

CCS-Fi: Widening Wi-Fi Sensing Bandwidth via Compressive Channel Sampling

Xin Li¹ Hongbo Wang^{2,1} Jingzhi Hu¹ Zhe Chen³ Zhiping Jiang⁴ Jun Luo¹

¹College of Computing and Data Science, Nanyang Technological University (NTU), Singapore

²Collaborative Initiative, Interdisciplinary Graduate Programme, Nanyang Technological University (NTU), Singapore

³Institute of Space Internet and School of Computer Science, Fudan University, China

⁴School of Computer Science and Technology, Xidian University, China

{l.xin, hongbo001, jingzhi.hu, junluo}@ntu.edu.sg, zhechen@fudan.edu.cn, zpj@xidian.edu.cn

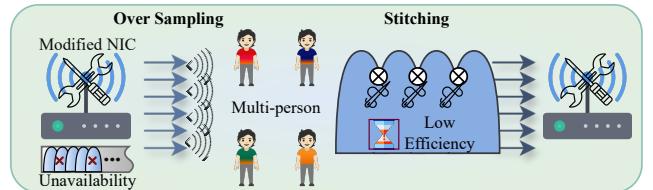
Abstract—Enabling *multi-person differentiation* is crucial for the wide adoption of Wi-Fi sensing, so it is imperative for Wi-Fi sensing to gain GHz-level bandwidth and thus to achieve sufficient *spatial resolution*. Whereas stitching wide bandwidth leveraging *continuous channel samples* appears to be plausible, it is inefficient (if not impossible) in both time and frequency domains. Fortunately, as physical phenomena to be sensed are often *sparse*, acquiring GHz-bandwidth from *sparse channel samples* can be feasible. To this end, we propose CCS-Fi as a novel scheme to widen Wi-Fi sensing bandwidth, exploiting sparse and irregular channel samples. We start with establishing a compressive sensing framework to analyze the potential of realizing GHz sensing capability with only sparse channel samples. Then we propose a model-driven deep learning strategy to implement the sparse recovery process, aiming to derive radar-like channel response as direct output while overcoming the impossibility of obtaining labels for training. Through comprehensive experimental evaluations, we demonstrate that CCS-Fi achieves centimeter-level resolution, effectively enabling indoor multi-person sensing.

Index Terms—Wi-Fi sensing, multi-person localization, compressive sensing, spatial resolution.

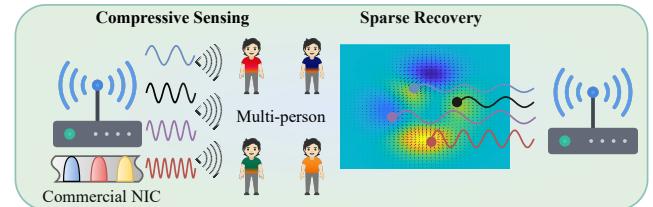
I. INTRODUCTION

Due to its ubiquitous and easily accessible presence, Wi-Fi is no longer regarded merely as a conventional communication medium by academia and industry. Rather, significant efforts have been directed towards harnessing its potential for enabling sensing applications [1]–[6]. Enabled by *channel state information* (CSI) [7], Wi-Fi sensing demonstrates satisfactory performance across diverse applications in single-person scenarios [8]–[12]. However, *how to achieve multi-person fine-grained sensing* remains an unsolved challenge hindering the practical adoption of Wi-Fi sensing. As multi-person sensing demands decimeter-level resolution and sufficient number of differentiable persons (a form of *physical diversity*), it has to be traded for by another form of physical diversity in sensing. Existing approaches relying on either links [13]–[15] or antennas [16]–[18] diversity can hardly be extended beyond differentiating very few persons, significantly restricting their applicability. Therefore, the only feasible option left is the diversity in time-frequency domain.

Meanwhile, exploring time-frequency diversity is confined to *continuous channel stitching* [19], [20]. However, it requires an extensive number of continuous channel samples, posing



(a) Continuous channel stitching framework.



(b) Compressive channel sampling framework.

Fig. 1: The continuous channel stitching framework demanding a large number of continuous channel samples has proven impractical. Instead, a compressive channel sampling framework holds promise for widening Wi-Fi sensing bandwidth.

three main issues, as illustrated in Fig. 1(a). Firstly, Wi-Fi channels may be occupied by other devices, making it impractical to consistently rely on a large number of channels, even if the continuity requirement can be lifted [21]. Secondly, as significant overlap between adjacent channels is necessary to counteract the boundary effect, channel stitching entails extremely heavy over-sampling and subsequent computation cost, making its practicality highly questionable. Last but not least, commercial network interface cards (NICs) allows for overlap between neighboring channels only in the 2.4 GHz band with less than 100 MHz total bandwidth, rendering GHz-bandwidth surely unachievable for sensing purpose even with the aforementioned efforts. Therefore, an efficient and effective framework has to be in place to widen the bandwidth of Wi-Fi sensing system.

Fortunately, according to *compressive sensing* theory, physical phenomena can be recovered with a limited number of samples if they are sparse [22]. This indicates the potential to estimate subject information from significantly fewer and arbitrary channel samples [23], providing three advantages.

Firstly, channel occupancy will no longer be a concern as each channel is not essential. Secondly, relying on sparse channel samples eliminates the need for excessive stitching due to boundary effects, exhibiting high computational efficiency. Finally, since overlap between channel samples is not required, the entire channels offered by commercial NICs, spanning up to 4.7 GHz, are available, as illustrated in Fig. 1(b). Although compressive sensing is a plausible strategy, sparse recovery from discrete samples to widened effective bandwidth of Wi-Fi remains a great challenge. Specifically, given the sparsity of sensed information, fully recovering continuous channel samples is neither necessary nor feasible, thus leaving the recovery representation for multi-person sensing an open issue. Furthermore, additional unknown parameters introduced by compressive sampling in Wi-Fi render conventional algorithms ineffective for extracting information from channel samples.

To address these challenges and fully leverage the advantages of compressive sensing, we propose a compressive channel sampling framework for Wi-Fi sensing, termed CCS-Fi. Initially, the inherent sparsity of information in Wi-Fi channel samples is analyzed to establish the compressive sensing model. Based on this foundation, a deep learning network is developed to resolve the underdetermined parameter estimation problem. Subsequently, a model-driven training strategy is designed to guide the network to filter out interference and recover sparse channel samples in radar-like spectra. Finally, we conduct thorough experiments on the prototype of CCS-Fi to assess its performance. In summary, the contributions of this paper are shown as follows:

- We propose a novel CCS-Fi framework for widening the Wi-Fi effective sensing bandwidth to GHz level.
- We construct a compressive sensing model to analyze the sensing capability of sparse channel samples, laying the theoretical foundation for information recovery.
- We design a model-driven training strategy focused on leveraging the meticulously crafted dataset to enable the neural network to recover radar-like spectra from sparse channel samples.
- We thoroughly evaluate CCS-Fi, demonstrating its performance and influencing factors.

The rest of this paper is organized as follows: Section II introduces the fundamentals of Wi-Fi sensing and outlines the background and motivation. Section III presents the compressive sensing model for sparse Wi-Fi channel samples. Section IV introduces CCS-Fi framework and delves into the deep learning strategy and data augmentation algorithms developed for parameter estimation from discrete channel samples. Section V showcases the prototype of the proposed CCS-Fi framework along with comprehensive evaluation results. Section VI discusses related work. Finally, the conclusion is presented in Section VII.

II. BACKGROUND AND MOTIVATION

In this section, we first establish the foundational model for Wi-Fi sensing and analyze the influence of bandwidth on

spatial resolution. Further, we explain the necessity of oversampling for the continuous channel stitching and elucidate the constraints it faces. Finally, we illustrate both the potential and challenges of compressive sensing through observations of spectra estimation from sparse channel samples.

