2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM) | 978-1-6654-0876-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/WoWMoM54355.2022.00027

ReWiS: Reliable Wi-Fi Sensing Through Few-Shot Multi-Antenna Multi-Receiver CSI Learning

Niloofar Bahadori†, Jonathan Ashdown‡ and Francesco Restuccia†

† Institute for the Wireless Internet of Things, Northeastern University, United States
‡ Air Force Research Laboratory, United States
Email: {n.bahadori, frestuc}@northeastern.edu, jonathan.ashdown@us.af.mil

Abstract*—Thanks to the ubiquitousness of Wi-Fi access points and devices, Wi-Fi sensing enables transformative applications in remote health care, home/office security, and surveillance, just to name a few. Existing work has explored the usage of machine learning on channel state information (CSI) computed from Wi-Fi packets to classify events of interest. However, most of these algorithms require a significant amount of data collection, as well as extensive computational power for additional CSI feature extraction. Moreover, the majority of these models suffer from poor accuracy when tested in a new/untrained environment. In this paper, we propose ReWiS, a novel framework for robust and environment-independent Wi-Fi sensing. The key innovation of ReWiS is to leverage few-shot learning (FSL) as the inference engine, which (i) reduces the need for extensive data collection and application-specific feature extraction; (ii) can rapidly generalize to new environments by leveraging only a few new samples. Moreover, ReWiS leverages multi-antenna, multi-receiver diversity, as well as fine-grained frequency resolution, to improve the overall robustness of the algorithms. Finally, we propose a technique based on singular value decomposition (SVD) to make the FSL input constant irrespective of the number of receive antennas. We prototype the ReWiS using off-the-shelf Wi-Fi equipment and showcase its performance by considering a compelling use case of human activity recognition. Thus, we perform an extensive data collection campaign in three different propagation environments with two human subjects. We evaluate the impact of each diversity component on the performance and compare ReWiS with an existing convolutional neural network (CNN)-based approach. Experimental results show that ReWiS improves the performance by about 40% with respect to existing single-antenna low-resolution approaches. Moreover, when compared to a CNN-based approach, ReWiS shows a 35% more accuracy and less than 10% drop in accuracy when tested in different environments, while the CNN drops by more than 45%. To allow reproducibility of our results and to address the current dearth of Wi-Fi sensing datasets, we pledge to release our 60 GB dataset and the entire code repository to the community.

I. Introduction

Wi-Fi has become one of the most pervasive wireless technologies ever invented. Indeed, today Wi-Fi is ubiquitous and provides wireless connectivity to almost any device of common use, including smartphones, tablets, laptop computers, and wearable devices. Just to give an idea of how fast Wi-Fi is growing, Cisco forecast that Wi-Fi 6 hotspots are expected to grow 13 fold from 2020 to 2023 [1]. Given their ever-increasing ubiquitousness, significant research efforts have investigated the usage of Wi-Fi waveforms to perform device-free classification, also called *Wi-Fi sensing*. An excellent survey on the topic can be found in [2]. In a nutshell, Wi-Fi sensing is based on passive monitoring of the changes in the channel frequency response (CFR) produced by the presence of scatterers located between the Wi-Fi transmitters and the Wi-Fi receivers. These sudden changes in CFR can be evaluated by estimating the channel state information (CSI) through the pilots contained in every Wi-Fi frame preamble [3]. This way, highly-innovative applications such as human activity recognition, remote health monitoring, and surveillance can be implemented [4]. Attesting to the relevance of these applications, in September 2020, IEEE 802.11 has approved a new technical group (TG) called IEEE 802.11bf [5]. According to the website, TGbf will define modifications to state-of-the-art IEEE 802.11 standards at both the Medium Access Control (MAC) and physical layer (PHY) to accommodate sensing operations between 1 GHz and 7.125 GHz, as well as above 45 GHz (i.e., millimeter-wave frequencies).

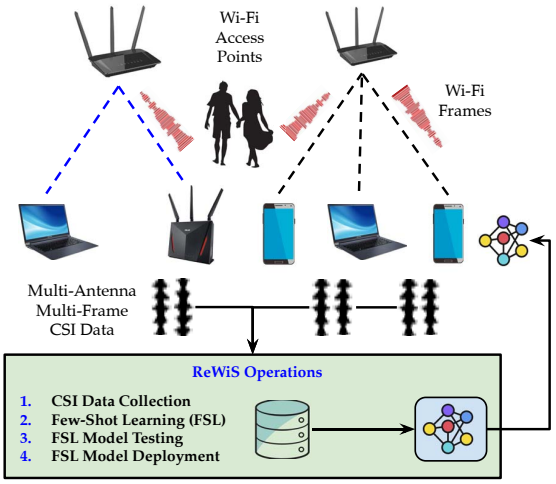


Fig. 1: High-level overview of ReWiS.

As critical Wi-Fi sensing applications come into the market, it becomes imperative to improve the *robustness* of Wi-Fi sensing operations to noise and interference, as well as their amenability to *generalize* to multiple operational environ-

Approved for Public Release; Distribution Unlimited: AFRL-2021-3408

978-1-6654-0876-9/22/$31.00 ©2022 IEEE
DOI 10.1109/WoWMoM54355.2022.00027

ments. For example, if a sensing application fails to detect intruders in a smart home scenario, or stops working when deployed in a new home, it may have severe repercussions in both commercialization efforts as well as the well-being of the final users. Although existing work – discussed in detail in Section II – has proposed CSI-based sensing, they fail to generalize to multiple environments and do not address how to improve the robustness of the algorithms.

To address the existing research gap, in this paper we propose ReWiS, a novel framework leveraging multi-antenna, multi-frame, multi-receiver CSI data to improve the robustness of Wi-Fi sensing operations. Figure 1 shows a high-level overview of ReWiS and its key operations. The fundamental difference of ReWiS with respect to existing work is that instead of relying on traditional convolutional neural network (CNN)-based learning [6, 7], ReWiS tackles the key problem of generalization through an approach based on few-shot learning (FSL), which (i) reduces the need of extensive data collection; (ii) allows ReWiS to rapidly generalize to new tasks by only leveraging a few new samples. Moreover, ReWiS leverages spatial diversity (i.e., multiple receivers and multiple antennas per receiver), time diversity (i.e., multiple CSI measurements), and increased subcarrier resolution to significantly improve the robustness of the sensing process. Although existing work has proposed FSL to address CSI learning [8], it is based on application-dependent embedding extraction using an long short-term memory (LSTM)-based technique called matching network (MatNet) [9], which ultimately limits its applicability to a single application and adds to the computational complexity of the approach. In this work, we rely on a new concept called prototypical networks (ProtoNets) [10] to do away with application-specific feature extraction, thus improving generalizability and computational burden.

