inputs

In algorithmic processing of CSI data, relevance and representativeness are crucial. CSI encompasses amplitude, phase, and time series, necessitating careful data selection based on algorithmic objectives.

Amplitude data is often preferred for its stability and minimal environmental impact, facilitating acquisition and processing. It also efficiently indicates key environmental changes. Phase differences between multiple antennas, another approach, yield nuanced movement insights, especially valuable in precision-dependent applications like gesture recognition [119–126].

Preprocessing of CSI datasets is imperative, encompassing filtering, denoising, and normalization. This step, including the removal of non-activity data, ensures data quality and suitability for algorithmic analysis. The transformation process involves feature extraction, leveraging statistical and time-frequency analysis techniques, format conversion, and data labeling, to render CSI data optimally utilizable for specific algorithmic needs.

### 3.4 Research Objectives and Assumptions

Assumptions refer to preliminary premises on which this research is based, generally derived from prior studies and theoretical foundations. The assumptions for non-contact sensing based on Wi-Fi signals can be considered from the following aspects:

- *Wireless signal propagation principles*: In indoor environments, wireless signals undergo multiple path effects like reflection and refraction, and the Doppler effect due to relative motion between the signal source and receiver causes frequency variations.
- *Signal processing techniques*: Techniques to extract features related to human posture from CSI signals, like amplitude, phase, and Doppler frequency shifts. Current signal processing technologies are relatively advanced, providing numerous methods for handling CSI signals.
- *Machine learning and recognition methods*: During training, the dataset used must be effective and the training algorithms must be excellent. For example, the accuracy of current deep learning algorithms like LSTM and GRU is relatively high.

Constraint conditions refer to factors that affect implementation during system design or project execution. The constraints for non-contact sensing based on Wi-Fi signals can be considered from the following aspects:

- *Hardware limitations*: The configuration of antennas and the number of subcarriers provided by signal capture tools can limit signal transmission. Additionally, compatibility issues might arise with iterative updates of Wi-Fi devices.
- *Environmental factors*: In home environments, numerous obstacles inevitably affect signal capture and transmission. Therefore, it is important to consider environmental factors during experiments.
- *Data constraints*: Currently, comprehensive and extensive open source datasets for human posture recognition are scarce. Although datasets can be obtained from Table 3, suitable datasets are still needed according to specific experimental requirements.

### 3.5 Existing Wi-Fi Sensing Methods

Currently, Wi-Fi sensing tasks can be categorized into two scales: large-scale sensing and small-scale sensing. Common small-scale sensing tasks include breathing detection, gesture recognition, and sleep monitoring. Large-scale sensing tasks encompass indoor positioning, human posture estimation, and anomaly detection or intrusion detection [127, 128]. The HAR tasks are categorized as shown in Figure 4.

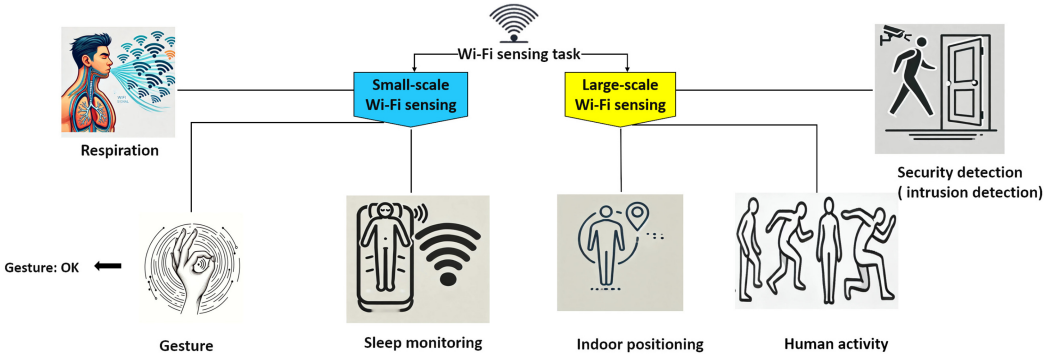


Fig. 4. HAR perception task.

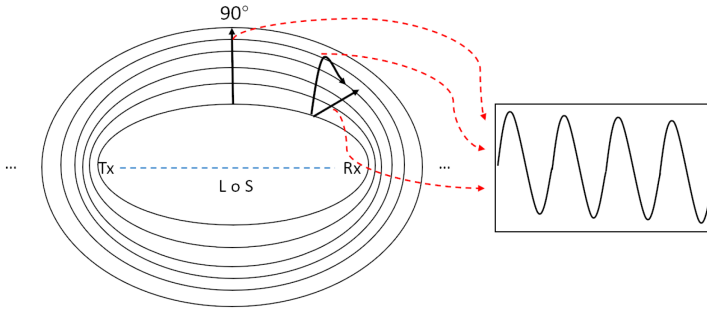


Fig. 5. The same signal at different positions in the Fresnel zone.

In the domain of small-scale sensing, Zhang et al. [129] introduced the application of the Fresnel diffraction model for human respiration monitoring and analysis within the first Fresnel zone. The study analyzed how moving objects within the first Fresnel zone affect received radio frequency signals and modeled human targets as cylinders of different sizes. The study revealed multiple unique characteristics of respiration sensing within this zone, predicting respiration detectability based on chest front-side and back-side movements and achieving high-accuracy respiration monitoring. Additionally, the study tested that front-side chest movement during respiration was the dominant factor affecting the test, but back-side chest movement during respiration also influenced the overall effectiveness of respiration sensing. Furthermore, differences in individuals' body thickness and variations in body posture during testing had varying effects on respiration sensing.

From a geometrical space perspective, the Fresnel zones establish the mathematical relationship between the CSI signal fluctuation patterns and the position and motion trajectory of the perceived target relative to the transmitting and receiving devices. Objects located in different Fresnel zones may produce identical or varied signals. Figure 5 illustrates a schematic where identical signals are emitted from different positions within the Fresnel zones. The figure shows that different locations within the Fresnel zone may emit identical signals, highlighting the challenges and unique characteristics of Wi-Fi signal detection. The gesture recognition for the HAR task is shown in Figure 6. The gesture signal detection and classification process is shown in Figure 7.

Regarding the use of Wi-Fi signals for gesture sensing, Moghaddam et al. [130] conducted relevant tests. The study utilized a dataset recording 13 different interactions between two people, with a focus on four interaction actions: approaching, clapping hands, kicking with the

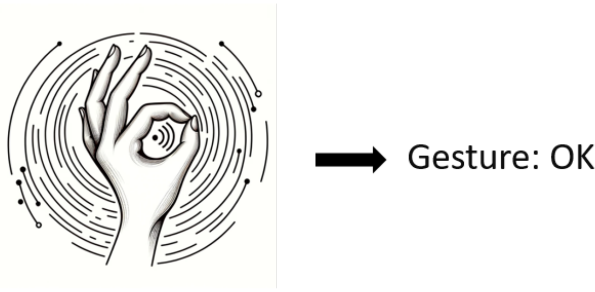


Fig. 6. Illustration of human gesture detection.

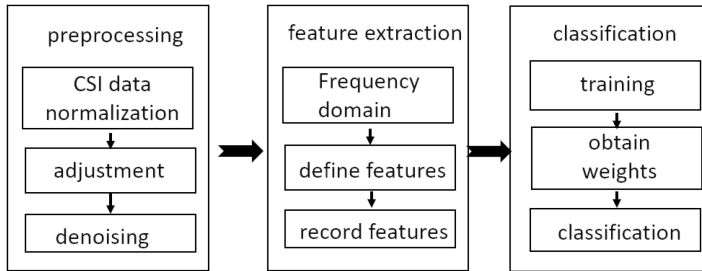


Fig. 7. Gesture signal detection and classification process.

left leg, and pointing with the right hand. The proposed method used CSI and RSSI data for recognition, incorporating seven frequency features from CSI data and one frequency feature from RSSI data. Extracted features were employed to train a K-nearest neighbors classifier on the dataset.

For sleep monitoring using Wi-Fi signals, Yu et al. [131] presented relevant insights and designs. Their study noted that wearing sensors throughout the night could cause discomfort to subjects, limiting real-life usage. The study utilized fine-grained channel state information from multiple antennas and advanced fusion and signal processing methods to extract accurate respiration and body movement information. A deep learning approach was introduced, combining clinical sleep medicine knowledge to achieve four-stage sleep monitoring with limited data sources (only respiration and body movement information). Wi-Fi-Sleep's performance was evaluated against the gold standard, polysomnography, in 19 comprehensive experiments involving 12 subjects in an environment resembling a residential bedroom.

In the field of indoor positioning using Wi-Fi signals, there are primarily two types of methods: geometry-based methods and fingerprint-based methods. Geometry-based methods include angle-based and distance-based techniques for estimating angles based on received base station signals. Fingerprint-based CSI positioning associates collected signal characteristics with specific positions within a room. Compared to other positioning techniques such as AOA (Angle of Arrival) and TOA (Time of Arrival), fingerprint positioning offers advantages of simplicity, high accuracy, and algorithm flexibility. Wi-Fi, as an easily accessible signal source, is widely employed in fingerprint positioning. Wi-Fi signal fingerprint positioning typically involves offline and online stages. In the offline stage, fingerprints from different locations are collected, and a fingerprint database is established, whereas in the online stage, the system matches received signal features with signals in the database for positioning [132–145]. The application of Wi-Fi signals in indoor positioning is shown in Figure 8.

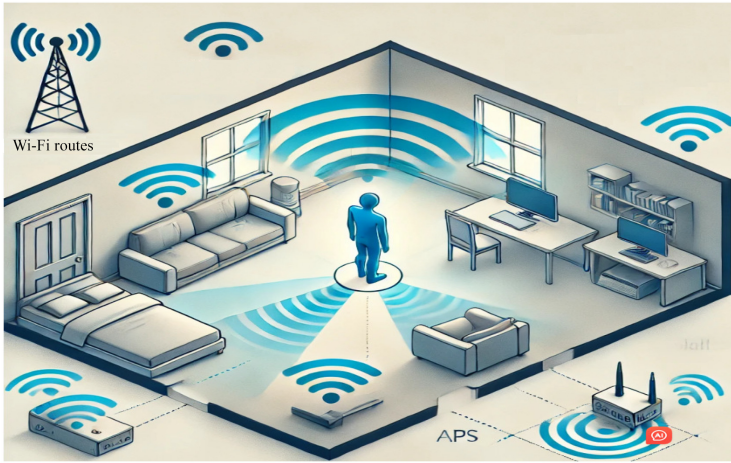


Fig. 8. Wi-Fi signals and indoor positioning.

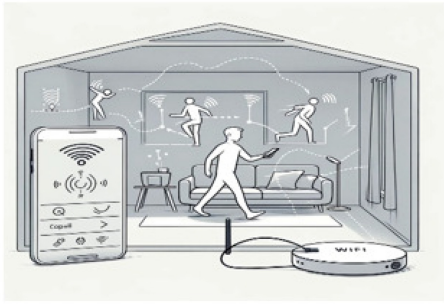


Fig. 9. Wi-Fi signals and HAR.

The recent advancements in human pose estimation utilizing Wi-Fi signals are significant. Lin et al. [146] demonstrate a system using smartphones and commercial Wi-Fi routers for recognizing human activities at various scales. This system leverages Wi-Fi CSI captured by smartphones, which are modified with custom firmware and software. Machine learning models, including **Convolutional Neural Networks (CNNs)**, Decision Trees, and LSTMs, are employed to classify activities based on features extracted from CSI data. The system's efficacy is evaluated through experiments involving 20 types of human movements with 60 participants, generating 36,000 data points.

Hao et al. [147] focus on the recognition of yoga postures using a GAN-based approach integrated with CNNs. Data from 10 individuals performing 14 yoga poses are used to train the model, which assesses the accuracy of the poses. Figure 9 shows the application of Wi-Fi signals in the HAR field.

### 3.6 Open Problems and Challenges

The challenges faced by non-contact sensing technology based on Wi-Fi are presented next.