A. Wi-Fi Sensing Fundamental

The k -th ($k \in \{1, 2, \dots, K\}$) path in a Wi-Fi sensing system can be described by its time of flight (ToF) τ_k and angle of arrival (AoA) θ_k . Therefore, the channel gain of the s -th ($s \in \{0, 1, \dots, S-1\}$) subcarrier at the Rx can be represented as:

$$\begin{aligned} h_{s,n} &= \sum_{k=1}^K \alpha_{s,n,k} \cdot u_k \cdot p_s^{\tau_k} \cdot p_n^{\theta_k} + \epsilon \\ &= \sum_{k=1}^K \alpha_{s,n,k} e^{-i2\pi f_c \tau_k} e^{-i2\pi f_b \tau_k} \left(e^{-i2\pi(n-1)d \cos \theta_k f_c / c} \right) + \epsilon, \end{aligned} \quad (1)$$

where $n \in \{1, 2, \dots, N\}$ represents the antenna index of Rx, assuming that antennas in the system are uniformly arranged with a distance of d . Constant α denotes the amplitude attenuation, f_c and f_b represent the channel center frequency and subcarrier bandwidth, respectively. Constant c is the speed of light and ϵ indicates noise.

According to [24], the spatial resolution of Wi-Fi sensing systems is $\Delta D = c/B$, where B represents the signal bandwidth. As $\Delta\tau = 1/B$ is the temporal resolution, expanding the bandwidth for ΔD essentially enhances the resolution of ToF. Meanwhile, a wide Wi-Fi bandwidth indicates the presence of more subcarriers in the signal, thereby also enhancing the resolution of AoA [25]. Thus, expanding the effective bandwidth of Wi-Fi systems holds great significance in comprehensively enhancing the performance of multi-person sensing.

B. Continuous Channel Stitching

To widen the Wi-Fi sensing bandwidth, ToneTrack [19] and Splicer [20] attempt to stitch continuous channel samples; they both require an overlap between adjacent channel samples to align inconsistent boundaries. Accordingly, we perform continuous channel scanning in the 2.4 GHz band, where standard channels overlap [26], aiming to differentiate two subjects with a separation of 3.8 m. We initially stitch the continuous channel samples, as depicted in Fig. 2. Subsequently, the MULTiple SIgnal Classification (MUSIC) [27], [28] is optionally selected from various feasible algorithms to estimate subject information from the stitched signals (details are provided in Section III-B), as shown in Fig. 3.

According to Section II-A, Wi-Fi signals with 80MHz bandwidth can provide sufficient spatial resolution to differentiate the two subjects. We initially stitch four 20 MHz channel samples. However, both the stitched phase and magnitude exhibit notable fluctuations, as shown in Fig. 2. The information estimated by the signal stitched from 4 channel samples, depicted in Fig. 3(a), unsurprisingly fails to differentiate the two subjects. After observing persistent fluctuations when increasing the number of channel samples to 8, we gradually

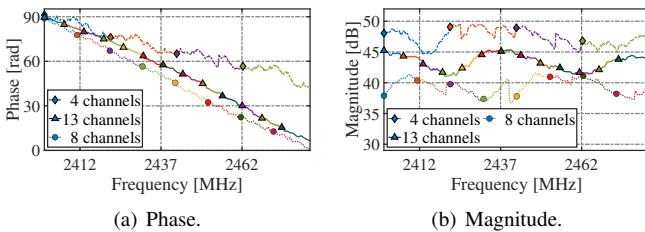


Fig. 2: Results of continuous channel stitching.

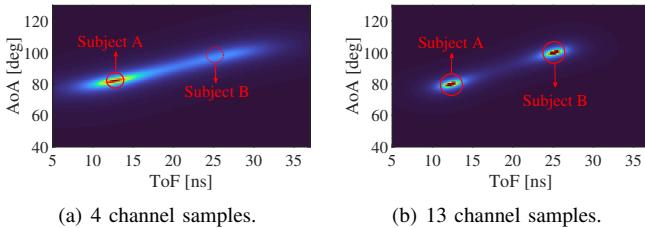


Fig. 3: ToF-AoA spectra of 4 and 13 channel samples.

increase it to 13, at which point the stitched phase and magnitude become significantly smoother due to improved calibration. The estimated ToF-AoA spectrum from them, as shown in Fig. 3(b), successfully differentiates two subjects.

This result indicates that continuous channel stitching demands a much larger number of channel samples than theoretically necessary to achieve satisfactory spatial resolution. Given that achieving an effective bandwidth of only 80 MHz requires 13 channel samples, widening it beyond 3 GHz for centimeter-level resolution demands approximately 500 samples. This not only prevents the sampling efficiency from meeting the channel coherence time [21] but also leads to a waste of spectrum resources. Furthermore, according to our evaluations performed on a MacBook Pro equipped with the M1 Pro chip, each stitching process takes approximately 10ms. Thus, stitching 500 channel samples takes approximately 5 s, significantly surpassing the budget allowed by Wi-Fi real-time sensing. Besides, the need for such a large number of channels to be available is impractical. Therefore, it is imperative to adopt a novel theoretical framework for efficiently widening the Wi-Fi sensing bandwidth.

C. Compressive Channel Sampling

The limitations of continuous channel stitching prompt us to explore the possibility of utilizing discrete and irregular channel samples to widen the Wi-Fi sensing bandwidth. Fortunately, physical phenomena sensed in real life are typically sparse [29], inspiring us that compressive sensing holds promise. To validate the feasibility of compressive sensing, we simulate three sparse 20 MHz channel samples spanning a frequency range of 80 MHz based on Eqn. (1) and estimate the ToF-AoA spectra with the MUSIC algorithm. The results are illustrated on the left side of Fig. 4. The highlighted regions reveal that peaks appear only at very few positions in each spectrum, confirming the sparsity of subject information. However, due to the varying carrier frequencies of the channel

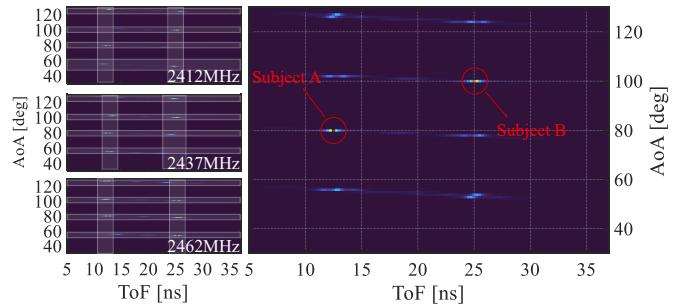


Fig. 4: Sensing with compressive channel sampling. Left: ToF-AoA spectra of channel samples with different frequencies, right: synthesized ToF-AoA spectrum.

samples, the ToF-AoA spectra exhibit different patterns of periodic aliasing, making it impossible to determine the locations of subjects based on a single channel sample. Therefore, we synthesize the three ToF-AoA spectra (details are shown in Section III-B), as illustrated on the right side of Fig. 4. We can utilize maximum likelihood estimation to identify the points with the highest occurrence of peak values, i.e., the maximum values of the ToF-AoA spectrum, representing the subjects' information. This validates the potential of sparse channel samples for widening the Wi-Fi sensing bandwidth.

However, the ToF-AoA spectrum from real-life sparse channel samples fails to reveal subject information, as shown in Fig. 5(a). In addition to the influence of noise, the primary reason for the obscurity is the discrete channel sampling, wherein the Wi-Fi NIC needs to restart with each channel switch, introducing random offsets that cannot be completely eliminated. Among them, the carrier phase offset (CPO) has a particularly severe impact on Wi-Fi sensing [1]. To illustrate this issue, we control two Wi-Fi NICs to collect 50 samples in a stable environment under two conditions: channel hopping and staying on the same channel, and their phase differences [25], [30] of the received CSIs from the two Rx antennas are shown in Fig. 5(b). The results indicate that, compared to staying on the same channel, the phase difference under the hopping condition shows significant random fluctuations. With each sampling, although diversity information about the subject is acquired, unknown interferences are also introduced, potentially overwhelming the signal. This poses a challenge for traditional signal processing algorithms to estimate subject information from sparse channel samples.