This paper makes the following novel contributions:

• We propose ReWiS, a novel framework for robust and environment-independent CSI-based Wi-Fi sensing (Section III). The core design principles behind the ReWiS are to (i) leverage multi-antenna, multi-receiver diversity, as well as fine-grained frequency resolution to improve the overall robustness of the algorithms; (ii) leverage a customized version of FSL to (a) remove the need of application-specific feature extraction; (b) help generalize to new environments by only leveraging a limited number of new samples (Section III-A). We also propose a technique based on singular value decomposition (SVD) to make the FSL input constant irrespective of the number of receive antennas and window size (Section III-B). To give a perspective, in our dataset, we reduce the input size by about 80% of the original size;

• We prototype ReWiS using off-the-shelf Wi-Fi equipment and showcase its generalizability to new environment by considering human activity recognition as a use case (Section IV). We perform an extensive data collection campaign in three different propagation environments with two human subjects (IRB approval available upon request). We evaluate the impact of each diversity component on the performance and compare ReWiS with a CNN-based approach [6]. Experimental results (Section V) show that the ReWiS improves the performance by about 40% with respect to existing single-antenna low-resolution approaches. Moreover, when compared to a CNN-based approach, ReWiS shows 35% more accuracy and less than 10% drop in accuracy when tested in unseen environments, while the CNN performance drops by more than 45%. *To allow full reproducibility of our results, the ReWiS and dataset can be found at* https://github.com/niloobah/ReWiS

II. Background and Related Work

In this section, we present some background notions on Wi-Fi sensing, as well as presenting related work and highlighting the novelty of this paper. First, we summarize Wi-Fi sensing and channel state information (CSI) learning in Sections II-A and II-B. Then, we summarize CSI collection methodologies and current CSI dataset availability in Section II-C.

A. Wi-Fi Sensing

The main idea behind Wi-Fi sensing is that the presence of moving objects alters the propagation environment throughout time and frequency. Thus, information on the presence/motion of objects in the environment can be captured by analyzing the received Wi-Fi signal. Thus, Wi-Fi sensing has been utilized in a wide range of sensing applications such as smart homes [11], device-free surveillance [12], activity recognition [7, 13, 14], and healthcare [15, 16] applications. For a comprehensive survey on the topic, please refer to [2]. Wi-Fi sensing approaches can be categorized into three main groups, (i) received signal strength (RSS), (ii) passive Wi-Fi radar (PWR) and (iii) CSI sensing. RSS sensing is performed through measuring RSS per packet, thus, is a coarse-grained parameter and provides a low-resolution information [13]. Although PWR sensing relies on calculating the difference between transmit and receive signal, its accuracy is highly dependant on the location of the bistatic [14]. Conversely, thanks to the new methodologies, CSI can be captured in real-time over up to 8 antennas and up to 160MHz channels for each packet [17], which provides us with fine-grained information in both time and frequency domains.

B. CSI Learning

Deep learning has been proved to be an effective tool for accurate Wi-Fi CSI sensing. To improve the performance of Wi-Fi sensing in terms of accuracy, some approaches have focused on preprocessing and feature extraction [15, 16, 18–20], while others applied more sophisticated machine learning (ML) models [6–8, 21, 22]. However, most of these approaches are unable to generalize and their performance deteriorates when tested in an unseen environment [7, 23]. Commercial CSI sensing algorithms are trained offline and the source model is expected to be adapted to target environments rapidly given a limited number of samples. To this end, *generalizing* the model to different environments is a crucial factor for the success of Wi-Fi sensing. A very limited number of existing papers have focused on this aspect of CSI sensing [7, 8, 22, 23].

Authors in [7] generalize to a broader range of features by designing a very sophisticated deep learning algorithm using attention-based bi-directional long short-term memory (ABLSTM) to extract significant sequential features from raw CSI measurement. To generalize to new users, CsiGAN in [22] leverages generative adversarial network (GAN) to generate artificial examples to account for the left-out user whose CSI data is not available in the training examples of the deep learning model. Although these frameworks successfully improved the accuracy of predictions, despite their computationally intensive training procedure, they fall short when it comes to generalizing to new environments. For example, as attested in [7], the accuracy of the model reduces to 32% when tested in an untrained environment. The authors in [23] suggested extracting the power distribution of different gestures' velocities from the Doppler spectrum. Then, a temporal learning model is applied to learn the extracted features and perform domain-independent gesture recognition. However, these approaches require extensive data pre-processing, careful feature extraction, and an abundance of training examples. For example, experimental results in [23] show that the accuracy of the algorithm drops significantly when the number of human subjects and training samples decreases.

Conversely, few-shot learning (FSL) is a novel ML technique that has shown remarkable results in generalizing from a few examples and classifying examples from previously unseen classes given only a handful of training data. Thanks to this unique feature, FSL can be an exceptional candidate for CSI learning problems. A recent work [8] addresses human activity recognition through FSL. A CSI feature extraction method along with a matching network (MatNet) [9] is used to remove the environment-dependent data and to make accurate predictions in new environments. However, MatNets require designing application-dependent embedding functions, through long short-term memory (LSTM) and attention-based LSTM networks. This adds to the computational complexity of the algorithm and limits its generalizability to only certain applications. *To address these challenges, we use state-of-the-art prototypical network (ProtoNet) [10], which (i) achieves the same level of performance with a simpler learning structure, and (ii) can be generalized to a variety of CSI sensing applications.*

C. CSI Collection and Datasets Availability

Reliable and high-resolution CSI datasets are key enablers of any learning-based CSI sensing applications. In Wi-Fi, CSI data can be estimated at the receiver over all subcarriers through pilot symbols contained in the physical layer (PHY) preamble. Being computed at the PHY, CSI is not accessible by the end-user through normal network interface cards (NICs), which makes CSI acquisition a challenging task. Some software-defined radio (SDR) Wi-Fi implementations do exist [24], but to the best of our knowledge, they are limited to 20 MHz bandwidth only (IEEE 802.11 a/g/p). For this reason, the majority of research works have used 802.11g and 802.11n standards for data collection, which limits the total bandwidth to 40 MHz [25]. To address this shortcoming, we leverage the recently released Nexmon CSI-extractor tool [17] to obtain CSI from IEEE 802.11ac transmissions at 80 MHz of bandwidth. By leveraging Multiple Input, Multiple Output (MIMO), we collected CSI with four receiver antennas. In Section V, we show that increasing the number of antennas, receivers, and subcarriers over which the CSI data is measured increases the CSI sensing performance significantly.