- *Dataset Quality.* Many current datasets are collected in idealized environments, which do not reflect real-world application scenarios. As a result, these datasets often lack interference factors present in real environments, such as walls, furniture, noise, and multi-person

interactions, failing to accurately capture Wi-Fi signal behavior in complex conditions. To enhance the practical effectiveness of models, it is essential to gather data from more realistic scenarios, considering factors like complex indoor structures, diverse devices, and sources of interference. This will facilitate the development of more robust and generalizable Wi-Fi signal datasets, thereby improving the accuracy and reliability of algorithms in real-world applications.

- *Deployment on Edge Devices.* The deployment challenges on edge devices mainly involve making models lightweight and addressing the complexities encountered in real-world applications. Due to the limited computational power, storage, and battery life of edge devices, models must be compressed and optimized to run efficiently on these devices. Additionally, real-world applications may involve complexities such as multi-device coordination, real-time processing, and network fluctuations. Therefore, model selection must carefully consider computational efficiency, energy consumption, and adaptability to complex environments, ensuring effective operation within the constraints of edge devices.
- *Detection Efficiency.* Detection efficiency is critical for practical applications, particularly in scenarios requiring rapid response. For instance, in applications such as real-time monitoring or smart device control, the model's reaction speed to various signals directly affects the overall system performance and user experience. If the model's response is too slow, data processing delays may occur, compromising the system's real-time capabilities and its ability to handle unexpected situations. Moreover, for edge devices and IoT systems, detection efficiency impacts not only computational load but also power consumption. Efficient detection ensures that devices can operate continuously under low power conditions, extending battery life and improving system reliability. Therefore, ensuring that models are capable of fast and accurate responses to different signal environments is a critical requirement in practical applications.
- *Model Generalization and the Application of Deep Learning Models.* Model generalization refers to the ability of machine learning models, particularly deep learning models, to perform well on unseen data. This capability is crucial in practical applications, as it ensures that the model not only excels on the training data but can also accurately handle new, previously unencountered data. This characteristic is critical for the deployment of models in a wide range of real-world scenarios, such as image classification, speech recognition, and natural language processing tasks.

Regarding dataset quality, we have enumerated several high-quality datasets for consideration in the preceding text, and an increasing number of researchers are developing superior datasets for utilization. The deployment of models on edge devices presents significant challenges, primarily related to the need for lightweight designs and the complexities of real-world applications, which impose stringent criteria for model selection. Additionally, ensuring detection efficiency, model generalization, and practical applicability are critical concerns. These topics will be elaborated in subsequent sections.

#### 4 Deep Learning Models and Generalization Issues

In Wi-Fi-based non-contact sensing and human pose estimation, deep learning exhibits substantial potential, largely due to its capability to analyze and extract features from complex Wi-Fi signal data, leading to high-accuracy recognition tasks. However, generalization remains a significant challenge. This issue manifests when models trained on specific datasets underperform on new, unobserved datasets, a common scenario in real-world deployments like home environments. Addressing the generalization capability of deep learning algorithms in human pose estimation is critical for practical applications.

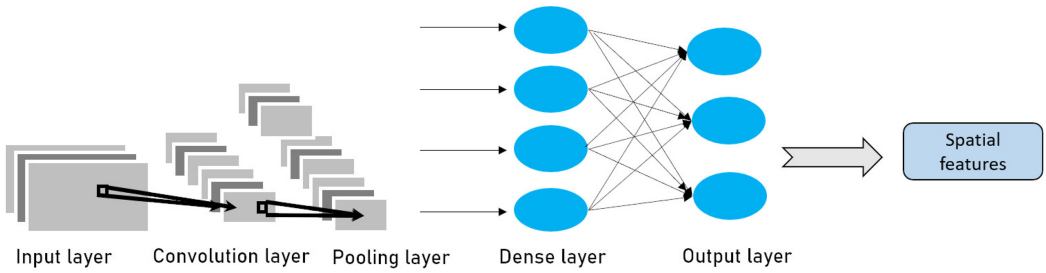


Fig. 10. Simplified diagram of a CNN network.

This section will focus on delineating neural network models prevalent in CSI signal based human pose estimation, highlighting their utility and addressing the intrinsic challenge of model generalization across diverse environments.

#### 4.1 Convolutional Neural Networks

CNNs, originally prevalent in image processing, have gained traction in CSI signal sensing due to their ability to autonomously learn spatial features from data. A typical CNN architecture includes layers such as input, convolutional, pooling, and fully connected layers. The convolution operation employs filters to extract features through dot products with input vectors, whereas pooling layers perform downsampling to reduce dimensionality.

In CSI sensing, the use of CNNs is particularly effective for extracting spatial features from the subcarrier dimensions of CSI data, crucial for detecting patterns and changes within the signal. Research has explored transforming CSI signals into multi-dimensional representations, leveraging CNNs for deep feature extraction [148]. Other approaches have focused on using the amplitude of CSI signals in combination with CNNs and GRUs for enhanced feature learning and recognition tasks [149].

The application of CNNs in Wi-Fi sensing offers advantages in terms of reduced training parameters and the ability to preserve key features in CSI data, including subcarrier and temporal dimensions. This capability is vital for complex detection tasks in varied environments, such as indoor human pose and gesture recognition. Despite its potential, challenges persist in areas like environmental adaptability, data quality, noise handling, and real-time processing demands, necessitating further research to optimize CNN applications in diverse Wi-Fi sensing scenarios [150]. Figure 10 presents a schematic diagram of a CNN network.

#### 4.2 LSTM Networks

LSTM networks, a class of deep learning models, are adept at processing CSI signals. LSTM networks feature gate mechanisms including forget gates, input gates, and output gates, which effectively manage and retain long-term information, thereby addressing the problem of gradient vanishing. The core of an LSTM network is its cell state, responsible for transmitting important information across sequences.

LSTM networks show great potential in processing CSI signals, particularly in applications like Wi-Fi sensing, indoor localization, and behavior recognition. Their ability to handle long-term sequential data enables them to effectively capture complex temporal patterns in CSI signals, filter out noise, and adapt to changes in dynamic environments. Additionally, LSTM networks are capable of real-time or near-real-time data processing, making them suitable for applications that require quick responses. In terms of behavior recognition, LSTM networks can accurately identify different human activities based on the temporal features of CSI signals, including walking,

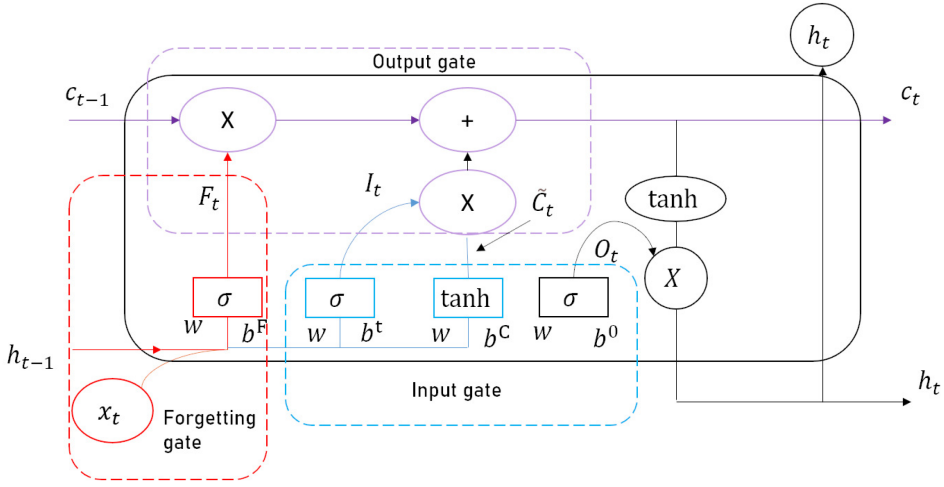


Fig. 11. Simplified diagram of an LSTM network.

running, sitting, and more, providing crucial support for analyzing human behaviors in smart environments. The application of LSTM in the HAR field has yielded numerous achievements due to its excellent ability in handling time series. Despite challenges related to computational complexity and data handling, LSTM networks hold promising prospects in CSI signal analysis [151–158]. A simplified diagram of an LSTM network is presented in Figure 11.

The operation of an LSTM network sequentially controls the flow of information through four key components: the forget gate, input gate, memory cell update, and output gate. First, the forget gate  $F_t$  determines how much of the previous memory state  $c_{t-1}$  should be retained by using the current input  $x_t$  and the previous hidden state  $h_{t-1}$ . The output of the forget gate is computed as  $\sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ . Next, the input gate  $I_t$  manages the amount of new information to be added to the memory cell by combining the current input and hidden state. It controls the updating of the candidate memory state  $\tilde{C}_t$ , which is calculated as  $\tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$ . The memory cell state  $C_t$  is then updated based on the outputs from both the forget gate  $F_t$  and input gate  $I_t$ . Specifically, the term  $F_t \cdot c_{t-1}$  determines how much of the previous memory state is kept, whereas  $I_t \cdot \tilde{C}_t$  controls how much new information is added.

Afterward, the output gate  $O_t$  determines the hidden state  $h_t$  at the current timestep, controlling how much of the memory information should be passed on. The output gate computes its value as  $\sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ , and the hidden state is updated as  $h_t = O_t \cdot \tanh(c_t)$ . The hidden state  $h_t$  is the output of the LSTM unit at this timestep and is passed forward to the next step in the sequence.

In LSTM networks, the weights and biases are denoted as  $W$  and  $b$ , respectively. Two activation functions are utilized in the formulas: the sigmoid function  $\sigma(\cdot)$ , responsible for gating mechanisms, and the hyperbolic tangent function  $\tanh(\cdot)$ , which manages memory cell updates and hidden states. The symbol  $*$  represents convolution operations when applied. The key parameters include  $h_t$  (the hidden state),  $I_t$  (input gate),  $F_t$  (forget gate),  $O_t$  (output gate),  $\tilde{C}_t$  (candidate cell state), and  $C_t$  (current cell state). Together, they form the mechanism that updates the input  $x_t$ , the hidden state  $h_{t-1}$ , and the previous cell state  $c_{t-1}$ , facilitating information flow through the LSTM network. Here are the specific formulas for an LSTM network:

$$F_t = \sigma \left( W^F [h_{t-1}, x_t] + b^F \right), \quad (3a)$$

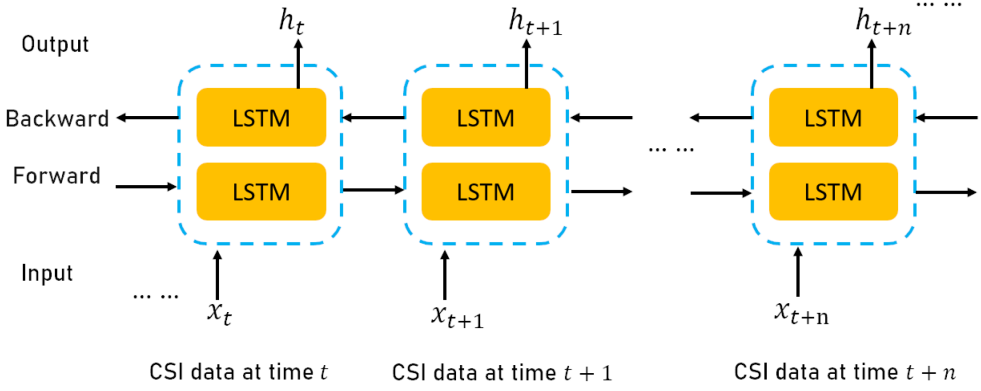


Fig. 12. Simplified diagram of a BLSTM network.

$$I_t = \sigma \left( W^I [h_{t-1}, X_t] + b^I \right), \quad (3b)$$

$$\tilde{C}_t = \tanh \left( W^C [h_{t-1}, X_t] + b^C \right), \quad (3c)$$

$$C_t = F_t \cdot C_{t-1} + I_t \cdot \tilde{C}_t, \quad (3d)$$

$$O_t = \sigma \left( W^O [h_{t-1}, X_t] + b^O \right). \quad (3e)$$

### 4.3 Bidirectional Long Short-Term Memory Network

**Bidirectional Long Short-Term Memory (Bi-LSTM or BLSTM)** networks are a type of deep learning model that enhance contextual understanding by processing sequence data both forward and backward. Unlike the previously mentioned LSTM networks, which only consider current and past channel state information, Bi-LSTM networks integrate both past and future information at each time point. This dual-layered approach, comprising one LSTM layer processing data in normal chronological order and another in reverse order, allows Bi-LSTM networks to excel in tasks requiring comprehensive context comprehension, such as text classification, named entity recognition, speech recognition, and machine translation, as highlighted in the work of Kadir et al. [160] and Kaushik et al. [161]. A simplified diagram of a BLSTM network is presented in Figure 12.