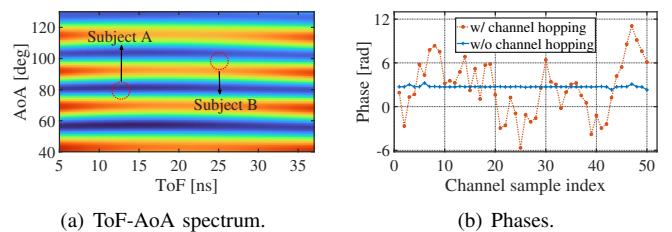


Fig. 5: Analysis of real-life channel samples. The additional interference introduced by channel hopping impedes the estimation of subject information from sparse channel samples.

In theory, deep learning models can fit any function through effective training [31], [32], thus enabling the sparse recovery of CSI signals. Aiming to emulate radar systems [33], [34], we employ a deep learning model to map sparse channel samples to the spectrum. The spectrum also offers several advantages. Firstly, as a high-dimensional data form, the spectrum can describe various quantities of subjects across different bins. Secondly, sparse recovery involves information decompression and reconstruction. The spectrum provides sufficient parameters to encode the CSI phase and amplitude, laying the foundation for various ubiquitous Wi-Fi sensing tasks. Finally, leveraging the bias-variance trade-off [35], the model's performance can be enhanced through distributing the estimated bias into parameters unrelated to subjects. The ToF-AoA spectrum presented in Fig. 4 offers a promising expression, and detailed discussions will be provided in Section III-B and Section IV-C.

III. COMPRESSIVE SENSING FRAMEWORK

In this section, we begin by analyzing the feasibility of compressive sensing with CSI signals. Next, a mathematical model is established for sparse recovery, outlining the transformation of sparse channel samples into the spectrum with subject information to lay the foundation for algorithm design.

A. Compressive Sensing of CSI

In compressive sensing, channel sampling relies on the information content of the signal rather than its bandwidth [22]. The subject information in real-life scenarios tends to be sparse, which allows for fine-granularity sensing with only a small size of signals. To demonstrate it, assume that when the effective bandwidth meets the spatial resolution ΔD , given by $B = c/\Delta D$, the channel sample is $H = \{H_1, H_2, \dots, H_Q\}$, where $H \in \mathbb{C}^Q$ and Q represents the quantity of standard 20 MHz channels, which is proportional to the bandwidth B . We temporarily omit the number of Rx antennas N , as our compressive sensing does not involve it.

To evaluate the feasibility of compressive channel sampling, the raw CSI signal is generally expressed as $H = \Psi X$, where Ψ is an $Q \times Q$ sparse basis matrix for sensing purposes, and $X = f(\tau, \theta)$ serves as the sparse representation of H . If $\|X\|_0 \leq r$, X is referred to as r -sparse and establishes the foundation for recovering H from a low-dimensional measurement Y . The *compressive sensing problem* described in this paper, which aims to achieve sufficient spatial resolution with only a small number of sparse channel samples, can be stated as follows: given measured signal $Y \in \mathbb{C}^M$ with only M discrete standard channels, recover the wideband channel sample H with Q channels, where $M \ll Q$. Suppose Φ is an $M \times Q$ measurement matrix that $Y = \Phi H$, then

$$Y = \Phi H = \Phi \Psi X = \Theta X, \quad (2)$$

where Θ is referred to as the sensing matrix.

According to the Restricted Isometry Property (RIP) [36], if there exists a constant $\delta_r \in (0, 1)$, assume Θ satisfy

$$(1 - \delta_r) \|X\|_2^2 \leq \|\Theta X\|_2^2 \leq (1 + \delta_r) \|X\|_2^2, \quad (3)$$

it is highly probable to accurately reconstruct X with Y , thus obtaining fine-granularity subject information. The sparsity order r of X is adjusted according to the RIP, while satisfying $r \leq C_r M / \log(Q/M)$, where C_r is a constant [36], [37]. Hence, we can conclude that the dimension M of measured channel samples only needs to meet:

$$M \geq r \log(Q/M)/C_r. \quad (4)$$

This implies widening the effective bandwidth to B may only require M channels if the information is sparse enough.

B. Sparse Recovery with CSI

Since sparse channel samples have the potential to provide sufficient subject information, the further task is to investigate how to effectively recover them. The ToF-AoA spectrum, inspired by radar systems, offers a promising approach for transforming sparse channel samples into spectral parameters. Assuming the k -th path signal is $u_k = e^{-i2\pi f_c \tau_k}$, the signal vector is $U = [u_1, u_2, \dots, u_K]^\top$. As shown in Eqn. (1), the signal received by the Rx is influenced by the phase shifts of subcarriers and antennas, denoted as $p_s^{\tau_k}$ and $p_n^{\theta_k}$, respectively. Thus, the steering vector can be expressed as $v(\tau_k, \theta_k) = [p_0^{\tau_k} p_1^{\theta_k}, \dots, p_{S-1}^{\tau_k} p_1^{\theta_k}, \dots, p_0^{\tau_k} p_N^{\theta_k}, \dots, p_{S-1}^{\tau_k} p_N^{\theta_k}]^\top$, which is then integrated into $V(\tau, \theta) = [v(\tau_1, \theta_1), \dots, v(\tau_K, \theta_K)]$ for all paths. The received signal vector $H = VU + \epsilon$ is poised for spectrum estimation. Subsequently,

$$P(\tau, \theta) = g(H), \quad (5)$$

where g represents a mapping operation that can be implemented by various estimation algorithms. Taking the MUSIC algorithm mentioned above as an example, by constructing a matrix E from eigenvectors corresponding to the l smallest eigenvalues of HH^\top , the ToF-AoA spectrum is $P(\tau, \theta) = \frac{1}{v^\top(\tau, \theta) E E^\top v(\tau, \theta)}$, where superscript H denotes the Hermite transpose of the matrix. Since E represents the noise vector subspace, which is orthogonal to the signal vector subspace, the subject information will appear as peaks in $P(\tau, \theta)$ [25].

Spectrum $P(\tau, \theta)$ can be directly applied to continuous subcarriers (as shown in Section II-B). However, for discrete samples, sparse recovery relies on the synthesized spectrum $\sum_j P_j(\tau, \theta)$. Section II-C illustrates that the periodic aliasing in the spectral parameters obtained from the ideal CSI signal poses a challenge to estimating the subject information. This is primarily attributed to the fact that the τ_k and θ_k are:

$$\tau_k = -\frac{\angle(u_k \cdot p_s^{\tau_k})}{2\pi(f_c + sf_b)} \mod \frac{1}{f_c + sf_b}, \quad (6)$$

$$\cos \theta_k = -\frac{\angle h_n^{\tau_k}}{2\pi(n-1)df_c/c} \mod \frac{c}{(n-1)df_c}. \quad (7)$$

The aliasing is not apparent in the ToF dimension of Fig. 4 due to the sufficient number of subcarriers. If $f_c = 2412$ MHz and $d = 0.32$ m, then the modulus of $\cos \theta_k$ (not θ_k itself) is 0.389. Since $\theta \in [0^\circ, 180^\circ]$, meaning $\cos \theta_k \in [-1, 1]$, the spectrum on the left side of Fig. 4 exhibits multiple peaks in the AoA dimension. When $f_c = 2437$ MHz, the

modulus of $\cos \theta_k$ is 0.385, and it is 0.381 for 2462 MHz. The aliasing peaks changes as f_c varies, whereas the true peaks remain unaffected. Synthesizing multiple spectra with different f_c values enables the detection of parameters corresponding to the subjects at the maximum peaks. This provides us with the mathematical model of sparse recovery and radar emulation that deep learning algorithms need to accommodate. It also offers an alternative understanding of resolution: the differences in aliasing peaks caused by adjacent channels are minimal, which prompts a large number of continuous channel samples to induce enough variation in f_c for enhancing spatial resolution. On the other hand, sparse channel samples with a wide span can directly facilitate the recovery of subject information based on high diversity. We will demonstrate how to achieve this goal under real-life scenarios in Section IV.