To the best of our knowledge, only a few CSI sensing datasets are publicly available. FalldeFi [26] is a fall detection dataset consisting of information collected in 6 rooms. human activity recognition (HAR) dataset is collected in [27] and [28] where measurements are collected inside 1 and 3 rooms, respectively. In all cases, 1 transmitter and 1 receiver are utilized. The size of each instance in all datasets is (2000, 3, 30), representing 2000 CSI matrices with three receive antennas and 30 subcarriers measured in two seconds. In this paper, for the first time, *we collect a large-scale dataset containing multi-antenna (up to 4) multi-receiver (up to 3) fine-grained (up to 242 subcarriers) CSI readings from multiple environments and activities, which we will release to the community.*

III. The ReWiS Framework

This section describes the proposed ReWiS framework for robust CSI sensing. Our framework's core focus is on (i) reducing the cost of data acquisition and labeling by adopting customized ProtoNet to learn and generalize with limited data; (ii) reducing the complexity of the learning algorithm and data pre-processing while maintaining the accuracy by collecting data over multiple antennas, multiple receivers at different locations and high-resolution CSI data over large channel bandwidth. We first explain the core design principles in Section III-A. Then in Section III-B, we describe the ReWiS CSI processing procedure. Finally, we present our FSL-based technique for robust CSI learning in Section II-B.

A. ReWiS Robust Design Principles

The ReWiS is a passive sensing system leveraging CSI data computed through listening to ongoing traffic exchange between Wi-Fi devices. The key motivation behind the ReWiS is that the accuracy of the CSI computation depends on the coherence time of the channel, the received power of the transmitted signal, interference from other transmissions and background noise, among other factors. For this reason, the ReWiS leverages multiple receivers and multiple antennas to collect robust CSI data and thus increases the accuracy. Below, we detail the impact of each factor in increasing the accuracy of CSI sensing.

(1) Spatial diversity. Due to the presence of multiple reflectors and scatterers in indoor environments, there is a significant probability that the communication channel is in a deep fade, and thus, that the CSI measurement may be erroneous. This motivated us to increase the reliability of the CSI sensing framework through adding spatial diversity. Since we do not have any control over the final configuration of Access Points (APs) in the target environment, we must

ensure that the Orthogonal Frequency Division Multiplexing (OFDM) symbols for CSI estimation pass through multiple signal paths, each of which fades independently, guaranteeing that reliable CSI measurement is possible even if some paths are in a deep fade. Therefore, in ReWiS we incorporated two types of diversity, namely, *macro-diversity* and *micro-diversity*.

(a) *Micro-diversity.* ReWiS improves performance by collecting CSI data with multiple receive antennas. As long as the antennas are placed sufficiently far apart, this would create independent propagation paths between different antenna pairs, thus reducing the chance of deep fade significantly. In indoor environments, the channel decorrelates over shorter spatial distances [29], and the typical Wi-Fi router's antenna separation of the half to one carrier wavelength is sufficient (about 3-6 cm at 5GHz band). To make ReWiS compatible with legacy Wi-Fi systems, we chose to leverage one spatial stream and use the full diversity on the receiving side by activating four receive antennas. **Our experimental results in Table I show that micro-diversity helps increase classification accuracy by up to 16% with respect to a single-antenna system**, as micro-diversity improves robustness to multipath fading and interference.

(b) *Macro-diversity.* ReWiS leverages the usage of multiple receivers to further improve diversity. Indeed, modern indoor Wi-Fi networks commonly leverage multiple APs to increase the wireless range, boost coverage, ensure reliability and support a large number of wireless applications/devices. Thus, ReWiS collects CSI simultaneously from multiple receivers. **The experimental results in Table I show that macro-diversity improves the prediction accuracy by 38% with respect to a single-receiver system**, as the receivers are located several wavelengths apart from each other. On the other hand, macro-diversity implies that the data from multiple receivers must be aligned both in time and frequency domains.

(2) Time diversity. The Wi-Fi propagation environment is subject to almost continuous change, mainly owing to the movement of obstacles between the transmitter and the receiver, as well as the presence of noise and interference from overlapping channels. For this reason, to improve robustness, ReWiS leverages the usage of multiple, subsequent CSI readings to boost the classification accuracy. To compensate for the increase in complexity, we use a novel and custom-tailored technique based on singular value decomposition (SVD), which is explained in Section III-B (Step 4: Dimension reduction). **We show that our technique helps reduce the input size by about 80%.** Moreover, by trading off delay for accuracy (the more CSI readings, the more delay), **we show in Figure 2 that time diversity increases accuracy by up to 35%**, as the number of CSI readings fed to the learning model help counteract the adverse channel conditions.

(3) Subcarrier resolution. CSI measurement at low channel bandwidth cannot resolve multipath propagations and thus limiting many CSI sensing applications that require higher precision [15, 30]. Inevitably, either feature extraction algorithms have been utilized or sampling frequency has been increased, to compensate for the low resolution in the frequency domain. In order to avoid time- and computational- intensive data pre-processing for complex feature extraction and efficient sampling, we proposed using the Nexmon tool [17] to extract high-resolution 802.11ac CSI data over the 80 MHz wide channel with 242 subcarriers. **The results in Figures 7 and 8 show that using higher subcarrier resolution improves prediction accuracy by up to 19%**, as finer-grained CSI data is fed to the learning model.

B. ReWiS CSI Processing

We now describe the procedure ReWiS converts the unprocessed CSI measurements into a dataset used as the input of the ML algorithm. We term this processing *CSI data preparation*, and is depicted in Figure 2.

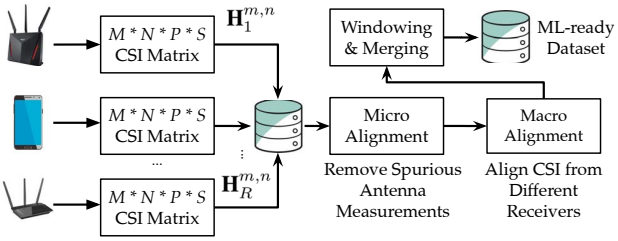


Fig. 2: ReWiS CSI dataset preparation and dataset construction.

As explained earlier, ReWiS leverages multiple receivers with multiple antennas, as well as fine-grained CSI estimated by the Wi-Fi OFDM receiver. In an OFDM system, the digital data stream is modulated over multiple overlapping, closely spaced orthogonal subcarriers to transmit data in parallel. Let us assume P packets are captured during the data collection campaign. In a $M \times N$ MIMO OFDM system with M transmit antennas, N receive antennas and S subcarriers, the extracted CSI matrix between the transmit antenna m and the receive antenna n located on receiver r can be written as

$$\mathbf{H}_r^{m,n} = \begin{bmatrix} h_{1,1}^{m,n} & \dots & h_{1,s}^{m,n} & \dots & h_{1,S}^{m,n} \\ \vdots & & \vdots & & \vdots \\ h_{p,1}^{m,n} & \dots & h_{p,s}^{m,n} & \dots & h_{p,S}^{m,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ h_{P,1}^{m,n} & \dots & h_{P,s}^{m,n} & \dots & h_{P,S}^{m,n} \end{bmatrix}, \quad \begin{matrix} 1 \le n \le N \\ 1 \le m \le M, \end{matrix}$$

where the element $h_{p,s}^{m,n}$ denotes the amplitude and phase information of the CSI obtained from p-th packet, the s-th OFDM subcarrier over the channel from the transmitter m to the receiver n. For example, when $P = 100$ and 80 MHz channels, the matrices $\mathbf{H}_r^{m,n}$ have $P \times S = 100 \times 242$ elements.