The Bi-LSTM architecture, with its capability to merge information from past and future contexts, shows superior performance in fields requiring a thorough understanding of sequential data [162]. However, this enhanced capability comes with increased computational costs and complexity, necessitating careful consideration and optimization in practical applications [163–165]. The input and output of a Bi-LSTM network can be represented as follows:

$$h_t = \overrightarrow{h}_t \oplus \overleftarrow{h}_t. \quad (4)$$

### 4.4 Gated Recurrent Unit

The GRU network, a streamlined variant of the LSTM, reduces computational requirements by integrating LSTM's forget and input gates. Efficient in processing CSI signals, GRUs excel in analyzing and predicting complex temporal data. Their gating mechanisms adeptly capture time dynamics and spatial variations in signals, vital for applications like indoor localization and human activity monitoring through Wi-Fi signals.

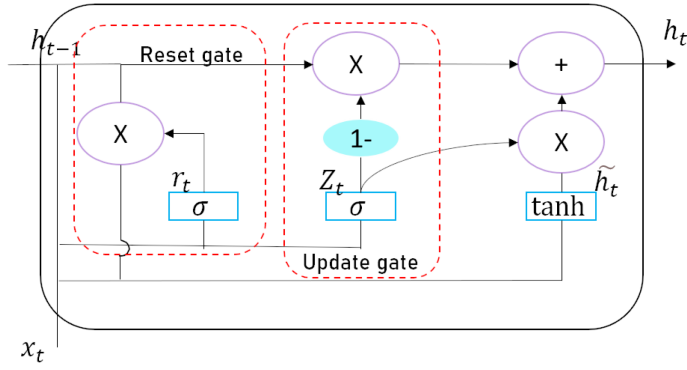


Fig. 13. Simplified diagram of a GRU network.

GRUs offer a balance between memory retention and new information acquisition, key in adapting to environmental changes and capturing essential features. Their simpler structure compared to LSTMs makes them suitable for environments with limited resources or small datasets, enabling quick training with effective outcomes [166–169]. This positions GRUs as a favorable option for real-time applications in CSI signal processing, such as motion recognition. A simplified diagram of a GRU network is presented in Figure 13.

Studies like the one by Tang et al. [149] have explored combining one-dimensional CNNs with GRUs for dual extraction of spatial and temporal features from CSI signals. The CNN-GRU model has been adapted for gesture recognition in systems like Widar [170]. Additionally, Zhang et al. [171] proposed a hybrid network of one-dimensional CNNs and BLSTMs, targeting small-sample gesture recognition scenarios. In the GRU network, the flow of information is controlled by two gates: the reset gate  $r_t$  and the update gate  $Z_t$ . First, the reset gate  $r_t$  is calculated by the formula  $r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r)$ , which determines how much information from the previous hidden state  $h_{t-1}$  needs to be “reset” or ignored, to contribute to the computation of the current candidate hidden state  $\tilde{h}_t$ . The candidate hidden state  $\tilde{h}_t$  is obtained by the formula  $\tilde{h}_t = \tanh(W_h \cdot [r_t \cdot h_{t-1}, x_t] + b_h)$ , representing the possible new hidden state at the current timestep, calculated by combining the current input  $x_t$  and the reset  $h_{t-1}$ . Then, the update gate  $Z_t$  is calculated by the formula  $Z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z)$ , which determines how much of the current hidden state  $h_t$  should retain the old hidden state  $h_{t-1}$  and how much should incorporate the new candidate hidden state  $\tilde{h}_t$ . Finally, the current hidden state  $h_t$  is computed as  $h_t = Z_t \cdot h_{t-1} + (1 - Z_t) \cdot \tilde{h}_t$ . The update gate is similar to the forget gate and input gate in the LSTM, controlling how much old information to retain and how much new information to accept, whereas the reset gate determines which parts of the past hidden state are irrelevant to the current computation. Through these two gates, the GRU effectively controls the transmission and update of information, addressing long-term dependency problems. Here are the formulas for the GRU network:

$$Z_t = \sigma(W_z \cdot [h_{t-1}, x_t]), \quad (5a)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]), \quad (5b)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]), \quad (5c)$$

$$h_t = (1 - Z_t) * h_{t-1} + Z_t * \tilde{h}_t. \quad (5d)$$

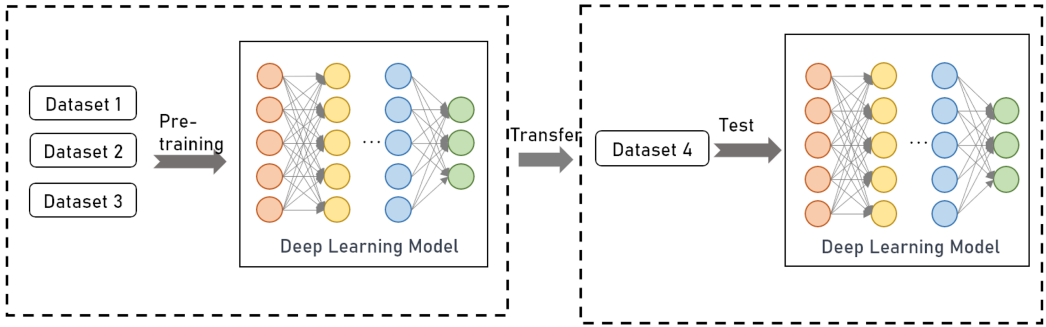


Fig. 14. Testing the transfer learning models for generalization ability.

Table 4. Testing Generalization Abilities

| Model                 | LSTM   | BLSTM  | ABLSTM | GRU    | CNN    |
|-----------------------|--------|--------|--------|--------|--------|
| Accuracy in Pattern A | 85.29% | 85.86% | 86.76% | 82.19% | 81.05% |
| Accuracy in Pattern B | 86.23% | 85.56% | 86.61% | 80.27% | 74.80% |
| Accuracy in Pattern C | 91.99% | 84.18% | 92.08% | 89.65% | 85.89% |
| Accuracy in Pattern D | 78.61% | 77.84% | 82.42% | 79.84% | 79.66% |
| Average Accuracy      | 85.53% | 83.36% | 86.97% | 82.99% | 80.35% |

Note: Patterns A, B, C, and D use Datasets 4, 3, 2, and 1 as test sets, respectively.

#### 4.5 Generalization Ability of Deep Learning Networks

Generalization ability refers to the performance of a machine learning model when predicting unseen new data. In the context of HAR, considering that new data cannot always be immediately acquired for training in certain environments, it is necessary to contemplate training on old datasets and then testing in new environments [172–184].

To explore the issue of generalization capabilities across various models, we have designed a transfer learning model and prepared relevant datasets. Figure 14 shows our rough model. Furthermore, to assess the accuracy and generalization of different models, we have conducted sequential testing on LSTM, BLSTM, GRU, and ABLSTM. We trained each model using three datasets and then evaluated their accuracy on a single test dataset to measure their generalization capabilities. We denote this process as follows. Pattern A involves training on datasets 1, 2, and 3, and testing generalization on dataset 4. Pattern B entails training on datasets 1, 2, and 4, and testing generalization on dataset 3. Pattern C comprises training on datasets 1, 3, and 4, and testing generalization on dataset 2. Last, Pattern D includes training on datasets 2, 3, and 4, and testing generalization on dataset 1. These patterns enable us to evaluate how well each model generalizes to new data by training them on different combinations of datasets and then assessing their performance on a separate dataset in each case. This approach helps us gauge the models' ability to adapt and perform on unseen data, which is a critical aspect of evaluating their generalization capabilities.

The results are shown in Table 4. From the charts, it is evident that in different testing modes, the accuracy achieved by LSTM is consistently higher than that of GRU and CNN networks. BLSTM's accuracy is on par with LSTM, but ABLSTM, which incorporates an attention mechanism, exhibits higher accuracy than all other networks. This indicates that ABLSTM possesses the strongest overall generalization capabilities and self-recognition accuracy among the tested models.

Additionally, we present the results in a radar chart in Figure 15, which more clearly shows that the generalization capability of the ABLSTM network surpasses all other networks. The

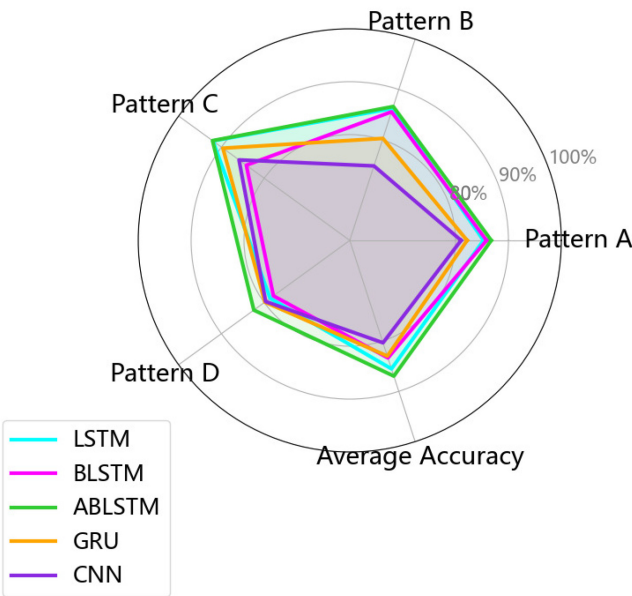


Fig. 15. Testing generalization abilities.

Table 5. Accuracy for Different Actions

| Model                 | LSTM   | BLSTM  | ABLSTM        | GRU           | CNN    |
|-----------------------|--------|--------|---------------|---------------|--------|
| Accuracy for Action 1 | 98.69% | 98.01% | <u>98.83%</u> | <u>98.84%</u> | 97.29% |
| Accuracy for Action 2 | 97.33% | 97.98% | <u>98.15%</u> | 97.99%        | 95.52% |
| Accuracy for Action 3 | 94.32% | 92.41% | <u>98.08%</u> | 96.40%        | 94.47% |
| Accuracy for Action 4 | 95.01% | 96.18% | <u>98.16%</u> | 97.44%        | 93.51% |
| Accuracy for Action 5 | 98.10% | 98.08% | <u>99.69%</u> | 98.57%        | 98.27% |
| Accuracy for Action 6 | 84.64% | 82.91% | <u>92.83%</u> | 89.37%        | 82.26% |
| Accuracy for Action 7 | 89.85% | 88.69% | <u>93.84%</u> | 92.43%        | 90.91% |

Note: Actions 1 through 7 sequentially refer to waving hands, clapping, walking, lying down, sitting, falling, and standing up after falling.

generalization ability of the BLSTM can approach that of the ABLSTM in certain modes, but the ABLSTM network with added attention mechanisms is noticeably more balanced and stable.

In addition, we conducted generalization tests for different actions across various models, and the results are presented in Table 5. This table displays the testing accuracy for different actions. From these results, we can observe that neural networks tend to have varying recognition accuracy for different actions, with most networks showing higher accuracy in recognizing small-scale movements like clapping or waving compared to recognizing larger-scale actions.

We have also created a radar chart, shown in Figure 16. In the chart, it is evident that only the ABLSTM network, which incorporates attention mechanisms, can stably recognize large-scale actions such as Action 6 (falling), demonstrating superior generalization ability. Other networks exhibit weaker generalization capabilities in recognizing this action.