IV. CCS-Fi ALGORITHM DESIGN

In this section, we first design the framework of CCS-Fi. Building upon this, we devise a background interference generation algorithm for input data to enhance the model's generalizability. Following that, radar-like spectrum outputs are designed as labels to train the neural network. Lastly, we introduce a spectrum enhancement algorithm aimed at improving the performance of the label.

A. CCS-Fi Framework Design

Leveraging sparse channel samples to achieve high spatial resolution is well-established through compressive sensing, yet the uncertain nature of real-life CSI signals impedes the recovery of subject information from them. Therefore, we design a novel CCS-Fi framework, as depicted in Fig. 6, utilizing a model-driven deep learning strategy to recover sparse channel samples into radar-like spectrum outputs. In this framework, our contributions encompass three aspects:

- Enhancing raw CSI data enables the deep learning model to effectively filter out diverse background multipath interferences, thereby attaining cross-domain capability.

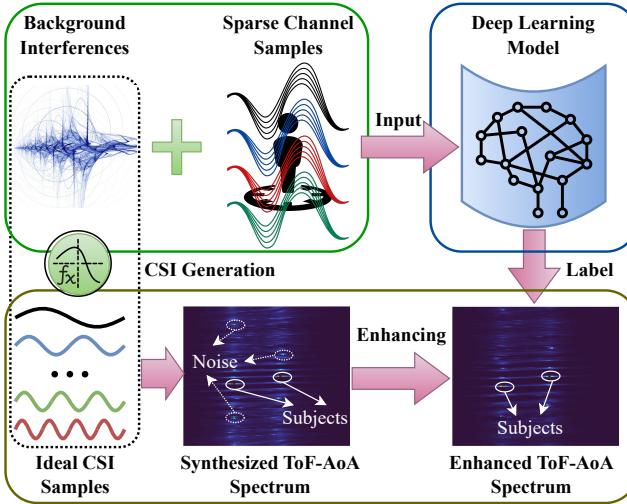


Fig. 6: Framework of CCS-Fi.

- Crafting spectral parameters as ground truth to train the neural network model for emulating radar system output.
- Optimizing the spectrum to highlight subject-related information, thereby improving model performance.

To address these challenges, we begin by generating various potential interference components that may exist in typical indoor environments and then incorporate them into the real-life sparse channel samples. Next, we transform the ground truth parameters into ToF-AoA spectra with diverse peak values associated with subject information and further enhance their significance. Finally, we establish a deep learning model and train it leveraging real-life CSI signals with diverse background interference as inputs, and enhanced ToF-AoA spectra as labels. This training strategy drives the model to perform denoising, calibration, and transformation of sparse channel samples into radar-like spectrum, enabling the utilization of peak value regions for ubiquitous sensing. We will detail the algorithms in the following sections.

B. Background Interference Generation

The Wi-Fi sensing model needs sufficient generalizability to adapt to diverse domains. The issue is typically addressed by training deep learning models on large datasets collected from various environments. However, it comes with high costs and often yields unsatisfactory performance in unknown environments. Besides facilitating precise subject information capture through high spatial resolution, the CCS-Fi framework offers an efficient solution by diversifying the training dataset with model-generated interferences. We analyze various types of background interference in typical indoor environments and estimate their reflection paths. Then, they are transformed into ideal CSI samples and added to the real-life samples to obtain a diverse dataset.

Assuming a background interference $\mathcal{T}_{B_l} = (\tau_{B_l}, \theta_{B_l})$, it can be transformed into an ideal CSI sample $H_{B_l^l}$ according

Algorithm 1: Diversification of Background Interference

Input: Real-life sparse channel sample dataset H_R and potential background interferences \mathcal{T}_B .
Output: Dataset with diverse interferences H .

```

1 Initialization:  $L$  is the size of  $H_R$  and  $H = \emptyset$ .
2 for  $l = 1, \dots, L$  do
3   Randomly select a tuple  $\mathcal{T}_{B_l} = (\tau_{B_l}, \theta_{B_l})$  from  $\mathcal{T}_B$ ;
4   Generate  $\alpha_l$  and  $\epsilon_l$ ;
5   Obtain the parameter vector  $\Omega_l = \{\tau_{B_l}, \theta_{B_l}, \alpha_l, \epsilon_l\}$ ;
6   Extract frequency vector  $f_c^l = \{f_{c_1}^l, f_{c_2}^l, \dots\}$  from  $H_{R_l} = \{H_{R_l^1}, H_{R_l^2}, \dots\}$ ;
7   for  $f_{c_j}^l \in f_c^l$  do
8      $H_{B_l^j} \leftarrow$  Eqn. (1) with parameters  $\Omega_l$  and  $f_{c_j}^l$ ;
9      $H_{l,j} = H_{R_l^j} + H_{B_l^j}$ ;
10    end
11     $H = H \cup H_l$ .
12 end

```

to Eqn. (1). Subsequently, by introducing variations in the amplitude α , noise ϵ , and fundamental frequency f_c , a diverse set of components $H_{B_l} = \{H_{B_l^1}, H_{B_l^2}, \dots\}$ is generated. Finally, a comprehensive dataset of background interference $H_B = \{H_{B_1}, H_{B_2}, \dots\}$ is gathered leveraging the augmentation of all potential interferences, i.e., $\mathcal{T}_B = \{\mathcal{T}_{B_1}, \mathcal{T}_{B_2}, \dots\}$. Moreover, by adjusting the parameters α and ϵ , the size of the generated background interference dataset matches that of the real-life dataset, allowing us to merge them through random mapping. The details are illustrated in Algorithm 1.

C. Model-Driven Label Generation

We aim to transform sparse channel samples into radar-like spectra leveraging a deep neural model. However, as described in Section II-C, real-life channel samples cannot be transformed into informative spectra through existing algorithms, leading to a lack of labels for training. Fortunately, we can generate ideal CSI samples by ground truths, i.e., ToF and AoA tuples, and then utilize the estimation algorithm to transform them into spectra. Subsequently, synthesizing multiple spectra based on compressive sensing yields the radar-like output. Notably, the sparsity of the generated data does not need to match that of the input data. Instead, it can leverage a larger number of channel samples for accurate spectrum recovery. Finally, we meticulously design parameter fluctuations within a tolerance range, thereby acquiring a label dataset with comprehensive information.