Step 1. Micro alignment. During our experiments, we noticed that in some instances, CSI packets are not captured by all antennas. In that case, those CSI measurements are identified and removed.

Step 2. Macro alignment. In addition to antenna-level alignment, the data collected from different receivers are aligned to ensure that each CSI element collected over different receivers represents the same time and frequency domain channel measurements. To this end, the Wi-Fi sequence number is used to match data from different receivers.

Step 3. Segmentation and integration. At this step, first, to remove noise and unwanted amplification, the CSI elements are normalized by the mean amplitude over all subcarriers. Next, the CSI measurements from each antenna, $\mathbf{H}_r^{m,n}$ is divided into fixed-size *data-segments* by sliding a non-overlapping window through the time-domain measurements. The data-segments, denoted by $\hat{\mathbf{H}}_r^{m,n}$, are matrices of dimension $W \times S$, where W is the number of OFDM packets in a window. Further, measurements from all antennas on the same receiver are stacked into one matrix to form a *data-frame* as

$$\mathbf{H}_r = \left[\hat{\mathbf{H}}_r^{m,1}, \cdots, \hat{\mathbf{H}}_r^{m,N}\right]^T \tag{1}$$

where data-frame $\mathbf{H}_r = \mathbf{H}_r^A e^{j\mathbf{H}_r^\phi}$ is a complex matrix. Note that the transmit antenna index m is omitted in the interest of readability. Moreover, both amplitude matrix $\mathbf{H}_r^A = \|\mathbf{H}_r\|$ and phase matrix $\mathbf{H}_r^\phi = \angle\mathbf{H}_r$ can be utilized for CSI sensing individually. For notation brevity, from here on, $\mathbf{H}_r$ is a real matrix representing both $\mathbf{H}_r^A$ and $\mathbf{H}_r^\phi$ matrices.

Step 4. Dimension reduction. Data-frames with size $N \times W \times S$ may be too large to be fed into the learning module. To improve the performance and processing time of the learning algorithm, data-frames are minimally pre-processed with singular value decomposition (SVD), which is a powerful tool to eliminate the less important variables of large-size data matrices and produce an approximation with lower dimensions. To preserve the subcarrier resolution, we only reduce the number of features in the time domain, i.e., CSI packets. By using SVD, the data-frame matrix in (1) is factored into the product of three matrices as $\mathbf{H}_r^T = U\Sigma V^T$, where diagonal values of Σ contains the singular values of the data-frame and U and V are known as the left- and right-singular vectors.

The key to understanding SVD functionality is that by multiplying $\mathbf{H}_r^T$ with the left-singular vector V, time-domain measurements over each subcarrier is mapped to the *subcarrier space*, to preserve only the useful packets over each subcarrier. Using eigenvalue analysis, we noticed that singular value contributions of the data-frames depend on the type of activity and the environment. Hence, we utilize all singular values to maintain the `ReWiS` generalizability. The compact data-frame, $\mathbf{H}'$ with dimension $S \times S$ is calculated as

$$\mathbf{H}'_r = \mathbf{H}_r^T \times V. \tag{2}$$

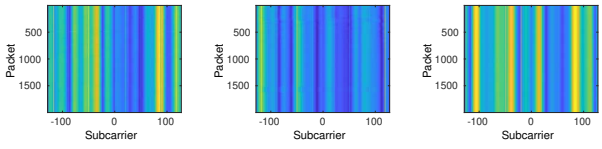
Further, to extract more features from the compact data-frame, the linear correlation among subcarriers is extracted simply by calculating the Pearson Correlation Coefficient (PCC) of the compact data-frame [31]. These correlation matrices, denoted as $\rho_{\mathbf{H}}$ with dimension $S \times S$ are the input of the `ReWiS` learning module. Therefore, we are able to reduce the dimensionality from $N \times W \times S$ to $S \times S$, *which makes the input constant with the window size and the number of receive antennas*. For example, since in our experiments we used 4 antennas and a window size of up to 300, we are able to reduce the input size *by about 80%*.

C. `ReWiS` Learning Models

We first discuss preliminaries about the FSL and related models. Then, we present the `ReWiS` CSI learning procedure.

FSL and ProtoNet. Traditionally, supervised learning approaches require a large labeled dataset for training. In applications such as Wi-Fi sensing, collecting and labeling large datasets may be challenging, thus the labeled data is severely limited [22]. For this reason, `ReWiS` uses few-shot learning (FSL), where the objective is to quickly adapt to new/unseen data given a limited number of samples. The purpose of FSL is to train a model with high accuracy when the data of the target task is small by using some kind of prior knowledge. In practice, FSL is useful when training examples are hard to find, or the cost of labeling is high [9].

Specifically, in K-way-N-shot learning, the model is trained through a set of sampled mini-batches of classes and data points, called *training tasks*. Each task is divided into a query set and limited support set with K different classes and N labeled examples (shots) of each class (typically 1-5). The small training set requires the fast adaptability of the model. The query set is classified using the knowledge from the support set. The performance of the model in generalizing to new classes is evaluated by the average test accuracy across many K-way-N-shot *test tasks* containing unseen/new classes. In a nutshell, by learning a generalizable metric space, FSL learns *how* to classify given a set of training tasks, and exploit this knowledge to classify new classes. It is worth noting that in CSI sensing applications a new environment can be interpreted as a new class. To illustrate this point, Figure 3 shows CSI measurements of empty rooms in 3 different environments. *Although all three are labeled as empty room, it can be seen that they show distinct patterns due to the unique propagation profile of each environment*.



(a) E_1 (b) E_2 (c) E_3

Fig. 3: CSI measurements of an empty room in the data collection environments described in Figure 5.

Different FSL approaches, such as MatNet [9], meta-learning [32] and ProtoNet [10], propose different metric spaces in which classification can be performed. Among these methods, ProtoNet applies a simple yet effective inductive bias in the form of class *prototypes* that leads to achieving impressive few-shot performance and reducing the network complexity. Next, we briefly explain the fundamentals of the ProtoNets.