Generalization is a significant challenge in our research on Wi-Fi-based human pose recognition. Different networks exhibit different generalization capabilities, and stronger generalization

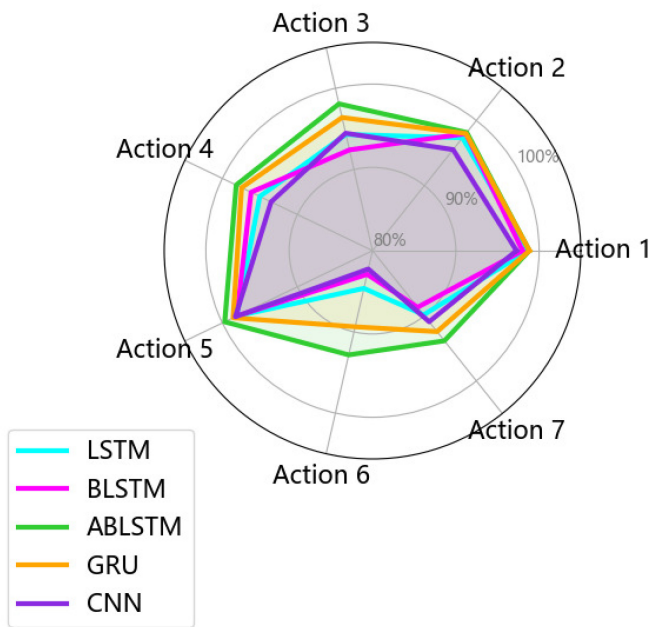


Fig. 16. Testing the accuracy for different actions.

Table 6. Summary of Network Models and Features in Wi-Fi-Based HAR Works

| Authors                     | Network Model                        | Key Features                                                          |
|-----------------------------|--------------------------------------|-----------------------------------------------------------------------|
| Fu et al. [186]             | CNN + Bi-LSTM                        | Through-wall recognition with >99% accuracy                           |
| Abuhoureyah et al. [187]    | Multi-Head Attention (MHAAM)         | Attention mechanism for complex activity recognition (98.5% accuracy) |
| Valaboju et al. [188]       | Inception Time + LSTM                | Improved hardware setups for HAR                                      |
| Padua et al. [189]          | SVM                                  | Multi-person gesture recognition with 93% accuracy                    |
| Abuhoureyah et al. [190]    | Multi-Layer LSTM                     | Signal decomposition for multi-user scenarios                         |
| Strohmayer and Kampel [191] | EfficientNetV2                       | Cross-scenario and cross-system generalization with data augmentation |
| Cui et al. [192]            | WiCAU (Wavelet + DNN)                | Domain adaptation for cross-environment HAR                           |
| Sheng et al. [193]          | Dual Adversarial Network             | Environment-independent feature extraction for HAR                    |
| Liu et al. [194]            | High-Dimensional Factor Model (HDFM) | Improved recognition accuracy by 6.8%                                 |

can assist us in addressing real-world issues such as variations in test results across different environments. This is a key focus and challenge in the field of Wi-Fi-based human pose recognition.

4.6 Introduction to the Latest HAR Methods and Applied Networks

Table 6 presents a comprehensive summary of the most recent advancements in HAR using Wi-Fi signals, focusing on cutting-edge neural network models and their key contributions. Each entry in the table emphasizes the most up-to-date research, with studies published in 2023 and 2024, showcasing the latest methodologies and technologies applied to HAR tasks.

The network models covered in these recent studies include advanced techniques such as CNN, Bi-LSTM, SVM (Support Vector Machines), and MHAAM (Multi-Head Adaptive Attention Mechanism). These models are employed to solve complex HAR challenges, including through-wall activity recognition and multi-user gesture detection. For instance, models like CNN and Bi-LSTM have been used for through-wall activity recognition, achieving remarkable accuracy levels exceeding 99%. The MHAAM model introduces attention mechanisms that significantly enhance the model's ability to recognize intricate human activities, with an accuracy rate reaching up to 98.5%. Additionally, SVM-based models have been effectively utilized for gesture recognition in multi-person scenarios, delivering strong performance with an accuracy of 93%.

Table 6 also highlights the innovative techniques used to address modern challenges such as cross-environment and cross-device generalization. State-of-the-art approaches, such as data augmentation and domain adaptation, have been applied to improve the adaptability of HAR systems across different environments and hardware setups. Models like EfficientNetV2 and HDFM (High-Dimensional Factor Model) further enhance performance by refining noisy CSI data, leading to a 6.8% improvement in recognition accuracy.

This detailed comparison of the latest 2023 and 2024 studies provides valuable insights for researchers and practitioners. It serves as a guide for understanding the capabilities of different network models in HAR applications, from smart home environments to healthcare and security systems. The inclusion of the most recent research ensures that the insights and methodologies are at the forefront of HAR technology advancements.

## 5 Potential Challenges and Research Directions

This article provides a comprehensive summary of the current state of human pose recognition, emphasizing the primary approaches employed in this field. It places particular emphasis on the current research status of human pose recognition based on Wi-Fi signals, discusses relevant datasets in the domain of human pose recognition, introduces neural network architectures based on deep learning, and assesses their generalization capabilities for human pose recognition. Certainly, a more concise and formal revision is discussed next.

In the realm of Wi-Fi-based human pose recognition, several challenges and future directions deserve attention. First, there is a scarcity of high-quality datasets covering diverse categories. Creating such datasets will require significant investment [185, 195–198].

Moreover, exploring methods like semi-supervised and unsupervised learning with limited data, as well as techniques like small-sample learning and transfer learning, can boost data utilization. Additionally, research on data generation methods like data fusion and generative adversarial approaches is ongoing [199, 200].

Furthermore, applying transfer learning, domain adaptation, and meta-learning can enhance cross-domain recognition in Wi-Fi sensing, making systems adaptable across various environments. Therefore, future Wi-Fi sensing research should focus on dataset creation and utilization [201–204].

In the current landscape, cloud-edge systems are progressively maturing, leading to a growing emphasis on real-time detection and on-site sensing in the realm of Wi-Fi sensing. Consequently, there is a burgeoning adoption of edge devices, prioritizing their portability and efficiency. Traditional algorithmic models are characterized by their substantial memory consumption, making them unsuitable for deployment on edge devices. To align with the memory and computational constraints of edge devices, it becomes imperative to explore techniques such as algorithm lightweighting, encompassing processes like model compression, pruning, quantization, and distillation. Hence, the future trajectory of Wi-Fi sensing research should prominently feature efforts directed toward algorithmic lightweighting [205–208].

Wi-Fi sensing is advantageous due to its wide coverage, cost-effectiveness, and immunity to non-LoS scenarios. However, existing Wi-Fi behavior sensing models often suffer from robustness issues. While they excel in specific scenarios, any changes in the environment or objects can significantly degrade their performance.

To address this challenge, integrating Wi-Fi sensing with other sensor modalities is a promising solution. This creates multimodal sensing data by combining Wi-Fi CSI data with other sensor data. Learning from these combined data enables algorithm models to adapt reliably to different task objectives. Notably, WiVi, introduced by Hao et al. [209], integrated Wi-Fi sensing with video information, marking the first use of multimodal learning in Wi-Fi sensing. In the work of Kefayati et al. [210], WiVi was enhanced with WiNN, a novel neural network that significantly improved activity recognition robustness.

Cross-modal learning, exemplified in the work of Wang et al. [211] and Yang et al. [212], involves translating information between different modalities. These works generated video frames from CSI data, pioneering cross-modal learning in Wi-Fi sensing. Multimodal learning enables Wi-Fi sensing to draw insights from various domains, including computer vision, radar, and radio frequency signals like laser-based systems.

Presently, Wi-Fi sensing research primarily focuses on enhancing sensing model precision and stability. While Wi-Fi sensing poses fewer privacy concerns than alternatives like computer vision, it is not immune to security challenges, especially as cloud-based systems advance and CSI signals are collected from edge devices.

To address these security issues, various approaches have been proposed. For example, in the work of Yang et al. [213], strengthening adversarial training helped counter potential attacks on Wi-Fi sensing systems. In a different work, Yang et al. [85], introduced a method employing variational autoencoders for CSI quantization, ensuring efficient communication and privacy protection. Additionally, Zhou et al. [214] used data unrecognizable by a general classifier enhances user privacy protection.

Furthermore, federated learning holds promise in multi-user integrated learning, potentially establishing large-scale, secure, and stable Wi-Fi sensing models in the future.

In conclusion, this article initially delineated sensor-based recognition, computer vision, and Wi-Fi signal based HAR methods. It then delved into tools for Wi-Fi-based HAR signal collection, listing several high-quality datasets. Following this, the document reviewed an array of sensing tasks facilitated by Wi-Fi signal recognition, with a focus on the deployment of deep learning networks for Wi-Fi signal detection. Next, we elucidated the experimental results from evaluating the capabilities of various networks. We then discussed challenges and future prospects within the HAR domain.

In summary, sustained innovation and interdisciplinary research are crucial for overcoming existing challenges and fully unleashing the potential of Wi-Fi-based HAR. Wi-Fi signal based HAR stands at the vanguard of technological evolution, heralding a future where our interaction with the digital realm is more seamless, intuitive, and secure. With a comprehensive understanding of this technology and its developmental trajectory, we should be better equipped to navigate the inherent issues and challenges.

## References

- [1] W. Jiao and C. Zhang. 2023. An efficient human activity recognition system using WiFi channel state information. *IEEE Systems Journal* 17, 4 (2023), 6687–6690.
- [2] S. Li, Y. Ge, M. Shentu, S. Zhu, M. Imran, and Q. Abbasi. 2021. Human activity recognition based on collaboration of vision and WiFi signals. In *Proceedings of the 2021 International Conference on UK-China Emerging Technologies (UCET'21)*. 204–208.

- [3] M. Miyazaki, S. Ishida, A. Fukuda, T. Murakami, and S. Otsuki. 2019. Initial attempt on outdoor human detection using IEEE 802.11ac WLAN signal. In *Proceedings of the 2019 IEEE Sensors Applications Symposium (SAS'19)*. 1–6.
- [4] J. Chen, K. Yang, X. Zheng, S. Dong, L. Liu, and H. Ma. 2023. WiMix: A lightweight multimodal human activity recognition system based on WiFi and vision. In *Proceedings of the 2023 IEEE 20th International Conference on Mobile Ad Hoc and Smart Systems (MASS'23)*. 406–414.
- [5] G. Zhao, Z. Zhou, Y. Huang, A. Nayak, W. Gong, and H. Zhou. 2023. ALSensing: Human activity recognition using WiFi based on active learning. In *Proceedings of the 2023 IEEE Wireless Communications and Networking Conference (WCNC'23)*. 1–6.
- [6] I. O. Joudeh, A.-M. Cretu, R. B. Wallace, R. A. Goubran, A. Alkhalid, and M. Allegue-Martinez. 2019. WiFi channel state information-based recognition of sitting-down and standing-up activities. In *Proceedings of the 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA'19)*. 1–6.
- [7] Z. Xiao, F. Mengyin, Y. Yi, and L. Ningyi. 2012. 3D human postures recognition using Kinect. In *Proceedings of the 2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics*. 344–347.
- [8] A. Leone, G. Rescio, A. Caroppo, P. Siciliano, and A. Manni. 2023. Human postures recognition by accelerometer sensor and ML architecture integrated in embedded platforms: Benchmarking and performance evaluation. *Sensors* 23, 2 (2023), 1039.
- [9] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu. 2021. Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities. *ACM Computing Surveys* 54, 4 (2021), Article 70, 40 pages.
- [10] L. M. Dang, K. Min, H. Wang, Md. J. Piran, C. H. Lee, and H. Moon. 2020. Sensor-based and vision-based human activity recognition: A comprehensive survey. *Pattern Recognition* 108 (2020), 107561.
- [11] S. K. Yadav, K. Tiwari, H. M. Pandey, and S. A. Akbar. 2021. A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions. *Knowledge-Based Systems* 223 (2021), 106970.
- [12] B. Fu, N. Damer, F. Kirchbuchner, and A. Kuijper. 2020. Sensing technology for human activity recognition: A comprehensive survey. *IEEE Access* 8 (2020), 83791–83820.
- [13] O. Nafea, W. Abdul, G. Muhammad, M. Alsulaiman, and M. A. Bencherif. 2021. Sensor-based human activity recognition with spatio-temporal deep learning. *Sensors* 21, 6 (2021), 2141.
- [14] S. Qiu, H. Zhao, N. Jiang, Y. An, Y. Li, and J. Liu. 2022. Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges. *Information Fusion* 80 (2022), 241–265.
- [15] E. Ramanujam, T. Perumal, and S. Padmavathi. 2021. Human activity recognition with smartphone and wearable sensors using deep learning techniques: A review. *IEEE Sensors Journal* 21, 12 (2021), 13029–13040.
- [16] N. Ahmed, J. I. Rafiq, and M. R. Islam. 2020. Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model. *Sensors* 20, 1 (2020), 317.
- [17] S. Zhang, Y. Li, S. Zhang, L. Shahabi, J. L. Deng, and B. Mortazavi. 2022. Deep learning in human activity recognition with wearable sensors: A review on advances. *Sensors* 22, 4 (2022), 1476.
- [18] S. K. Challa, A. Kumar, and V. B. Semwal. 2022. A multibranch CNN-BiLSTM model for human activity recognition using wearable sensor data. *Visual Computer* 38, 12 (2022), 4095–4109.
- [19] Y. Li, G. Yang, Z. Su, L. Wang, M. Zhou, and X. Chen. 2023. Human activity recognition based on multienvironment sensor data. *Information Fusion* 91, (2023) 47–63.
- [20] I. Dirgová Luptáková, M. Kubovčík, and J. Pospíchal. 2022. Wearable sensor-based human activity recognition with transformer model. *Sensors* 22, 5 (2022), 1911.
- [21] K. Xia, J. Huang, and H. Wang. 2020. LSTM-CNN architecture for human activity recognition. *IEEE Access* 8 (2020), 56855–56866.
- [22] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi. 2020. Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey. *IEEE Access* 8 (2020), 210816–210836.
- [23] S. Mekruksavanich and A. Jitpattanakul. 2021. LSTM networks using smartphone data for sensor-based human activity recognition in smart homes. *Sensors* 21, 5 (2021), 1636.
- [24] Y. Xu and T. T. Qiu. 2021. Human activity recognition and embedded application based on convolutional neural network. *Journal of Artificial Intelligence and Technology* 1, 1 (2021), 51–60.
- [25] V. Bijalwan, V. B. Semwal, and V. Gupta. 2022. Wearable sensor-based pattern mining for human activity recognition: Deep learning approach. *Industrial Robot: The International Journal of Robotics Research and Application* 49, 1 (2022), 21–33.
- [26] S. K. Yadav, K. Tiwari, H. M. Pandey, and M. J. R. Alam. 2021. A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions. *Knowledge-Based Systems* 223 (2021), 106970.