Assuming the real-life sparse channel sample used for training is H_{R_l} , where the ground truth is represented as $\mathcal{T}_{R_l} = (\tau_{R_l}, \theta_{R_l})$. The tuple can be transformed into an ideal CSI vector $\hat{H}_{R_l} = \{\hat{H}_{R_l^1}, \hat{H}_{R_l^2}, \dots\}$ with various f_c according to Eqn. (1). Subsequently, they are further estimated into the ToF-AoA spectra $\{P_{l_1}, P_{l_2}, \dots\}$ as per Eqn. (5), and the synthesized spectrum $P_l = \sum_j P_{l_j}$ serves as the radar-like label containing subject information. The ToF-AoA spectrum, defined by a discrete grid of $\Delta = (\Delta\tau, \Delta\theta)$,

Algorithm 2: Spectrum Label Generation

Input: Ground truth dataset \mathcal{T}_R , fundamental frequency vector $f_c = \{f_{c_1}, f_{c_2}, \dots\}$.
Output: Spectral label dataset \mathcal{P} .

```

1 Initialization:  $L$  and  $L_c$  are the sizes of  $\mathcal{T}_R$  and  $f_c$  respectively, and  $\mathcal{P} = \emptyset$ .
2 for  $l = 1, \dots, L$  do
3   Initialization:  $P_l = 0$ ;
4    $\tau_{R_l} \in \{\tau_{R_l} - \Delta\hat{\tau}, \tau_{R_l}, \tau_{R_l} + \Delta\hat{\tau}\}$ ;
5    $\theta_{R_l} \in \{\theta_{R_l} - \Delta\hat{\theta}, \theta_{R_l}, \theta_{R_l} + \Delta\hat{\theta}\}$ ;
6   for  $f_{c_j} \in f_c$  do
7      $H_{l_j} \leftarrow$  Eqn. (1) with  $\tau_{R_l}, \theta_{R_l}, f_{c_j}$ ;
8      $P_{l_j} \leftarrow$  Eqn. (5) with  $H_{l_j}$ ;
9      $P_l = P_l + P_{l_j}$ ;
10  end
11   $\mathcal{P} = \mathcal{P} \cup (P_l / L_c)$ .
12 end
```

Algorithm 3: ToF-AoA Spectrum Enhancement

Input: Original ToF-AoA spectrum P ; parameters κ , γ , $L_o \in \mathbb{O}$, and σ for the enhancement.
Output: Enhanced ToF-AoA spectrum \mathfrak{P} .

```

1 for  $i, j = -(L_o - 1)/2, \dots, (L_o - 1)/2$  do
2    $\Gamma_{i,j} = \frac{i^2 + j^2 - 2\sigma^2}{2\pi\sigma^6} e^{-\frac{i^2 + j^2}{2\sigma^2}}$ .
3 end
4  $\mathfrak{P} = \kappa |P|^\gamma$ ;
5  $\mathfrak{P} = \mathfrak{P} \odot (\mathfrak{P} * \Upsilon)$ , where  $\odot$  and  $*$  represent the Hadamard product and convolution, respectively;
6  $\mathfrak{P} = \frac{\max(P)}{\max(\mathfrak{P})} \mathfrak{P}$ .
```

implies the existence of a tolerance range, enabling the spectra to encode comprehensive information. Assuming a random perturbation $\hat{\Delta}(\Delta\hat{\tau}, \Delta\hat{\theta})$, where $0 < \Delta - \hat{\Delta} < \Delta$, multiple spectra can be generated utilizing $\{\tau - \Delta\hat{\tau}, \tau, \tau + \Delta\hat{\tau}\}$ and $\{\theta - \Delta\hat{\theta}, \theta, \theta + \Delta\hat{\theta}\}$. Randomly selecting one of these spectra as the corresponding output for the input helps improve the generalization of the deep learning model. The details are provided in Algorithm 2.

D. ToF-AoA Spectrum Enhancement

In general, estimation algorithms, e.g., MUSIC, are influenced by the geometric structure of the antenna array, even in the ideal scenario without noise and hardware-related offsets, which means that satisfactory spectral parameters cannot be guaranteed for the subject under all conditions. Additionally, as shown in Fig. 4, Eqn. (6), and Eqn. (7), the synthesized spectrum still exhibits periodic aliasing, which may impact the training of the deep neural model. Therefore, the ToF-AoA spectrum should be further enhanced to ensure high-quality labels. To address this, the gamma transformation is firstly employed to amplify the contrast of spectral parameters. Then, we apply a specially designed operator to smoothly attenuate interference components, thereby further magnifying the differences between subject and background parameters. The details are presented in Algorithm 3.

V. EXPERIMENT AND EVALUATION

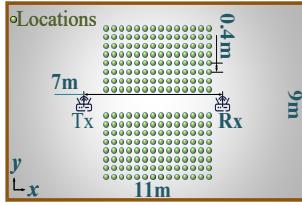
In this section, we first introduce the implementation of CCS-Fi prototype and outline the experiment setup. Next, we conduct a comprehensive evaluation of CCS-Fi, showcasing its overall performance in multi-person fine-granularity sensing. Finally, we analyze various factors that influence the performance of CCS-Fi.

A. Prototype and Experiment Setup

1) *Prototype of CCS-Fi*: CCS-Fi is deployed on two laptops equipped with Intel's AX210 NICs, which support 2.4, 5, and 6 GHz bands, offering a wide range of channel options. We install the PicoScenes platform [38] on the two laptops running the Linux kernel version 5.15.0-60 and collect sparse channel samples with a custom plugin. Next, we construct an end-to-end regression model leveraging the ResNet [39] architecture



(a) A training case.



(b) Measurement locations.



(c) A multi-person case.



(d) Resolution testing scenario.

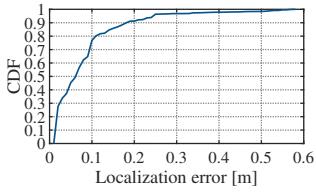
Fig. 7: Experiment scenarios.

based on experience to transform CSI data into ToF-AoA spectra. Finally, we train the model according to the pipeline proposed in Section IV. Notably, our main contribution lies in the pipeline for training the deep neural model. With its support, any model capable of the transformation can be effective, allowing users to employ more complex neural models for further improvement.

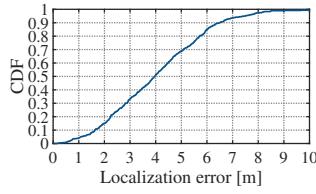
2) *Experiment Setup*: We recruit 8 subjects with varying body sizes and collect sparse channel samples in the rehearsal room (RR), lecture room (LR), reading room (RD), conference room (CR), discussion room (DR), and private office (PO). To verify the cross-domain capability of CCS-Fi, we use the data collected in the RR for training, where the scene is shown in Fig. 7(a). The Tx and Rx (i.e., the two laptops) are placed at opposite sides of the room, and the locations of the subjects are illustrated in Fig. 7(b), where the ground truth is associated with the center of the subject. The data collected in other environments are utilized for testing. Fig. 7(c) illustrates one of the multi-person testing scenarios. To attain sufficient proximity for testing the resolution limit of CCS-Fi, we additionally employ two electric pendulums as subjects, as depicted in Fig. 7(d). The details will be discussed below.

B. Overall Performance

1) *Localization Accuracy*: For localization, CCS-Fi initially collects 20 sparse channel samples and transforms them into a ToF-AoA spectrum. Subsequently, maximum likelihood estimation is employed to identify the points corresponding to

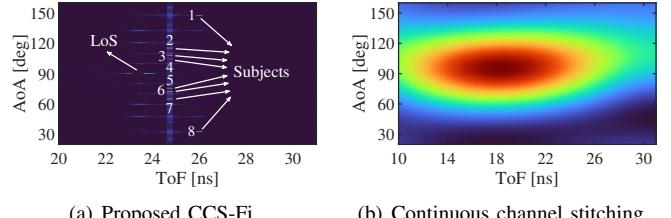


(a) Proposed CCS-Fi.

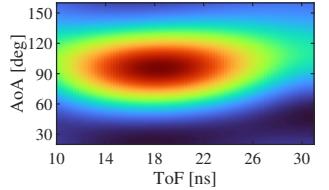


(b) Continuous channel stitching.

Fig. 8: Localization accuracy.



(a) Proposed CCS-Fi.



(b) Continuous channel stitching.

Fig. 9: Localization resolution.