ProtoNet model. During e-th training *episode*, a *training task* comprising a mini-batch of classes and data points are sampled as $\mathcal{T}_e = \{\mathcal{D}_1, \ldots, \mathcal{D}_K\}$. $\mathcal{D}_k = \{(\mathbf{x}_i, y_i) | y_i = k\}$ denotes the set of data points labeled with class k, where

each $\mathbf{x}_i$ is a data point and $y_i \in \{1, \ldots, K\}$ is the corresponding label. Next, $\mathcal{T}_e$ is further divided into *support* and *query* sets, i.e., $\mathcal{T}_e = \{\mathcal{S}_e, \mathcal{Q}_e\}$. A subset of each class set, $\mathcal{D}_k$ with N instances is selected as the support set: $\mathcal{S}_e = \{(\mathbf{s}_1, y_1), \ldots, (\mathbf{s}_N, y_1), \ldots, (\mathbf{s}_N, y_K)\}$ and the rest of the examples are used as the query set, $\mathcal{Q}_e = \{(\mathbf{q}_1, y_1), \ldots, (\mathbf{q}_L, y_1), \ldots, (\mathbf{q}_L, y_K)\}$ where L is the number of samples in the query set.

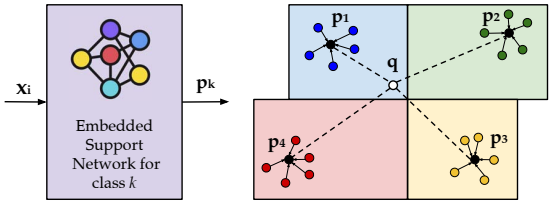


Fig. 4: Embedded Prototype Network.

To avoid suffering from high variance caused by the high dimensionality of the $\mathbf{x}_i$, the support and query data are mapped into a feature space by embedding function $f_\theta(\mathbf{x}_i)$ with cleanable parameters θ, as shown in Figure 4. Learning a proper embedding model is a paramount task [33] that will be discussed later in-depth in this section.

The principle idea behind ProtoNet is that data points cluster around a single prototype representation for each class that is simply the mean of the embedded support samples of each class [10]. Therefore, the *prototype* of each class, $\mathbf{p}_k$, is computed as

$$\mathbf{p}_k = \frac{1}{|\mathcal{D}_k|} \cdot \sum_{(\mathbf{s}_i, y_k) \in D_k} f_{\boldsymbol{\phi}}(\mathbf{s}_i) \tag{3}$$

These prototypes can be used to classify the query samples, as illustrated in Figure 4. For a query point $\mathbf{q}$, the ProtoNet produces a distribution over classes using a softmax over distances to the prototypes in the embedding space as

$$\mathcal{L}(\mathcal{Q}_e) = \frac{-1}{|\mathcal{Q}_e|} \sum_{(\mathbf{q}_i, y_i) \in \mathcal{Q}_e} \log \left(\frac{\exp(-\|f_{\boldsymbol{\theta}}(\mathbf{q}_i) - \mathbf{p}_k)\|^2)}{\sum_{k'} \exp(-\|f_{\boldsymbol{\theta}}(\mathbf{q}_i) - \mathbf{p}_{k'}\|^2)} \right). \tag{4}$$

The ProtoNet is trained through minimizing the loss function via stochastic gradient descend (SGD) over training episodes. At test time, sample $\mathbf{x}_i$ is classified using the nearest-neighbouring prototype computed from the support set of the test episode as $y(\mathbf{x}_i) = \arg\min_{j \in \{1,\ldots,K\}} \|\mathbf{x}_i - \mathbf{p}_j\|^2$. It is worth noting that `ReWiS` adopts Euclidean distance rather than cosine similarity used in [8] as it is shown through extensive experiments that it outperforms other distance metrics [10].

Learning embedding model f_θ. The ProtoNet's goal is to learn a transferable embedding that generalizes to new tasks. Previous work [8] has adopted a MatNet and employed two distinct deep neural networks (DNNs) for query samples and support samples. In addition, attention-based LSTM is used to encode a full context embedding of support samples. *Unlike [8], we use only one embedding function for both support and query sets* rather than multiple embedding functions. *This reduces the number of hyper-parameters, simplifies the learning process of CSI sensing and reduces the time- and computational- complexity.* In addition, learning the embeddings according to the dataset *makes `ReWiS` application-independent*. Inspired by [33], `ReWiS` learns the embedding function f_θ by training a neural network on the entire training set, i.e., we merge all the support mini-batches into a single set as $\mathcal{S} = \cup\{\mathcal{S}_1, \ldots, \mathcal{S}_e, \ldots, \mathcal{S}_E\}$, where E is the total number of training tasks. The embedding model parameters is then achieved by

$$\theta = \arg\min_{\theta} \mathcal{L}^{ce}(\mathcal{S}; \theta), \tag{5}$$

where $\mathcal{L}^{ce}$ denotes the cross-entropy loss between predictions and ground-truth labels.

D. `ReWiS` Training and Inference

`ReWiS` learns from the training dataset collected in the *source* environment to perform the related CSI sensing application in the *target* environment. The dataset of each receiver r consists of CSI data-frames of size $S \times S$ and corresponding labels, as explained in Section III-B. Using these datasets, a learning model is trained for each receiver. The training dataset is sampled into mini-batches of training and testing tasks. Each task includes a support and a query set of K classes and N examples (shots) as explained in Section III-C.

Training. First, the support mini-batches are fed into an embedding learning module to train a proper embedding function. The embedding is trained using a convolutional neural network (CNN) with four convolutional blocks. Each block comprises a 64-filter 3×3 convolution, batch normalization layer [34], a ReLU nonlinearity, and a 2×2 max-pooling layer is applied after each of the blocks. Finally and a global average-pooling layer is on top of the fourth block to generate the feature embedding. All of our models were trained via SGD optimizer with Adam [35]. The learning rate is initialized as 10^{-3} and is cut in half every 2000 episodes. Following learning the proper embedding model, the learning proceeds by minimizing the loss function in (4).

Testing. During testing time, the algorithm only requires N samples (also called *shots*) of the desired activities to function in the target environment. In a real-world scenario, this can be realized through a mobile application installed on the end-users smartphone. As a part of the initialization process, the user is asked to perform some activities for a short amount of time (less than a minute) and label the activities. The learning algorithm utilizes the provided labeled data by the user as the support set of the testing task and adapts to the target environment. Note that initialization and collecting data in the target environment is a common practice of commercial activity recognition tools, such as Qualcomm's positioning units [36], and does not impact the practicality of this approach. As explained earlier, there exists a trained model per receiver. At testing time, each model returns a probability distribution vector over all classes as $P_r = [p_1, \ldots, p_k]$. By defining R as the number of receivers, the final decision on the label of a data

point y is made through the superposition of all probability vectors and finding the maxima $y = \arg\max_k \sum_{r=1}^{R} P_r$.

IV. Prototype and Experimental Data Collection

To demonstrate the robustness of ReWiS in generalizing to new environments, we designed a testbed with commercially-available off-the-shelf Wi-Fi devices. We evaluate the performance of ReWiS by building a prototype and analyzing a use-case application of CSI sensing, i.e., activity recognition. First, we present our evaluation methodology including environment setup, measurement tools, data collection campaigns, and training/testing procedure. Then, we show our results which investigate the following crucial aspects:

1) The role of ProtoNet in enabling ReWiS generalizing to unseen environments.
2) The impact of macro-, micro- diversity, and subcarrier resolution on the accuracy of ReWiS predictions.