- [27] H. Wang, J. Zhao, J. Li, L. Tian, P. Tu, T. Cao, Y. An, K. Wang, and S. Li. 2020. Wearable sensor-based human activity recognition using hybrid deep learning techniques. *Security and Communication Networks* 2020, 1 (2020), 2132138.
- [28] W. Gao, L. Zhang, Q. Teng, and W. Li. 2021. DanHAR: Dual attention network for multimodal human activity recognition using wearable sensors. *Applied Soft Computing* 111 (2021), 107728.
- [29] M. Strackiewicz, P. James, and J. P. Onnela. 2021. A systematic review of smartphone-based human activity recognition methods for health research. *NPJ Digital Medicine* 4, 1 (2021), 148.
- [30] W. Gao, L. Zhang, W. Huang, and X. Liu. 2021. Deep neural networks for sensor-based human activity recognition using selective kernel convolution. *IEEE Transactions on Instrumentation and Measurement* 70, (2021), 1–13.
- [31] N. Dua, S. N. Singh, and V. B. Semwal. 2021. Multi-input CNN-GRU based human activity recognition using wearable sensors. *Computing* 103, 7 (2021), 1461–1478.
- [32] Z. Qi and Y. Lou. 2023. 3D keypoint detection of lying human body using an RGB-D camera. In *Proceedings of the 2023 42nd Chinese Control Conference (CCC'23)*. 7459–7464.
- [33] G. J. Rani, M. F. Hashmi, and A. Gupta. 2023. Surface electromyography and artificial intelligence for human activity recognition—A systematic review on methods, emerging trends, applications, challenges, and future implementation. *IEEE Access* 11 (2023), 105140–105169.
- [34] A. A. Ghany, B. Uguen, and D. Lemur. 2020. A robustness comparison of measured narrowband CSI vs RSSI for IoT localization. In *Proceedings of the 2020 IEEE 92nd Vehicular Technology Conference (VTC'20 Fall)*. 1–5.
- [35] B. Xiang, F. Yan, Y. Zhu, T. Wu, W. Xia, and J. Pang. 2020. UAV assisted localization scheme of WSNs using RSSI and CSI information. In *Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications (ICCC'20)*. 718–722.
- [36] L. Zhang, B. Wang, Y. Xu, J. Liu, H. Jiang, and Y. Luo. 2022. Indoor location method based on RSSI probability distribution and CSI modified model. In *Proceedings of the 2022 Global Conference on Robotics, Artificial Intelligence, and Information Technology (GCRAIT'22)*. 429–433.
- [37] B. Wei, H. Song, J. Katto, and T. Kikkawa. 2023. RSSI-CSI measurement and variation mitigation with commodity Wi-Fi device. *IEEE Internet of Things Journal* 10, 7 (2023), 6249–6258.
- [38] M. K. Huynh and D. A. Nguyen. 2019. A research on automated guided vehicle indoor localization system via CSI. In *Proceedings of the 2019 International Conference on System Science and Engineering (ICSSE'19)*. 581–585.
- [39] R. S. Rosli, M. H. Habaebi, and M. R. Islam. 2018. Characteristic analysis of received signal strength indicator from ESP8266 WiFi transceiver module. In *Proceedings of the 2018 7th International Conference on Computer and Communication Engineering (ICCCE'18)*. 504–507.
- [40] R. S. Rosli, M. H. Habaebi, and M. R. Islam. 2018. Comparative analysis of digital filters for received signal strength indicator. In *Proceedings of the 2018 7th International Conference on Computer and Communication Engineering (ICCCE'18)*. 214–217.
- [41] I.-C. Severin. 2020. Head posture monitor based on 3 IMU sensors: Consideration toward healthcare application. In *Proceedings of the 2020 International Conference on e-Health and Bioengineering (EHB'20)*. 1–4.
- [42] T. K. Vashishth, V. Sharma, B. Kumar, and R. Panwar. 2023. Exploring the role of computer vision in human emotion recognition: A systematic review and meta-analysis. In *Proceedings of the 2023 2nd International Conference on Augmented Intelligence and Sustainable Systems (ICAISS'23)*. 1071–1077.
- [43] Z. Zhang, Y. Liu, A. Li, and M. Wang. 2014. A novel method for user-defined human posture recognition using Kinect. In *Proceedings of the 2014 7th International Congress on Image and Signal Processing*. 736–740.
- [44] D. R. Beddiar, B. Nini, M. Sabokrou, and A. Hadid. 2020. Vision-based human activity recognition: A survey. *Multi-media Tools and Applications* 79, 41 (2020), 30509–30555.
- [45] P. Pareek and A. Thakkar. 2021. A survey on video-based human action recognition: Recent updates, datasets, challenges, and applications. *Artificial Intelligence Review* 54, 3 (2021), 2259–2322.
- [46] Z. Qin, Y. Zhang, S. Meng, and H. Zhao. 2020. Imaging and fusing time series for wearable sensor-based human activity recognition. *Information Fusion* 53 (2020), 80–87.
- [47] S. Wan, L. Qi, X. Xu, and W. Zhang. 2020. Deep learning models for real-time human activity recognition with smartphones. *Mobile Networks and Applications* 25, 2 (2020), 743–755.
- [48] N. Jaouedi, N. Boujnah, and M. S. Bouhlel. 2020. A new hybrid deep learning model for human action recognition. *Journal of King Saud University—Computer and Information Sciences* 32, 4 (2020), 447–453.
- [49] M. Oudah, A. Al-Naji, and J. Chahl. 2020. Hand gesture recognition based on computer vision: A review of techniques. *Journal of Imaging* 6, 8 (2020), 73.
- [50] A. M. F and S. Singh. 2021. Computer vision-based survey on human activity recognition system, challenges and applications. In *Proceedings of the 2021 3rd International Conference on Signal Processing and Communication (ICPSC'21)*. 110–114.

- [51] Y. Wang, H. Liu, K. Cui, A. Zhou, W. Li, and H. Ma. 2021. m-Activity: Accurate and real-time human activity recognition via millimeter wave radar. In *Proceedings of the 2021 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'21)*. 8298–8302.
- [52] D. Srivastav, A. Bajpai, and A. Singhal. 2022. A temporal convolutional neural network based activity recognition model using a real-time two-dimensional single pose estimation framework. In *Proceedings of the 2022 12th International Conference on Cloud Computing, Data Science, and Engineering (Confluence'22)*. 434–438.
- [53] D. R. Beddiar and B. Nini. 2017. Vision based abnormal human activities recognition: An overview. In *Proceedings of the 2017 8th International Conference on Information Technology (ICIT'17)*. 548–553.
- [54] S. Cui, A. Ma, Y. Wan, Y. Zhong, B. Luo, and M. Xu. 2022. Cross-modality image matching network with modality-invariant feature representation for airborne-ground thermal infrared and visible datasets. *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), Article 5606414, 14 pages.
- [55] R. R. Deshpande, M. R. Bhatt, and C. H. R. Madhavi. 2021. Accuracy in depth recovery and 3D image synthesis from single image using multi-color filter aperture and shallow depth of field. *IEEE Access* 9 (2021), 123528–123540.
- [56] M. Mohammadi, S. Mohammadi, H. Hadizadeh, M. Olfati, F. Moradi, G. Tanzifi, and S. Ghaderi. 2024. Brain metastases from breast cancer using magnetic resonance imaging: A systematic review. *Journal of Medical Radiation Sciences* 71, 1 (2024), 133–141.
- [57] Y. Huang, X. Xu, Z. He, Y. Wang, Z. Lu, and F. Shi. 2023. A lightweight road crack and damage detection method using Yolov5s for IoT applications. In *Proceedings of the 2023 IEEE/CIC International Conference on Communications in China (ICCC'23)*. 1–6.
- [58] T. Molet, Z. Huang, R. Boulic, and D. Thalmann. 1997. An animation interface designed for motion capture. In *Proceedings of Computer Animation'97*. 77–85.
- [59] O. Rashid, A. Al-Hamadi, and B. Michaelis. 2009. A framework for the integration of gesture and posture recognition using HMM and SVM. In *Proceedings of the 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*. 572–577.
- [60] F. Adib and D. Katabi. 2013. See through walls with WiFi! In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM (SIGCOMM'13)*. 75–86.
- [61] F. Gringoli, M. Schulz, J. Link, and M. Hollick. 2019. Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets. In *Proceedings of the 2019 Annual International Conference on Mobile Computing and Networking (MobiCom'19)*. 21–28.
- [62] D. Zhao, A. HajiRassouliha, E. J. L. P. Tang, M. P. Nash, M. F. P. Nielsen, and A. J. Taberner. 2019. A camera-based system for highly accurate 3D displacement field measurement and contactless force sensing. In *Proceedings of the 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC'19)*. 1–6.
- [63] S. Qiu, H. Zhao, N. Jiang, Y. An, Y. Li, and J. Liu. 2022. Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges. *Information Fusion* 80 (2022), 241–265.
- [64] C. I. Patel, D. Labana, S. Pandya, K. Modi, H. Ghayvat, and M. Awais. 2020. Histogram of oriented gradient-based fusion of features for human action recognition in action video sequences. *Sensors* 20, 24 (2020), 7299.
- [65] Z. Ren, F. Fang, N. Yan, and Y. Wu. 2022. State of the art in defect detection based on machine vision. *International Journal of Precision Engineering and Manufacturing-Green Technology* 9, 2 (2022), 661–691.
- [66] H. Haresamudram, A. Beedu, V. Agrawal, P. L. Grady, I. Essa, J. Hoffman, and T. Plotz. 2020. Masked reconstruction based self-supervision for human activity recognition. In *Proceedings of the 2020 ACM International Symposium on Wearable Computers*. 45–49.
- [67] J. Materzynska, T. Xiao, R. Herzig, H. Xu, X. Wang, and T. Darrell. 2020. Something-Else: Compositional action recognition with spatial-temporal interaction networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1049–1059.
- [68] Z. Liu, H. Zhang, Z. Chen, Z. Wang, and W. Ouyang. 2020. Disentangling and unifying graph convolutions for skeleton-based action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 143–152.
- [69] A. Mujahid, M. J. Awan, A. Yasin, M. A. Mohammed, R. Damasevicius, R. Maskeliunas, and K. H. Abdulkareem. 2021. Real-time hand gesture recognition based on deep learning YOLOv3 model. *Applied Sciences* 11, 9 (2021), 4164.
- [70] Y. Huang, X. Xu, Z. He, Y. Wang, Z. Lu, and F. Shi. 2023. A lightweight road crack and damage detection method using Yolov5s for IoT applications. In *Proceedings of the 2023 IEEE/CIC International Conference on Communications in China (ICCC'23)*. 1–6.
- [71] P. Zhang, C. Lan, W. Zeng, J. Xing, J. Xue, and N. Zheng. 2020. Semantics-guided neural networks for efficient skeleton-based human action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 1112–1121.