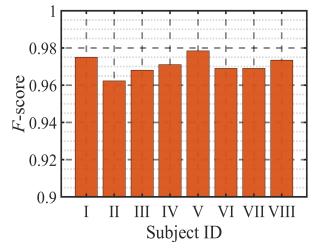
the subjects from peak values. Finally, the polar coordinates described by ToF and AoA are converted into Cartesian coordinates to determine the subjects' positions. The accuracy is defined as the error between the measured values and the ground truth, whose cumulative distribution is illustrated in Fig. 8(a). Continuous channel stitching is also tested as the baseline, as depicted in Fig. 8(b). However, due to its requirement for overlapped channel samples and not involving cross-domain capability, we present its error for a single environment and subject based on the channels within the 2.4 GHz band. CCS-Fi achieves a median error of approximately 6 cm and a 90th percentile error of 18 cm, significantly surpassing the errors of 4.0 m and 6.5 m obtained by the baseline. The accuracy can effectively meet the requirements of typical indoor localization scenarios.

2) *Localization Resolution*: The spatial resolution is the capability to distinguish multiple subjects at different distances. To validate it, we test 8 subjects seated on both sides of the table in the CR using CCS-Fi and baseline methods, respectively. The spectrum obtained by CCS-Fi, as shown in Fig. 9(a), displays multiple distinct peak values, allowing for the identification of all 8 subjects. However, the spectrum from the baseline method, as depicted in Fig. 9(b), does not provide any information related to these subjects. The result proves the successful expansion of effective sensing bandwidth by our CCS-Fi, thereby achieving satisfactory spatial resolution.

3) *Gesture Detection*: We instruct 8 subjects in the CR to simultaneously perform 6 different gestures, including forward-backward (FB), up-down (UD), left-right (LR), draw circle (DC), zig-zag (ZZ), and clap (CL). We track multiple consecutive ToF-AoA spectra obtained by CCS-Fi and extract several magnitude values within the region corresponding to the subject, thereby obtaining a time series signal. Then, the model provided by Widar 3.0 [40] is employed to identify

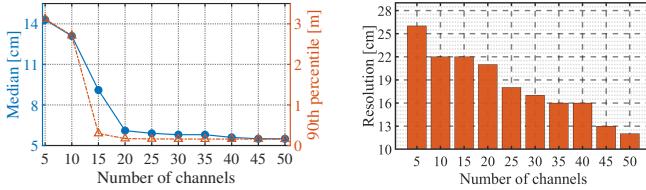
	FB	.99	.00	.00	.00	.00	.00
UD	.00	.97	.02	.00	.00	.00	.00
LR	.01	.04	.95	.00	.00	.00	.00
DC	.00	.00	.00	.98	.00	.01	.00
ZZ	.01	.00	.00	.00	.99	.00	.00
CL	.00	.01	.00	.01	.01	.96	.00

(a) Confusion matrix of gestures.



(b) F-scores for each subject.

Fig. 10: Gesture detection results.



(a) Localization accuracy.

(b) Localization resolution.

Fig. 11: Impact of channel number.

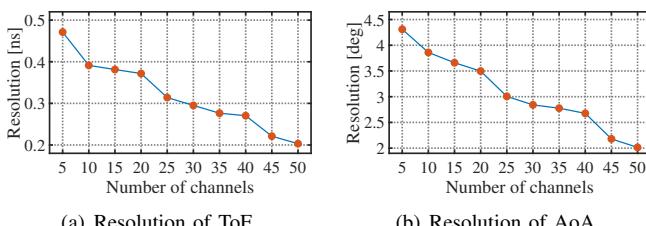
gestures from these data. The overall performance, depicted in Fig. 10(a), exceeds 97%. We further analyze the F -score [41] of each subject to comprehensively evaluate their recognition results, as depicted in Fig. 10(b). Although the F -score of each subject fluctuates due to factors such as body size or occlusion, they all exceed 0.96, indicating satisfactory results. Furthermore, as a platform capable of providing multi-person fine-granularity sensing, CCS-Fi can utilize the above pipeline to accomplish various sensing tasks, whereas the baseline method is confined solely to localization.

C. Impact Factors

1) *Number of Channels*: We first analyze the impact of the number of channels on localization accuracy and resolution. For the accuracy analysis, CCS-Fi randomly collects 5 to 50 sparse channel samples and evaluates performance leveraging median and 90th percentile errors. The results, as shown in Fig. 11(a), indicate that both the errors decrease as the number of channels increases, but these changes level off after reaching 20 channel samples.

We employ the same setup to analyze spatial resolution. In order to ensure a sufficiently small distance between the two subjects, we conduct tests on the two pendulums depicted in Fig. 7(d). The results, shown in Fig. 11(b), indicate that resolution improves with an increase in the sampled channel quantity, reaching sub-decimeter level at 50 channels. For time budget and performance considerations, CCS-Fi samples 20 channels, achieving a resolution of approximately 20 cm, which is sufficient for indoor scenarios to distinguish normal-sized adults. Notably, even with only 5 channel samples, CCS-Fi still outperforms the baseline shown in Fig. 8(b) and Fig. 9(b). This is primarily attributed to CCS-Fi's ability to achieve a wider effective bandwidth through sparse recovery compared to continuous channel stitching.

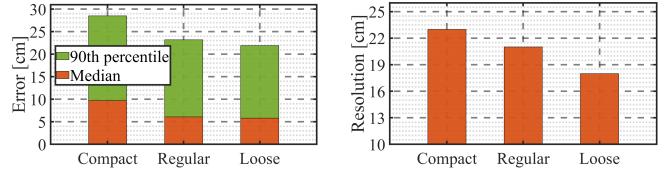
We delve into the impact of the number of channels on the equivalent ToF and AoA resolutions, as shown in Fig. 12.



(a) Resolution of ToF.

(b) Resolution of AoA.

Fig. 12: Resolutions of ToF and AoA.



(a) Localization accuracy.

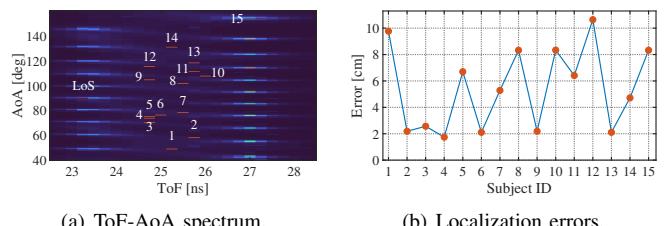
(b) Localization resolution.

Fig. 13: Impact of channel placement.

Apart from ToF, the resolution of AoA also improves as the number of channels increases, which is expected to be correlated with the number of antennas. The primary reason is that increasing the number of sampled channels in a Wi-Fi sensing system means augmenting the number of subcarriers. This actually provides a greater number of independent measurements, enabling super-resolution analysis in the AoA dimension by reconstructing the sparse channel samples into a sensor array [25], a process also learned by CCS-Fi.

2) *Placement of Channels*: Following the same experimental setup as previously described using 20 channel samples, we proceed to investigate the impact of channel placement on accuracy and spatial resolution. We design three scenarios: randomly selected channel samples (Regular), manually selected compact channel samples (Compact), and loose channel samples spanning 2.4 GHz-6 GHz bands (Loose). The results, as depicted in Fig. 13, demonstrate that both accuracy and resolution achieve the best performance under the loose placement, while performing the worst under the compact scenario. The main reason is that the broader span of the loose placement provides greater diversity, enabling the recovery of a wider effective bandwidth from sparse channel samples. Therefore, when the availability of all channels is known, manually setting a loose placement of channels can assist CCS-Fi in achieving better performance.

3) *Number of Subjects*: To demonstrate the number of subjects that CCS-Fi can distinguish, we increase to 15 people, enough to fill the CR, standing randomly in various positions. The result, as shown in Fig. 14(a), clearly displays multiple distinct peaks in the ToF-AoA spectrum (indicated by the dark warm-colored points), enabling us to identify all 15 people. Furthermore, we also conduct additional analysis on the localization errors of the 15 subjects, as depicted in Fig. 14(b). The maximum error is only 10.6 cm, while the minimum error is less than 2 cm. These findings underscore the stable performance of CCS-Fi, even in scenarios involving



(a) ToF-AoA spectrum.