Note that our main goal is to demonstrate the ability of ReWiS in generalizing to new environments, not recognizing human activities. Thus, we suffice our experiment to only four activities, namely *empty room*, *walking*, *jumping*, and *standing*, each of which has its own specific challenges. Notice that even identifying an empty room is not a trivial task, as changes in environment/time of experiment lead to significant alterations in the CSI measurements as shown in Figure 3. Considering various activities with more complexities will be left to our future work. The implementation of the ReWiS prototype requires a pair of Wi-Fi transmitter-receiver to establish a traffic link, in addition to a set of Wi-Fi routers equipped with an extraction tool to process the CSI data. In the following, we detail the ReWiS components as well as the experiment setup.

A. CSI Extraction, Hardware and Testbed Setup

We have used Nexmon CSI [17], the state-of-the-art CSI extraction tool to collect CSI measurements using Asus RT-AC86U Wi-Fi routers. In a real-world scenario, ReWiS relies on already existing Wi-Fi transmissions. In our controlled experiment, we dedicated a pair of transmitter-receiver to emulate the traffic generation. To establish the Wi-Fi link, a Netgear R7800 Wi-Fi router with a Qualcomm Atheros chipset is used in AP mode. An off-the-shelf laptop acts as the client. The Hostapd tool [37] is utilized on the AP to ensure that the traffic is generated using 802.11ac at the desired bandwidth. UDP packets between the AP and the client are generated with a rate of 1 Mbit/s via iperf3 tool [38] and transmitted through a single spatial stream.

To implement our prototype ReWiS, we have used 3 Asus RT-AC86U WiFi routers, each equipped with $N = 4$ antennas. The Asus routers extract the CSI packets using the Nexmon firmware, by computing the CSI on the UDP frames transmitted from the AP to the client. CSI is computed at a rate of 100 Hz. The Nexmon tool enabled us to collect 802.11ac channel measurements at 5GHz with 20 and 80 MHz bandwidth over $S = 52$ and $S = 242$ subcarriers. The measurements at 20 MHz are used to evaluate the effectiveness of higher subcarrier resolutions as compared with legacy CSI measurements. Note that originally 802.11ac 20 MHz and 80 MHz channels consist of 64 and 256 subcarriers. However, in our measurements, we discarded the CSI from the guard and null subcarriers as they contain arbitrary values [17].

To evaluate the capability of ReWiS to generalize to different environments, we performed CSI measurements in three different environments. Specifically, the measurements are carried in an office area E_1, a meeting room E_2, and a classroom E_3 on different days and times. We carefully picked the environment with exclusive furniture arrangements, size, and construction material to ensure that the target environments are mutually exclusive from the source environment in terms of propagation characteristics. However, visually, E_1 is more similar to E_3 than E_2. Figure 5 shows the environment layout, as well as the position of AP, client and CSI sniffers. The position of ReWiS components may influence the test outcome. Therefore, we attempted to loosely change the configuration in different environments and design the testbed to align with realistic home/office HAR scenarios and be close to those used in other CSI activity recognition studies [22, 39].

B. Data Collection and Dataset Preparation

The two subjects involved in the experiments were instructed about the type, duration, and location of the activities including jumping, walking, and standing. Each measurement campaign involves 180 seconds of data collection for each activity performed by two people (IRB approval available upon request). The measurements are repeated 10 times with a time interval of at least 2 hours in between measurements. The collected raw data is processed by applying the CSI data preparation algorithm presented in Section III-B. Upon aligning the data collected from 4 antennas of each Asus router, a window size of $W = 300$ samples (tantamount to 3s) is used to segment the raw data into data-segments. Further, the data-segments are integrated using (1) to form the CSI data-frame $\mathbf{H}_r$. As a result, through the explained measurement campaign with $M = 1$ spatial stream, $N = 4$ receive antennas, the CSI data-frame $\mathbf{H}_r$ at 80 MHz is a matrix of size 1200×242. Thanks to the dimension reduction algorithm proposed in Section III-B, the size of the data-frame is reduced to 242×242, which is about 80%. We collected data in three different propagation environments, which are shown in Figure 5.

Environment E_1 is selected as the source environment, while E_2 and E_3 are considered as the target environments, and their dataset is entirely used for testing purposes. The ReWiS learning algorithm presented in Section III-C has been trained using the prepared dataset, where 70% of the dataset is used for training while the remainder is used for testing and evaluation. Unless otherwise mentioned, the model is trained using 4-way-5-shot training tasks, i.e., 4 classes and 5 examples in both support and query tasks. The accuracy of the algorithm is tested through sampling 1000 randomly generated testing tasks from the test sets. We have used two measures to

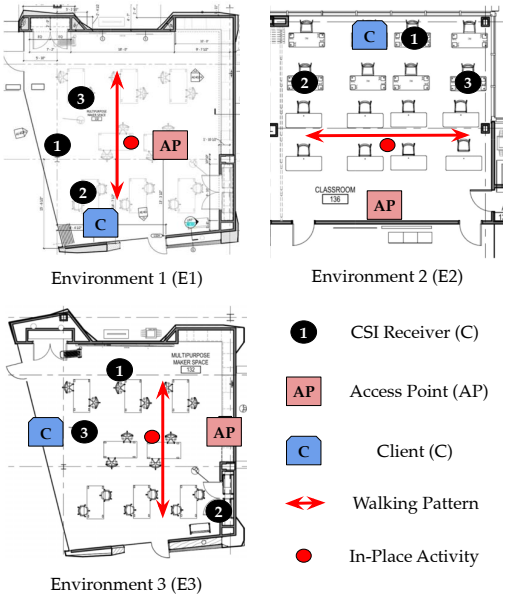


Fig. 5: Data collection and testing environments used in our experiments. We also report the legend of symbols used in the figure.

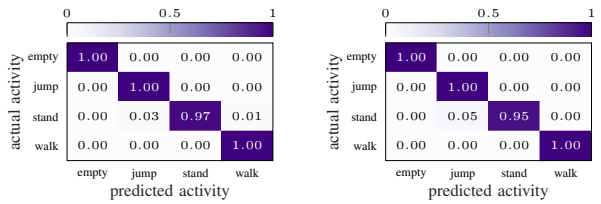
report the results, namely accuracy and F1 score. The accuracy refers to the correct predictions divided by total predictions, while the F1 score is defined as

$$\text{F1 score} = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

where Recall $:=$ TP$/$(TP $+$ TN) and Precision $:=$ TP$/$(TP $+$ FP). TP, TN and FP stand for true positive, true negative and false positive, respectively.