- [72] A. M. Ali and A. Abdelhafeez. 2022. DeepHAR-Net: A novel machine intelligence approach for human activity recognition from inertial sensors. *Sustainable Machine Intelligence Journal* 1, 1 (2022), 1–12.
- [73] C. Plizzari, M. Cannici, and M. Matteucci. 2021. Spatial temporal transformer network for skeleton-based action recognition. In *Pattern Recognition: ICPR International Workshops and Challenges*. Lecture Notes in Computer Science, Vol. 12663. Springer, 694–701.
- [74] M. Xu, C. Zhao, D. S. Rojas, A. Thabet, and B. Ghanem. 2020. G-TAD: Sub-graph localization for temporal action detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10156–10165.
- [75] Y. Li, B. Ji, X. Shi, J. Zhang, B. Kang, and L. Wang. 2020. Tea: Temporal excitation and aggregation for action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 909–918.
- [76] Y. F. Song, Z. Zhang, C. Shan, and L. Wang. 2020. Stronger, faster and more explainable: A graph convolutional baseline for skeleton-based action recognition. In *Proceedings of the 28th ACM International Conference on Multimedia*. 1625–1633.
- [77] S. Xu, J. Wang, W. Shou, T. Ngo, A.-M. Sadick, and X. Wang. 2021. Computer vision techniques in construction: A critical review. *Archives of Computational Methods in Engineering* 28 (2021), 3383–3397.
- [78] A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, Y. Liu, E. Topol, J. Dean, and R. Socher. 2021. Deep learning-enabled medical computer vision. *NPJ Digital Medicine* 4, 1 (2021), 5.
- [79] C. C. Angelica, H. Purnama, and F. Purnomo. 2021. Impact of computer vision with deep learning approach in medical imaging diagnosis. In *Proceedings of the 2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI'21)*. 37–41.
- [80] W. Peng, Z.-S. Ye, and N. Chen. 2020. Bayesian deep-learning-based health prognostics toward prognostics uncertainty. *IEEE Transactions on Industrial Electronics* 67, 3 (2020), 2283–2293.
- [81] C. Wang, X. Bai, S. Wang, J. Zhou, and P. Ren. 2019. Multiscale visual attention networks for object detection in VHR remote sensing images. *IEEE Geoscience and Remote Sensing Letters* 16, 2 (2019), 310–314.
- [82] S. Chaudhuri, R. Saha, A. Chatterjee, S. Mookherjee, and D. Sanyal. 2020. Development of a motion sensing system based on visual servoing of an eye-in-hand electrohydraulic parallel manipulator. *IEEE Sensors Journal* 20, 14 (2020), 8108–8116.
- [83] Y. Wang, J. Guang, F. Yichen, H. Bing, Z. Liang, and L. Xuhan. 2023. Research on the development of on-site human activity recognition and supervision tools based on fuzzy algorithms. In *Proceedings of the 2023 IEEE 5th International Conference on Power, Intelligent Computing, and Systems (ICPICS'23)*. 636–641.
- [84] B. Fu, N. Damer, F. Kirchbuchner, and A. Kuijper. 2020. Sensing technology for human activity recognition: A comprehensive survey. *IEEE Access* 8 (2020), 83791–83820.
- [85] J. Yang, X. Chen, H. Zou, D. Wang, and L. Xie. 2022. AutoFI: Towards automatic WiFi human sensing via geometric self-supervised learning. *IEEE Internet of Things Journal* 10, 8 (2022), 7416–7425.
- [86] Y. Zeng, J. Liu, J. Xiong, Z. Liu, D. Wu, and D. Zhang. 2021. Exploring multiple antennas for long-range WiFi sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 4 (2021), Article 190, 30 pages.
- [87] D. Yang, T. Wang, Y. Sun, and Y. Wu. 2018. Doppler shift measurement using complex-valued CSI of WiFi in corridors. In *Proceedings of the 2018 3rd International Conference on Computer and Communication Systems (ICCCS'18)*. 367–371.
- [88] Z. Gao, Y. Gao, S. Wang, D. Li, and Y. Xu. 2021. CRISLoc: Reconstructable CSI fingerprinting for indoor smartphone localization. *IEEE Internet of Things Journal* 8, 5 (2021), 3422–3437.
- [89] I. Galdino, J. C. H. Soto, E. Caballero, V. Ferreira, T. C. Ramos, and C. Albuquerque. 2023. eHealth CSI: A Wi-Fi CSI dataset of human activities. *IEEE Access* 11 (2023), 71003–71012.
- [90] F. Miao, Y. Huang, W. Qian, F. Shi, and Z. Lu. 2023. Intelligent surface defect detection of hot rolled strip steel using YOLOv8 framework. In *Proceedings of the 2023 IEEE 23rd International Conference on Communication Technology (ICCT'23)*. 678–682, DOI: <https://doi.org/10.1109/ICCT59356.2023.10419492>
- [91] M. R. M. Bastwesi, K. Kai, H. Choi, S. Ishida, and Y. Arakawa. 2023. Wi-Nod: Head nodding recognition by Wi-Fi CSI toward communicative support for quadriplegics. In *Proceedings of the 2023 IEEE Wireless Communications and Networking Conference (WCNC'23)*. 1–6.
- [92] Y. Dong, T. Arslan, and Y. Yang. 2022. Real-time NLOS/LOS identification for smartphone-based indoor positioning systems using WiFi RTT and RSS. *IEEE Sensors Journal* 22, 6 (2022), 5199–5209.
- [93] J. Wang and J. Park. 2021. An enhanced indoor positioning algorithm based on fingerprint using fine-grained CSI and RSSI measurements of IEEE 802.11n WLAN. *Sensors* 21, 8 (2021), 2769.
- [94] H. Boudlal, M. Serrhini, and A. Tahiri. 2023. Design and deployment of a practical wireless sensing system for HAR with WiFi CSI in the 5GHz band. In *Proceedings of the 2023 IEEE International Conference on Networking, Sensing, and Control (ICNSC'23)*. 1–6.
- [95] P. F. Moshiri, R. Shahbazian, M. Nabati, and S. A. Ghorashi. 2021. A CSI-based human activity recognition using deep learning. *Sensors* 21, 21 (2021), 7225.

- [96] J. Liu, Y. Zeng, T. Gu, L. Wang, and D. Zhang. 2021. WiPhone: Smartphone-based respiration monitoring using ambient reflected WiFi signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), Article 23, 19 pages.
- [97] X. Zhu, T. Qiu, W. Qu, X. Zhou, Y. Wang, and D. O. Wu. 2021. Path planning for adaptive CSI map construction with A3C in dynamic environments. *IEEE Transactions on Mobile Computing* 22, 5 (2021), 2925–2937.
- [98] C. Zhang and W. Jiao. 2023. ImgFi: A high accuracy and lightweight human activity recognition framework using CSI image. *IEEE Sensors Journal* 23, 18 (2023), 21966–21977.
- [99] J. Zhang and Y. Zhang. 2020. Vehicular localization based on CSI-fingerprint and vector match. *IEEE Transactions on Intelligent Transportation Systems* 22, 12 (2020), 7736–7746.
- [100] X. Zhu, W. Qu, X. Zhou, L. Zhao, Z. Ning, and T. Qiu. 2022. Intelligent fingerprint-based localization scheme using CSI images for Internet of Things. *IEEE Transactions on Network Science and Engineering* 9, 4 (2022), 2378–2391.
- [101] Y. Wang, L. Yao, Y. Wang, and Y. Zhang. 2021. Robust CSI-based human activity recognition with augment few shot learning. *IEEE Sensors Journal* 21, 21 (2021), 24297–24308.
- [102] J. Guo, C. K. Wen, S. Jin, and G. Y. Li. 2022. Overview of deep learning-based CSI feedback in massive MIMO systems. *IEEE Transactions on Communications* 70, 12 (2022), 8017–8045.
- [103] B. A. Alsaify, M. M. Almazari, R. Alazrai, and M. I. Daoud. 2020. A dataset for Wi-Fi-based human activity recognition in line-of-sight and non-line-of-sight indoor environments. *Data in Brief* 33 (2020), 106534.
- [104] Z. He, X. Zhang, Y. Wang, Y. Lin, G. Gui, and H. Gacanin. 2023. A robust CSI-based Wi-Fi passive sensing method using attention mechanism deep learning. *IEEE Internet of Things Journal* 10, 19 (2023), 17490–17499.
- [105] X. Ding, T. Jiang, Y. Zhong, S. Wu, J. Yang, and W. Xue. 2021. Improving WiFi-based human activity recognition with adaptive initial state via one-shot learning. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC'21)*. 1–6.
- [106] Y. Zhang, X. Wang, Y. Wang, and H. Chen. 2022. Human activity recognition across scenes and categories based on CSI. *IEEE Transactions on Mobile Computing* 21, 7 (2022), 2411–2420.
- [107] R. Alazrai, A. Awad, B. A. Alsaify, M. Hababeh, and M. I. Daoud. 2020. A dataset for Wi-Fi-based human-to-human interaction recognition. *Data in Brief* 31 (2020), 105668.
- [108] J. Yang, Y. Liu, Z. Liu, Y. Wu, T. Li, and Y. Yang. 2021. A framework for human activity recognition based on WiFi CSI signal enhancement. *International Journal of Antennas and Propagation* 2021, 1 (2021), 6654752.
- [109] Y. Zhang, Y. Zheng, K. Qian, G. Zhang, Y. Liu, and C. Wu. 2022. Widar3.0: Zero-effort cross-domain gesture recognition with Wi-Fi. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 11 (2022), 8671–8688.
- [110] J. Yang, X. Chen, H. Zou, D. Wang, Q. Xu, and L. Xie. 2022. EfficientFI: Towards large-scale lightweight WiFi sensing via CSI compression. *IEEE Internet of Things Journal* 9, 15 (2022), 13086–13095.
- [111] D. Wang, J. Yang, W. Cui, L. Xie, and S. Sun. 2021. Multimodal CSI based human activity recognition using GANs. *IEEE Internet of Things Journal* 8, 24 (2021), 17345–17355.
- [112] G. Forbes, S. Massie, and S. Craw. 2020. WiFi-based human activity recognition using Raspberry Pi. In *Proceedings of the 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI'20)*. 722–730.
- [113] P. F. Moshiri, R. Shahbazian, M. Nabati, and S. A. Ghorashi. 2021. A CSI-based human activity recognition using deep learning. *Sensors* 21, 21 (2021), 7225–7244.
- [114] S. Xu, Z. R. He, Y. Wang, W. Shi, G. Gui, and T. Ohtsuki. 2022. Cross-person activity recognition method using snapshot ensemble learning. In *Proceedings of the 2022 IEEE 96th Vehicular Technology Conference (VTC'22 Fall)*. 1–5.
- [115] Y. Ma, S. Arshad, S. Muniraju, E. Torkildson, E. Rantala, K. Doppler, and G. Zhou. 2021. Location- and person-independent activity recognition with WiFi, deep neural networks, and reinforcement learning. *ACM Transactions on Internet of Things* 2, 1 (2021), 1–25.
- [116] S. Tan and J. Yang. 2021. Multi-user activity recognition and tracking using commodity WiFi. *arXiv preprint arXiv:2106.00865* (2021).
- [117] N. Damodaran and J. Schäfer. 2019. Device free human activity recognition using WiFi channel state information. In *Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart City (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI'19)*.
- [118] B. A. Alsaify, M. M. Almazari, R. Alazrai, and M. I. Daoud. 2020. A dataset for Wi-Fi-based human activity recognition in line-of-sight and non-line-of-sight indoor environments. *Data in Brief* 33 (2020), 106534.
- [119] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang. 2019. FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), Article 121, 26 pages.
- [120] Y. LeCun, Y. Bengio, and G. Hinton. 2015. Deep learning. *Nature* 521, 7553 (2015), 436–444.
- [121] C. Zhu and W. Sheng. 2011. Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans* 41, 3 (2011), 569–573.