(b) Localization errors.

Fig. 14: Maximum number of recognizable subjects.

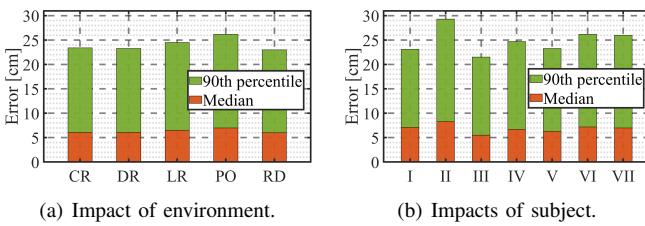


Fig. 15: Impacts of environment and subject.

a substantial number of sensed subjects.

4) *Cross-domain*: To further validate the cross-domain capability of CCS-Fi, we analyze the performance of 7 untrained subjects in 5 unknown environments. Fig. 15(a) shows the localization errors of all subjects across various environments. The results indicate that accuracy remains relatively stable, and even in the most complex PO environment, the localization error stays within an acceptable range. Fig. 15(b) illustrates the localization errors of each subject across all environments. The results show that their median and 90th percentile errors are approximately 6 cm and 20 cm, respectively. The slightly larger errors for Subject II may arise from the smaller body size or higher level of noise interference. The above results adequately demonstrate that the performance of CCS-Fi is minimally affected by the environment and subject. This is primarily attributed to the sparse channel sampling strategy effectively widening the sensing bandwidth and accurately separating subjects from complex background interference.

5) *Training Strategy*: To demonstrate the importance of Algorithm 1 (Diversification) and Algorithm 3 (Enhancement), we separately test their performance without utilizing either algorithm. Fig. 16(a) illustrates the cumulative distribution of localization errors for CCS-Fi without Algorithm 1, where the median and 90th percentile errors increase to 22 cm and 3.4 m, respectively. The results demonstrate that Algorithm 1 effectively assists CCS-Fi in enhancing cross-domain capability, which is also crucial for reducing the workload associated with training data collection. Fig. 16(b) illustrates the localization errors of CCS-Fi without Algorithm 3, where the median and 90th percentile errors increase to 12 cm and 4.4 m, respectively. This indicates that the enhancement algorithm is necessary to achieve accurate spectrum-like labels for training.

VI. RELATED WORK

In this section, we categorize existing proposals related to CCS-Fi into two main types: localization and ubiquitous

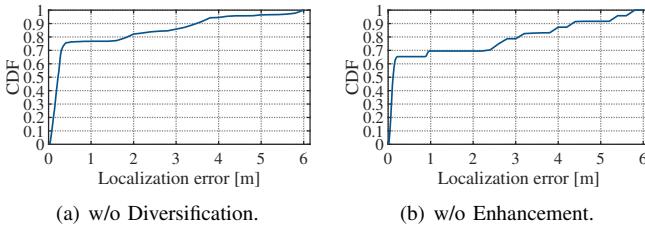


Fig. 16: Ablation studies on localization error.

sensing, based on the development of Wi-Fi sensing systems.

a) *Wi-Fi Localization System*: As a widely recognized pioneering work, SpotFi [25] treats each subcarrier as an independent measurement and reconstructs it into a sensor array, allowing for super-resolution analysis of AoA with only three antennas. To further improve ToF accuracy, Tone-Track [19] and Splicer [20] collect overlapped channel samples to expand effective bandwidth. Chronos [21] continuously samples standard channels to estimate the position of user device. M^3 [8] leverages channel hopping to enable the estimation of AoA and ToF for sub-meter accuracy. To further identify NLoS paths, LiFS [42] establishes a multi-link system and selects subcarriers unaffected by multipath interference. Widar [43] achieves tracking of subjects by establishing a quantitative relationship between CSI dynamics and their positions. Widar2.0 [9] integrates AoA, ToF, and DFS to construct a model capable of single-person tracking, and mD-Track [17] goes a step further by incorporating AoD into consideration.

b) *Wi-Fi Human Sensing System*: As a further advancement, the Wi-Fi ubiquitous sensing system greatly enhances user experiences. SLNet [3] employs a deep learning model for high-resolution spectra generation, establishing a versatile wireless sensing framework that surpasses time-frequency uncertainty constraints and enables applications such as HAR. Multi-Track [13] removes static components from the CSI to extract DFS from dynamic components associated with subject motion, thereby achieving multi-person sensing. MultiSense [16] leverages the ICA to estimate signals for facilitating multi-person respiration recovery. SenCom [4] calibrates the CSI of different Wi-Fi communication modes and provides unified measurements for sensing. In addition to algorithmic efforts, MUSE-Fi [14] utilizes signals emitted by multiple devices for sensing subjects in close proximity.

VII. CONCLUSION

We have introduced a novel CCS-Fi framework based on compressive sensing theory, which widens the effective bandwidth of the Wi-Fi sensing system. We initially design a diversification algorithm to enrich background interference within limited training data, thereby endowing CCS-Fi with cross-domain capabilities. Subsequently, to emulate radar systems, we transform tuple-structured ground truth into spectrum-like labels to train the neural network, while alleviating the fitting pressure and encoding phase information relevant to subjects in CSI. Finally, we develop an enhancement algorithm for the label, further improving its performance. We have constructed a prototype system based on these algorithms and conducted thorough evaluations. The results demonstrate that CCS-Fi successfully achieves multi-person fine-granularity sensing applications in typical indoor scenarios.

ACKNOWLEDGMENT

This research is supported in part by MOE Tier 1 grant RG16/22, Natural Science Foundation of China (NSFC) No. 62172320, and Interdisciplinary Graduate Programme (IGP). We also sincerely thank the anonymous reviewers for their constructive and insightful suggestions.