V. Performance Evaluation

In this section, we show the end-to-end performance evaluation of ReWiS. Figure 6 shows the overall performance of ReWiS in terms of accuracy of predictions with 3 receivers, each with 4 antennas, at 80 MHz channels when tested in environment E_2 and E_3. It can be seen that ReWiS not only learned very well from the source dataset but successfully generalized to the target environments, as it is able to achieve accuracy close to 100% in unseen environments by taking advantage of all three components of macro/micro diversity and subcarrier resolution, in addition to implementing ProtoNet with modified embedding learning.

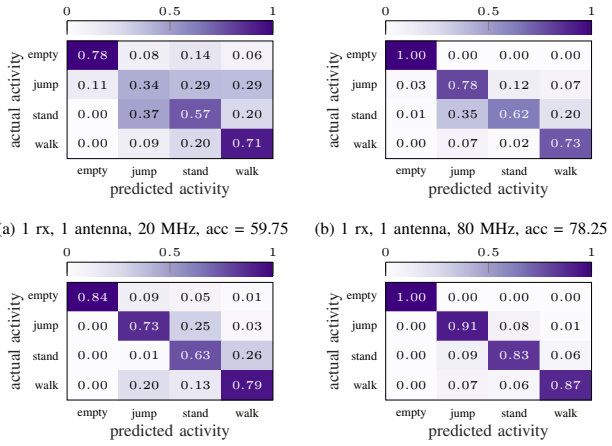


(a) E_3, acc = 98.85 (b) E_2, acc = 99.25

Fig. 6: The performance of ReWiS in generalizing to E_2 and E_3 environment, with 3 receivers, 4 antennas, 80 MHz.

1) Impact of Macro- and Micro-diversity: Table I shows the impact of macro- and micro- diversity on the accuracy of the predictions. It can be seen that increasing the number of antennas from 1 to 4 increases the accuracy by 12% and 10% in environments E_2 and E_3, respectively. In addition, increasing the number of receivers from 1 to 3 improves the accuracy by 14% and 18% in environments E_2 and E_3, respectively. By looking closely at these results, we can see that in general, micro-diversity (number of antennas) improves the accuracy of in-place activities, such as jump and standing, while macro-diversity improves the accuracy of walking. Interestingly, macro-diversity helps discriminate walking from standing and jumping and micro-diversity increases the precision of differentiating jumping from standing.

2) Impact of Subcarrier Resolution: Figures 7 and 8 show the impact of subcarrier resolution and the number of antennas by comparing the confusion matrices of ReWiS in 20 MHz and 80 MHz channels, respectively with 1 receiver and 3 receivers. It can be seen that at 20 MHz with 1 receiver the accuracy diminishes by close to 20% in the worst case. The number of receivers improves the performance, but we still notice that higher resolution implies 6% better accuracy. In general, differentiating standing from jumping and walking is a challenging task, since our human subjects were allowed to make small body movements like moving arms and head (as it is naturally). Based on this result, we conclude that higher frequency can be beneficial for detecting activities with micro-movements like respiratory detection [15].



(a) 1 rx, 1 antenna, 20 MHz, acc = 59.75 (b) 1 rx, 1 antenna, 80 MHz, acc = 78.25

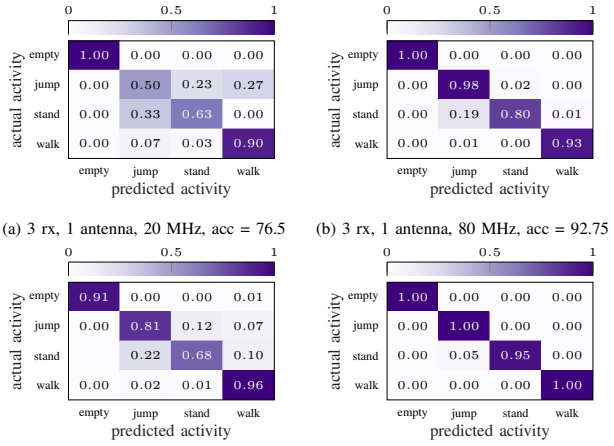
(c) 1 rx, 4 antennas, 20 MHz, acc = 74.83 (d) 1 rx, 4 antennas, 80 MHz, acc = 90.42

Fig. 7: Impact of frequency resolution and diversity, with one receiver. E_2 is the target environment.

3) Comparison of CNN vs. ReWiS: Figure 9 compares the performance of ReWiS learning framework with a baseline CNN classifier [6]. The classifier utilizes a CNN for feature extraction with the same structure as the one ReWiS uses for embedding function training, explained in Section III-C. Further, a 3-layer fully connected network is used for classification. Overall, the total number of parameters of the baseline classifier is comparable with the ReWiS learning framework.

			empty		jumping		standing		walking		mean	
area	rx.	antenna	Accuracy	F1 score	Accuracy	F1 score	Accuracy	F1 score	Accuracy	F1 score	Accuracy	F1 score
E_1	1	1	100	0.9967	85.34	0.9412	82.67	0.8575	79.00	0.8531	87.00	0.9076
	1	4	100	0.9852	98.00	0.9484	94.00	0.9292	85.67	0.9212	94.41	0.9460
	3	1	100	1.0000	100	0.9891	98.63	0.9905	98.67	0.9928	99.32	0.9931
	3	4	100	1.0000	100	1.0000	99.28	0.9960	100	0.9980	99.82	0.9977
E_2	1	1	100	0.9804	78.00	0.7091	62.00	0.6392	73.00	0.8022	78.25	0.7827
	1	4	100	1.0000	91.00	0.8778	83.33	0.8460	87.33	0.8927	90.42	0.9041
	3	1	100	1.0000	98.00	0.9018	80.00	0.8775	93.00	0.9884	92.75	0.9417
	3	4	100	1.0000	100	0.9800	95.40	0.9853	100	1.0000	98.85	0.9085
E_3	1	1	100	0.9967	74.37	0.7348	87.67	0.7799	59.00	0.6933	80.25	0.8012
	1	4	100	0.9852	95.00	0.8810	92.00	0.9049	75.67	0.8599	90.67	0.9078
	3	1	100	1.0000	100	0.9747	94.63	0.9649	97.67	0.9827	98.07	0.9806
	3	4	100	1.0000	98.67	0.9852	97.00	0.9848	100	1.0000	99.25	0.9925

TABLE I: Impact of macro- and macro- diversity on the accuracy of the predictions. ReWiS is trained in E_1 and is tested in E_2 and E_3.



(a) 3 rx, 1 antenna, 20 MHz, acc = 76.5

(b) 3 rx, 1 antenna, 80 MHz, acc = 92.75

(c) 3 rx, 4 antennas, 20 MHz, acc = 84.25

(d) 3 rx, 4 antennas, 80 MHz, acc = 98.85

Fig. 8: Impact of frequency resolution and diversity, with one receiver. E_2 is the target environment.