- [122] D. Khan and I. W.-H. Ho. 2021. Deep learning of CSI for efficient device-free human activity recognition. In *Proceedings of the 2021 IEEE 7th World Forum on Internet of Things (WF-IoT'21)*. 19–24.
- [123] J. Wang, G. Gui, T. Ohtsuki, B. Adebisi, H. Gacanin, and H. Sari. 2021. Compressive sampled CSI feedback method based on deep learning for FDD massive MIMO systems. *IEEE Transactions on Communications* 69, 9 (2021), 5873–5885.
- [124] J. Yang, X. Chen, H. Zou, D. Wang, Q. Xu, and L. Xie. 2022. EfficientFi: Toward large-scale lightweight WiFi sensing via CSI compression. *IEEE Internet of Things Journal* 9, 15 (2022), 13086–13095.
- [125] W. Li, M. J. Bocus, C. Tang, R. J. Piechocki, K. Woodbridge, and K. Chetty. 2021. On CSI and passive WiFi radar for opportunistic physical activity recognition. *IEEE Transactions on Wireless Communications* 21, 1 (2021), 607–620.
- [126] H. Gu, J. Yang, G. Gui, and H. Gacanin. 2023. Triplet MatchNet based indoor position method using CSI fingerprint similarity comparison. *IEEE Transactions on Vehicular Technology* 72, 12 (2023), 16905–16910.
- [127] V. K. Singh, P. Kar, S. A. M. Sohini, M. Rangaiah, S. Chakraborty, and M. Maity. 2023. Monitoring engagement in online classes through WiFi CSI. In *Proceedings of the 2023 15th International Conference on Communication Systems and Networks (COMSNETS'23)*. 462–465.
- [128] J. Zhang, F. Wu, B. Wei, Q. Zhang, H. Huang, and S. W. Shah. 2021. Data augmentation and dense-LSTM for human activity recognition using WiFi signal. *IEEE Internet of Things Journal* 8, 6 (2021), 4628–4641.
- [129] F. Zhang, D. Zhang, J. Xiong, H. Wang, K. Niu, B. Jin, and Y. Wang. 2018. From Fresnel diffraction model to fine-grained human respiration sensing with commodity WiFi devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), Article 53, 23 pages.
- [130] M. G. Moghaddam, A. A. N. Shirehjini, and S. Shirmohammadi. 2022. A WiFi-based system for recognizing fine-grained multiple-subject human activities. In *Proceedings of the 2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC'22)*. 1–6.
- [131] B. Yu, Y. Wang, K. Niu, Y. Zeng, T. Gu, and L. Wang. 2021. WiFi-Sleep: Sleep stage monitoring using commodity WiFi devices. *IEEE Internet of Things Journal* 8, 18 (2021), 13900–13913.
- [132] W. Liu, Q. Cheng, Z. Deng, H. Chen, X. Fu, and X. Zheng. 2019. Survey on CSI-based indoor positioning systems and recent advances. In *Proceedings of the 2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN'19)*. 1–8.
- [133] M. Gao, X. Yang, and M. Zhou. 2021. Curve fitting based CSI compression and reconstruction for indoor positioning. In *Proceedings of the 2021 IEEE 94th Vehicular Technology Conference (VTC'21 Fall)*. 1–5.
- [134] M. Bullmann, T. Fetzter, F. Ebner, M. Ebner, F. Deinzer, and M. Grzegorzec. 2020. Comparison of 2.4 GHz WiFi FTM- and RSSI-based indoor positioning methods in realistic scenarios. *Sensors* 20, 16 (2020), 4515.
- [135] T. T. Khanh, V. D. Nguyen, X. Q. Pham, and E.-N. Huh. 2020. Wi-Fi indoor positioning and navigation: A cloudlet-based cloud computing approach. *Human-centric Computing and Information Sciences* 10, 1 (2020), 32.
- [136] T. K. Geok, K. Z. Aung, M. S. Aung, M. T. Soe, A. Abdaziz, C. P. Liew, F. Hossain, C. P. Tso, and W. H. Yong. 2020. Review of indoor positioning: Radio wave technology. *Applied Sciences* 11, 1 (2020), 279.
- [137] Y. Yu, R. Chen, L. Chen, S. Wu, W. Li, and Y. Wu. 2020. Precise 3-D indoor localization based on WiFi FTM and built-in sensors. *IEEE Internet of Things Journal* 7, 12 (2020), 11753–11765.
- [138] D. Wu, Y. Zeng, R. Gao, S. Li, Y. Li, and R. C. Shah. 2021. WiTraj: Robust indoor motion tracking with WiFi signals. *IEEE Transactions on Mobile Computing* 22, 5 (2021), 3062–3078.
- [139] R. Chen, Z. Li, F. Ye, G. Guo, S. Xu, and L. Qian. 2021. Precise indoor positioning based on acoustic ranging in smartphone. *IEEE Transactions on Instrumentation and Measurement* 70 (2021), 1–12.
- [140] S. Sadowski, P. Spachos, and K. N. Plataniotis. 2020. Memoryless techniques and wireless technologies for indoor localization with the Internet of Things. *IEEE Internet of Things Journal* 7, 11 (2020), 10996–11005.
- [141] L. Chen, X. Zhou, F. Chen, L.-L. Yang, and R. Chen. 2021. Carrier phase ranging for indoor positioning with 5G NR signals. *IEEE Internet of Things Journal* 9, 13 (2021), 10908–10919.
- [142] L. Bai, F. Ciravegna, R. Bond, and M. Mulvenna. 2020. A low cost indoor positioning system using Bluetooth Low Energy. *IEEE Access* 8 (2020), 136858–136871.
- [143] Z. Li, X. Zhao, Z. Zhao, and T. Braun. 2021. WiFi-RITA positioning: Enhanced crowdsourcing positioning based on massive noisy user traces. *IEEE Transactions on Wireless Communications* 20, 6 (2021), 3785–3799.
- [144] X. Zhu, W. Qu, T. Qiu, L. Zhao, M. Atiquzzaman, and D. O. Wu. 2020. Indoor intelligent fingerprint-based localization: Principles, approaches and challenges. *IEEE Communications Surveys & Tutorials* 22, 4 (2020), 2634–2657.
- [145] F. Gu, S. Valaee, K. Khoshelham, J. Shang, and R. Zhang. 2020. Landmark graph-based indoor localization. *IEEE Internet of Things Journal* 7, 9 (2020), 8343–8355.
- [146] G. Lin, W. Jiang, S. Xu, X. Zhou, X. Guo, and Y. Zhu. 2023. Human activity recognition using smartphones with WiFi signals. *IEEE Transactions on Human-Machine Systems* 53, 1 (2023), 142–153.
- [147] Z. Hao, J. Niu, X.-C. Dang, and Z. Qiao. 2021. Wi-Piga: A personnel-independent method for actions recognition based on Wi-Fi. In *Proceedings of the 2021 7th International Conference on Big Data Computing and Communications (BigCom'21)*. 52–59.

- [148] R. Alazrai, M. Hababeh, B. A. Alsaify, M. Z. Ali, and M. I. Daoud. 2020. An end-to-end deep learning framework for recognizing human-to-human interactions using Wi-Fi signals. *IEEE Access* 8 (2020), 7695–7710.
- [149] Z. Tang, A. Zhu, Z. Wang, K. Jiang, Y. Li, and F. Hu. 2020. Human behavior recognition based on WiFi channel state information. In *Proceedings of the 2020 Chinese Automation Congress (CAC'20)*. 1157–1162.
- [150] D. Kollias and S. Zafeiriou. 2021. Exploiting multi-CNN features in CNN-RNN based dimensional emotion recognition on the OMG in-the-wild dataset. *IEEE Transactions on Affective Computing* 12, 3 (2021), 595–606.
- [151] X. Shu, L. Zhang, Y. Sun, and J. Tang. 2021. Host-Parasite: Graph LSTM-in-LSTM for Group Activity Recognition. *IEEE Transactions on Neural Networks and Learning Systems* 32, 2 (2021), 663–674.
- [152] R. C. Staudemeyer and E. R. Morris. 2019. Understanding LSTM: A tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586* (2019).
- [153] Y. Yu, X. Si, C. Hu, and J. Zhang. 2019. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation* 31, 7 (2019), 1235–1270.
- [154] K. Smagulova and A. P. James. 2019. A survey on LSTM memristive neural network architectures and applications. *European Physical Journal Special Topics* 228, 10 (2019), 2313–2324.
- [155] K. Khalil, O. Eldash, A. Kumar, and M. Bayoumi. 2019. Economic LSTM approach for recurrent neural networks. *IEEE Transactions on Circuits and Systems II: Express Briefs* 66, 11 (2019), 1885–1889.
- [156] Z. Wang, X. Su, and Z. Ding. 2020. Long-term traffic prediction based on LSTM encoder-decoder architecture. *IEEE Transactions on Intelligent Transportation Systems* 22, 10 (2020), 6561–6571.
- [157] C. G. Huang, H. Z. Huang, and Y. F. Li. 2019. A bidirectional LSTM prognostics method under multiple operational conditions. *IEEE Transactions on Industrial Electronics* 66, 11 (2019), 8792–8802.
- [158] M. Ma and Z. Mao. 2020. Deep-convolution-based LSTM network for remaining useful life prediction. *IEEE Transactions on Industrial Informatics* 17, 3 (2020), 1658–1667.
- [159] Z. Chen, L. Zhang, C. Jiang, Z. Cao, and W. Cui. 2019. WiFi CSI based passive human activity recognition using attention based BLSTM. *IEEE Transactions on Mobile Computing* 18, 11 (2019), 2714–2724.
- [160] R. Kadir, R. Saha, M. A. Awal, and M. I. Kadir. 2021. Deep bidirectional LSTM network learning-aided OFDMA downlink and SC-FDMA uplink. In *Proceedings of the 2021 International Conference on Electronics, Communications, and Information Technology (ICECIT'21)*. 1–4.
- [161] P. Kaushik, A. Gupta, P. P. Roy, and D. P. Dogra. 2019. EEG-based age and gender prediction using deep BLSTM-LSTM network model. *IEEE Sensors Journal* 19, 7 (2019), 2634–2641.
- [162] T. Hayashi, S. Watanabe, T. Toda, T. Hori, J. Le Roux, and K. Takeda. 2017. Duration-controlled LSTM for polyphonic sound event detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 25, 11 (2017), 2059–2070.
- [163] F. Qiao, B. Li, Y. Zhang, H. Guo, W. Li, and S. Zhou. 2020. A fast and accurate recognition of ECG signals based on ELM-LRF and BLSTM algorithm. *IEEE Access* 8 (2020), 71189–71198.
- [164] S. Gupta, A. D. Dileep, and T. A. Gonsalves. 2020. Online sparse BLSTM models for resource usage prediction in cloud datacentres. *IEEE Transactions on Network and Service Management* 17, 4 (2020), 2335–2349.
- [165] Y. Jiang, W. Chen, M. Liu, Y. Wang, and E. Meijering. 2021. 3D neuron microscopy image segmentation via the ray-shooting model and a DC-BLSTM network. *IEEE Transactions on Medical Imaging* 40, 1 (2021), 26–37.
- [166] Q. Niu and X. Li. 2020. A high-performance web attack detection method based on CNN-GRU model. In *Proceedings of the 2020 IEEE 4th Information Technology, Networking, Electronic, and Automation Control Conference (ITNEC'20)*. 804–808.
- [167] A. Shewalkar, D. Nyavanandi, and S. A. Ludwig. 2019. Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU. *Journal of Artificial Intelligence and Soft Computing Research* 9, 4 (2019), 235–245.
- [168] M. Xia, H. Shao, X. Ma, and C. W. de Silva. 2021. A stacked GRU-RNN-based approach for predicting renewable energy and electricity load for smart grid operation. *IEEE Transactions on Industrial Informatics* 17, 10 (2021), 7050–7059.
- [169] Z. Que, X. Jin, and Z. Xu. 2021. Remaining useful life prediction for bearings based on a gated recurrent unit. *IEEE Transactions on Instrumentation and Measurement* 70 (2021), 1–11.
- [170] H. Nana, D. Lei, W. Lijie, H. Ying, D. Zhongjian, and W. Bo. 2019. Short-term wind speed prediction based on CNN-GRU model. In *Proceedings of the 2019 Chinese Control and Decision Conference (CCDC'19)*. 2243–2247.
- [171] Y. Zhang, Y. Zheng, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang. 2022. Widar3.0: Zero-effort cross-domain gesture recognition with WiFi. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 11 (2022), 8671–8688.
- [172] F. He, T. Liu, and D. Tao. 2020. Why ResNet works? Residuals generalize. *IEEE Transactions on Neural Networks and Learning Systems* 31, 12 (2020), 5349–5362.
- [173] X. Peng, F.-Y. Wang, and L. Li. 2023. Towards better generalization of deep neural networks via non-typicality sampling scheme. *IEEE Transactions on Neural Networks and Learning Systems* 34, 10 (2023), 7910–7920.