REFERENCES

- [1] Z. Chen, T. Zheng, C. Hu, H. Cao, Y. Yang, H. Jiang, and J. Luo, “ISACoT: Integrating Sensing with Data Traffic for Ubiquitous IoT Devices,” *IEEE Communications Magazine*, vol. 61, no. 5, pp. 98–104, 2022.
- [2] C. Wu, X. Huang, J. Huang, and G. Xing, “Enabling Ubiquitous WiFi Sensing with Beamforming Reports,” in *Proc. of 37th ACM SIGCOMM*, 2023, pp. 20–32.
- [3] Z. Yang, Y. Zhang, K. Qian, and C. Wu, “SLNet: A Spectrogram Learning Neural Network for Deep Wireless Sensing,” in *Proc. of the 20th USENIX NSDI*, 2023, pp. 1221–1236.
- [4] Y. He, J. Liu, M. Li, G. Yu, J. Han, and K. Ren, “SenCom: Integrated Sensing and Communication with Practical WiFi,” in *Proc. of the 29th ACM MobiCom*, 2023, pp. 1–16.
- [5] J. Wang, N. Tan, J. Luo, and S. J. Pan, “WOLoc: WiFi-only Outdoor Localization using Crowdsensed Hotspot Labels,” in *Proc. of the 36th IEEE INFOCOM*, 2017, pp. 1–9.
- [6] Y. He, G. Yu, Y. Cai, and H. Luo, “Integrated Sensing, Computation, and Communication: System Framework and Performance Optimization,” *IEEE Transactions on Wireless Communications*, vol. 23, no. 2, pp. 1114–1128, 2023.
- [7] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, “Tool Release: Gathering 802.11n Traces with Channel State Information,” *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 1, p. 53, 2011.
- [8] Z. Chen, G. Zhu, S. Wang, Y. Xu, J. Xiong, J. Zhao, J. Luo, and X. Wang, “M³: Multipath Assisted Wi-Fi Localization with a Single Access Point,” *IEEE Transactions on Mobile Computing*, vol. 20, no. 2, pp. 588–602, 2021.
- [9] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, “Widar2.0: Passive Human Tracking with a Single Wi-Fi Link,” in *Proc. of the 16th ACM MobiSys*, 2018, p. 350–361.
- [10] S. Ding, Z. Chen, T. Zheng, and J. Luo, “RF-Net: A Unified Meta-Learning Framework for RF-enabled One-Shot Human Activity Recognition,” in *Proc. of the 18th ACM SenSys*, 2020, pp. 517–530.
- [11] Y. Zhang, Y. Zheng, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, “Widar3.0: Zero-effort Cross-domain Gesture Recognition with WiFi,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 8671–8688, 2021.
- [12] P. Hillyard, A. Luong, A. S. Abrar, N. Patwari, K. Sundar, R. Farney, J. Burch, C. Porucznik, and S. H. Pollard, “Experience: Cross-technology Radio Respiratory Monitoring Performance Study,” in *Proc. of the 24th ACM MobiCom*, 2018, pp. 487–496.
- [13] S. Tan, L. Zhang, Z. Wang, and J. Yang, “MultiTrack: Multi-user Tracking and Activity Recognition using Commodity WiFi,” in *Proc. of the 37th ACM CHI*, 2019, pp. 1–12.
- [14] J. Hu, T. Zheng, Z. Chen, H. Wang, and J. Luo, “MUSE-Fi: Contactless MUti-person SEnsing Exploiting Near-field Wi-Fi Channel Variation,” in *Proc. of the 29th ACM MobiCom*, 2023, pp. 1–15.
- [15] J. Hu, H. Jiang, T. Zheng, J. Hu, H. Wang, H. Cao, Z. Chen, and J. Luo, “M²-Fi: Multi-person Respiration Monitoring via Handheld WiFi Devices,” in *Proc. of the 43rd IEEE INFOCOM*, 2024.
- [16] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, “MultiSense: Enabling Multi-person Respiration Sensing with Commodity WiFi,” *Proc. of the 22nd ACM UbiComp*, vol. 4, no. 3, pp. 1–29, 2020.
- [17] Y. Xie, J. Xiong, M. Li, and K. Jamieson, “mD-Track: Leveraging Multi-Dimensionality for Passive Indoor Wi-Fi Tracking,” in *Proc. of the 25th ACM MobiCom*, 2019, pp. 8:1–16.
- [18] C. R. Karanam, B. Korany, and Y. Mostofi, “Tracking from One Side: Multi-person Passive Tracking with WiFi Magnitude Measurements,” in *Proc. of the 18th ACM/IEEE IPSN*, 2019, pp. 181–192.
- [19] J. Xiong, K. Sundaresan, and K. Jamieson, “ToneTrack: Leveraging Frequency-Agile Radios for Time-Based Indoor Wireless Localization,” in *Proc. of the 21st ACM MobiCom*, 2015, pp. 537–549.
- [20] Y. Xie, Z. Li, and M. Li, “Precise Power Delay Profiling with Commodity WiFi,” in *Proc. of the 21st ACM MobiCom*, 2015, pp. 53–64.
- [21] D. Vasishth, S. Kumar, and D. Katabi, “Decimeter-Level Localization with a Single WiFi Access Point,” in *Proc. of the 13th USENIX NSDI*, 2016, p. 165–178.
- [22] E. J. Candès, J. Romberg, and T. Tao, “Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [23] X. Li, H. Wang, Z. Chen, Z. Jiang, and J. Luo, “UWB-Fi: Pushing WiFi towards Ultra-wideband for Fine-Granularity Sensing,” in *Proc. of the 22nd ACM MobiSys*, 2024, pp. 42–55.
- [24] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, “3D Tracking via Body Radio Reflections,” in *Proc. of the 11th USENIX NSDI*, 2014, pp. 317–329.
- [25] Kotaru, Manikanta and Joshi, Kiran and Bharadia, Dinesh and Katti, Sachin, “SpotFi: Decimeter Level Localization Using WiFi,” in *Proc. of 29th ACM SIGCOMM*, 2015, p. 269–282.
- [26] E. Perahia and R. Stacey, *Next Generation Wireless LANs: 802.11 n and 802.11 ac*. Cambridge University Press, 2013.
- [27] R. Schmidt, “Multiple Emitter Location and Signal Parameter Estimation,” *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.
- [28] K. Jiokeng, G. Jakllari, A. Tchana, and A.-L. Beylot, “When FTM Discovered MUSIC: Accurate WiFi-based Ranging in the Presence of Multipath,” in *Proc. of the 39th IEEE INFOCOM*, 2020, pp. 1857–1866.
- [29] J. Pegoraro, J. O. Lacruz, M. Rossi, and J. Widmer, “SPARCS: A Sparse Recovery Approach for Integrated Communication and Human Sensing in mmWave Systems,” in *Proc. of the 21st ACM/IEEE IPSN*, 2022, pp. 79–91.
- [30] Y. Ma, G. Zhou, and S. Wang, “WiFi Sensing With Channel State Information: A Survey,” *ACM Computing Surveys (CSUR)*, vol. 52, no. 3, pp. 1–36, 2019.
- [31] D.-X. Zhou, “Universality of Deep Convolutional Neural Networks,” *Applied and Computational Harmonic Analysis*, vol. 48, no. 2, pp. 787–794, 2020.
- [32] P. Kidger and T. Lyons, “Universal Approximation with Deep Narrow Networks,” in *PMLR Conference on Learning Theory*, 2020, pp. 2306–2327.
- [33] T. Zheng, Z. Chen, S. Zhang, C. Cai, and J. Luo, “MoRe-Fi: Motion-robust and Fine-grained Respiration Monitoring via Deep-Learning UWB Radar,” in *Proc. of the 19th ACM SenSys*, 2021, pp. 111–124.
- [34] T. Zheng, Z. Chen, J. Luo, L. Ke, C. Zhao, and Y. Yang, “SiWa: See into Walls via Deep UWB Radar,” in *Proc. of the 27th ACM MobiCom*, 2021, pp. 323–336.
- [35] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. MIT press, 2012.
- [36] E. J. Candes, “The Restricted Isometry Property and its Implications for Compressed Sensing,” *Comptes Rendus. Mathematique*, vol. 346, no. 9–10, pp. 589–592, 2008.
- [37] A. Cohen, W. Dahmen, and R. DeVore, “Compressed Sensing and Best k-term Approximation,” *Journal of the American Mathematical Society*, vol. 22, no. 1, pp. 211–231, 2009.
- [38] Z. Jiang, T. H. Luan, X. Ren, D. Lv, H. Hao, J. Wang, K. Zhao, W. Xi, Y. Xu, and R. Li, “Eliminating the Barriers: Demystifying WiFi Baseband Design and Introducing the PicoScenes WiFi Sensing Platform,” *IEEE Internet of Things Journal*, pp. 1–21, 2021.
- [39] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *Proc. of the 29th IEEE/CVF CVPR*, 2016, pp. 770–778.
- [40] Y. Zhang, Y. Zheng, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, “Widar3.0: Zero-effort Cross-Domain Gesture Recognition with WiFi,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 8671–8688, 2021.
- [41] A. Tharwat, “Classification Assessment Methods,” *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, 2020.
- [42] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang, and B. Xie, “LiFS: Low Human-Effort, Device-Free Localization with Fine-Grained Subcarrier Information,” in *Proc. of the 22nd ACM MobiCom*, 2016, p. 243–256.
- [43] K. Qian, C. Wu, Z. Yang, Y. Liu, and K. Jamieson, “Widar: Decimeter-Level Passive Tracking via Velocity Monitoring with Commodity WiFi,” in *Proc. of the 18th ACM MobiHoc*, 2017, pp. 6:1–10.