- [174] J. Zheng, C. Lu, C. Hao, D. Chen, and D. Guo. 2021. Improving the generalization ability of deep neural networks for cross-domain visual recognition. *IEEE Transactions on Cognitive and Developmental Systems* 13, 3 (2021), 607–620.
- [175] Y. Liu, H. Zhao, J. Gu, Y. Qiao, and C. Dong. 2023. Evaluating the generalization ability of super-resolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 12 (2023), 14497–14513.
- [176] D. Wang, J. Zeng, and S.-B. Lin. 2021. Random sketching for neural networks with ReLU. *IEEE Transactions on Neural Networks and Learning Systems* 32, 2 (2021), 748–762.
- [177] M. Zhang and W. Shi. 2020. A feature difference convolutional neural network-based change detection method. *IEEE Transactions on Geoscience and Remote Sensing* 58, 10 (2020), 7232–7246.
- [178] Q. Wang, J. Wan, and X. Li. 2019. Robust hierarchical deep learning for vehicular management. *IEEE Transactions on Vehicular Technology* 68, 5 (2019), 4148–4156.
- [179] Y. Tian, X. Cao, K. Huang, C. Fei, Z. Zheng, and X. Ji. 2022. Learning to drive like human beings: A method based on deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems* 23, 7 (2022), 6357–6367.
- [180] A. Bouguettaya, H. Zarzour, A. Kechida, and A. M. Taberkit. 2022. Vehicle detection from UAV imagery with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems* 33, 11 (2022), 6047–6067.
- [181] H. Guo, S. Li, K. Qi, Y. Guo, and Z. Xu. 2020. Learning automata based competition scheme to train deep neural networks. *IEEE Transactions on Emerging Topics in Computational Intelligence* 4, 2 (2020), 151–158.
- [182] X. Zhang, X. Chen, Y. Wang, G. Gui, B. Adebisi, and H. Sari. 2023. Lightweight automatic modulation classification via progressive differentiable architecture search. *IEEE Transactions on Cognitive Communications and Networking* 9, 6 (2023), 1519–1530.
- [183] Y. Wang, G. Gui, Y. Lin, H.-C. Wu, C. Yuen, and F. Adachi. 2022. Few-shot specific emitter identification via deep metric ensemble learning. *IEEE Internet of Things Journal* 9, 24 (2022), 24980–24994.
- [184] X. Zhang, L. Hao, G. Gui, Y. Wang, B. Adebisi, and H. Sari. 2024. An automatic and efficient malware traffic classification method for secure Internet of Things. *IEEE Internet of Things Journal* 11, 5 (2024), 8448–8458.
- [185] X. Ding, T. Jiang, Y. Zhong, S. Wu, J. Yang, and W. Xue. 2021. Improving WiFi-based human activity recognition with adaptive initial state via one-shot learning. In *Proceedings of the 2021 IEEE Wireless Communications and Networking Conference (WCNC'21)*. 1–6.
- [186] X. Fu, C. Wang, and S. Li. 2024. Wi-SensiNet: Through-wall human activity recognition based on WiFi sensing. *IEEE Access*. Forthcoming.
- [187] F. S. Abuhoureyah, W. Y. Chiew, and M. S. Zitouni. 2024. WIFI based human activity recognition using multi-head adaptive attention mechanism. *Journal of Intelligent & Fuzzy Systems*. Prepress.
- [188] S. K. C. Valaboju, B. Venkateswarlu, S. Vemuri, V. N. S. S. Pasupuleti, V. B. Burra, and P. Tumuluru. 2024. Activity identification via Wi-Fi channel state information with neural networks. *Research Square*. Forthcoming.
- [189] A. G. Padua, C. C. Corona, G. E. S. Miranda, R. C. C. Surara, and J. J. Reyes. 2024. People counting and gesture recognition system using indoor Wi-Fi sensing. In *Proceedings of the International Conference on Information and Communication Technology (IAICT'24)*. IEEE.
- [190] F. Abuhoureyah, K. S. Sim, and Y. C. Wong. 2024. Multi-user human activity recognition through adaptive location-independent WiFi signal characteristics. *IEEE Access*. Forthcoming.
- [191] J. Strohmayer and M. Kampel. 2024. Data augmentation techniques for cross-domain WiFi CSI-based human activity recognition. *arXiv preprint arXiv:2401.00964* (2024).
- [192] W. Cui, K. Wu, M. Wu, X. Li, and Z. Chen. 2024. WiCAU: Comprehensive partial adaptation with uncertainty-aware for WiFi-based cross-environment activity recognition. *IEEE Transactions on Instrumentation and Measurement* 73 (2024), Article 5503110, 10 pages.
- [193] L. Sheng, Y. Chen, S. Ning, S. Wang, B. Lian, and Z. Wei. 2023. DA-HAR: Dual adversarial network for environment-independent WiFi human activity recognition. *Pervasive and Mobile Computing* 101 (2023), 101850.
- [194] J. Liu, F. Wang, Z. Li, R. Xiong, T. Mi, and R. C. Qiu. 2023. A Wi-Fi signal-based human activity recognition using high-dimensional factor models. *arXiv preprint arXiv:2311.05921* (2023).
- [195] F. Deng, E. Jovanov, H. Song, W. Shi, Y. Zhang, and W. Xu. 2024. WiLDAR: WiFi signal-based lightweight deep learning model for human activity recognition. *IEEE Internet of Things Journal* 11, 2 (2024) 2899–2908.
- [196] W. Liang, R. Tang, S. Jiang, R. Wang, Y. Zhao, and C.-Z. Xu. 2024. LiWi-HAR: Lightweight WiFi-based human activity recognition using distributed AIoT. *IEEE Internet of Things Journal* 11, 1 (2024), 597–611.
- [197] R. Alizadeh, Y. Savaria, and C. Nerguizian. 2024. Characterization and selection of WiFi channel state information features for human activity detection in a smart public transportation system. *IEEE Open Journal of Intelligent Transportation Systems* 5 (2024), 55–69.
- [198] D. Gong, K. Liu, D. Pei, H. Zhang, S. Zhang, and M. Chen. 2024. Wi-Watch: Wi-Fi-based vigilant-activity recognition for ship bridge watchkeeping officers. *IEEE Transactions on Instrumentation and Measurement* 73 (2024), Article 9503717, 17 pages.

- [199] D. Wang, J. Yang, W. Cui, L. Xie, and S. Sun. 2024. AirFi: Empowering WiFi-based passive human gesture recognition to unseen environment via domain generalization. *IEEE Transactions on Mobile Computing* 23, 2 (2024), 1156–1168.
- [200] L.-H. Shen, A.-H. Hsiao, F.-Y. Chu, and K.-T. Feng. 2024. Time-selective RNN for device-free multiroom human presence detection using WiFi CSI. *IEEE Transactions on Instrumentation and Measurement* 73 (2024), Article 2505817, 17 pages.
- [201] G. Yin, J. Zhang, G. Shen, and Y. Chen. 2024. FewSense: Towards a scalable and cross-domain WiFi sensing system using few-shot learning. *IEEE Transactions on Mobile Computing* 23, 1 (2024), 453–468.
- [202] B. Yan, W. Cheng, G. Huang, Z. S. Zhu, and X. Gao. 2021. Activity recognition using the joint of Wi-Fi 2.4G and 5G frequency bands. In *Proceedings of the 2021 IEEE 21st International Conference on Communication Technology (ICCT'21)*. 1266–1270.
- [203] H. Ambalkar, X. Wang, and S. Mao. 2021. Adversarial human activity recognition using Wi-Fi CSI. In *Proceedings of the 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE'21)*. 1–5.
- [204] J. Ding, Y. Wang, H. Si, S. Gao, and J. Xing. 2022. Multimodal fusion-AdaBoost based activity recognition for smart home on WiFi platform. *IEEE Sensors Journal* 22, 5 (2022), 4661–4674.
- [205] P. Huang, X. Zhang, S. Yu, and L. Guo. 2022. IS-WARS: Intelligent and stealthy adversarial attack to WiFi-based human activity recognition systems. *IEEE Transactions on Dependable and Secure Computing* 19, 6 (2022), 3899–3912.
- [206] Y. Zhang, Q. Liu, Y. Wang, and G. Yu. 2022. CSI-based location-independent human activity recognition using feature fusion. *IEEE Transactions on Instrumentation and Measurement* 71 (2022), Article 5503312, 12 pages.
- [207] S. Zaravashan, S. Arefizadeh, and S. Torabi. 2023. Efficient sub-carrier relationship extraction for human activity recognition via EEGNet in wireless sensing. In *Proceedings of the 2023 13th International Conference on Computer and Knowledge Engineering (ICCKE'23)*. 259–264.
- [208] H. Zou, J. Yang, H. P. Das, H. Liu, Y. Zhou, and C. J. Spanos. 2019. WiFi and vision multimodal learning for accurate and robust device-free human activity recognition. In *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'19)*. 1–5.
- [209] Y. Hao, Z. Shi, and Y. Liu. 2020. A wireless-vision dataset for privacy-preserving human activity recognition. In *Proceedings of the 2020 4th International Conference on Multimedia Computing, Networking, and Applications (MCNA'20)*. 97–105.
- [210] M. H. Kefayati, V. Pourahmadi, and H. Aghaeinia. 2020. Wi2Vi: Generating video frames from WiFi CSI samples. *IEEE Sensors Journal* 20, 19 (2020), 11463–11473.
- [211] F. Wang, S. Zhou, S. Panev, J. Han, and D. Huang. 2019. Person-in-WiFi: Fine-grained person perception using WiFi. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 5452–5461.
- [212] J. Yang, H. Zou, and L. Xie. 2022. RobustSense: Defending adversarial attack for secure device-free human activity recognition. *arXiv preprint arXiv:2204.01560* (2022).
- [213] J. Yang, X. Chen, H. Zou, D. Wang, Q. Xu, and L. Xie. 2022. EfficientFi: Towards large-scale lightweight WiFi sensing via CSI compression. *IEEE Internet of Things Journal* 9, 15 (2022), 13086–13095.
- [214] S. Zhou, W. Zhang, D. Peng, Y. Liu, X. Liao, and H. Jiang. 2020. Adversarial WiFi sensing for privacy preservation of human behaviors. *IEEE Communications Letters* 24, 2 (2020), 259–263.

Received 14 July 2024; revised 9 October 2024; accepted 4 November 2024