We notice that the baseline model is able to learn from the CSI data collected in the source environment E_1, however, is not able to make accurate predictions in target environments. This happens since the baseline classifier's goal is to classify the CSI measurement. However, the ReWiS learning framework tends to learn *how* to learn and achieves 35% better accuracy on average.

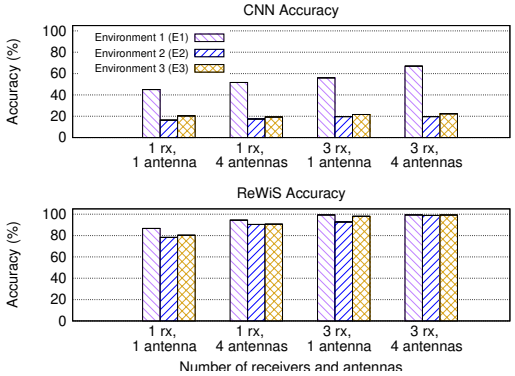


Fig. 9: CNN vs ReWiS accuracy performance.

4) Impact of windowing: Figure 10 compares the performance of the algorithm with different window sizes at 80 MHz channel. It can be seen that with a very small window size i.e., $W = 50$ the accuracy of the predictions is very low and it does not improve much by increasing the diversity. Specifically, Figure 10 shows that (i) by increasing the window size W, the number of receivers, and the number of antennas per receiver, the performance improves by up to 35%, 12% and 10%, respectively. However, with a higher number of antennas, the window size can be decreased and still achieve acceptable accuracy. It is worth noting that reducing the window size reduces pre-processing time and complexity.

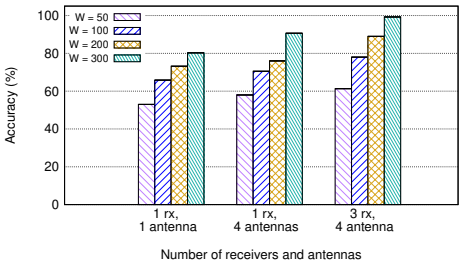


Fig. 10: Impact of diversity and window size.

VI. Conclusions and Future Work

In this paper, we have proposed ReWiS, a novel framework for robust and environment-independent Wi-Fi sensing. ReWiS leverages multi-antenna multi-receiver diversity to improve the overall robustness, and leverages a customized version of FSL to eliminate the need for application-specific feature extraction. ReWiS has been prototyped using off-the-shelf Wi-Fi equipment, and its performance has been showcased by considering a human activity recognition. We have performed an extensive data collection campaign, evaluated the impact of each diversity component on the performance, and compared ReWiS with a CNN-based approach. Experimental results have shown that ReWiS improves the performance by about 40% with respect to existing single-antenna low-resolution approaches, and increases accuracy by 35% with respect to a CNN when tested in different environments. As part of the novel contributions, we will release our 60 GB dataset and the entire code repository to the community. We believe that ReWiS improves the state of the art in

Wi-Fi sensing, by demonstrating superior generalization and robustness capabilities with respect to existing work. As part of future work, we are planning to test `ReWiS` in the presence of more complex classification tasks, for example, multiple human subjects or additional activities.

VII. Acknowledgment and Disclaimer

This work is funded in part by the AFRL Visiting Faculty Research Program (VFRP) contract number FA8750-20-3-1003 and National Science Foundation (NSF) grants CNS-2134973 and CNS-2120447. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the U.S. Government.

References

[1] U. Cisco, "Cisco annual internet report (2018–2023) white paper," *Online (accessed March 26, 2021) https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/whitepaper-c11-741490.html*, 2020.

[2] 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.

[3] O. Bejarano, E. W. Knightly, and M. Park, "IEEE 802.11ac: from Channelization to Multi-user MIMO," *IEEE Communications Magazine*, vol. 51, no. 10, pp. 84–90, 2013.

[4] H. Jiang, C. Cai, X. Ma, Y. Yang, and J. Liu, "Smart Home Based on WiFi Sensing: A Survey," *IEEE Access*, vol. 6, pp. 13317–13325, 2018.

[5] IEEE 802.11bf Task Group (TG), "IEEE 802.11bf (TGbf) Project Authorization Request (PAR)." https://tinyurl.com/TGbfPAR, 2021.

[6] M. Muaaz, A. Chelli, M. W. Gerdes, and M. Pätzold, "Wi-Sense: A Passive Human Activity Recognition System Using Wi-Fi and Convolutional Neural Network and Its Integration in Health Information Systems," *Annals of Telecommunications*, pp. 1–13, 2021.

[7] Z. Chen, L. Zhang, C. Jiang, Z. Cao, and W. Cui, "WiFi CSI Based Passive Human Activity Recognition Using Attention Based BLSTM," *IEEE Transactions on Mobile Computing*, vol. 18, no. 11, pp. 2714–2724, 2018.

[8] Z. Shi, J. A. Zhang, Y. D. R. Xu, and Q. Cheng, "Environment-Robust Device-free Human Activity Recognition with Channel-State-Information Enhancement and One-Shot Learning," *IEEE Transactions on Mobile Computing*, 2020.

[9] O. Vinyals, C. Blundell, T. Lillicrap, D. Wierstra, *et al.*, "Matching Networks for One Shot Learning," *Advances in Neural Information Processing Systems*, vol. 29, pp. 3630–3638, 2016.

[10] J. Snell, K. Swersky, and R. S. Zemel, "Prototypical Networks for Few-Shot Learning," *arXiv preprint arXiv:1703.05175*, 2017.

[11] H. Jiang, C. Cai, X. Ma, Y. Yang, and J. Liu, "Smart Home Based on WiFi Sensing: A Survey," *IEEE Access*, vol. 6, pp. 13317–13325, 2018.

[12] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-ID: Human Identification Using WiFi Signal," in *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 75–82, IEEE, 2016.

[13] L. Yao, Q. Z. Sheng, X. Li, T. Gu, M. Tan, X. Wang, S. Wang, and W. Ruan, "Compressive Representation for Device-free Activity Recognition with Passive RFID Signal Strength," *IEEE Transactions on Mobile Computing*, vol. 17, no. 2, pp. 293–306, 2017.

[14] W. Li, R. J. Piechocki, K. Woodbridge, C. Tang, and K. Chetty, "Passive WiFi Radar for Human Sensing Using a Stand-alone Access Point," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 3, pp. 1986–1998, 2020.

[15] S. Shi, Y. Xie, M. Li, A. X. Liu, and J. Zhao, "Synthesizing wider wifi bandwidth for respiration rate monitoring in dynamic environments," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pp. 181–189, IEEE, 2019.

[16] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "MultiSense: Enabling Multi-person Respiration Sensing with Commodity WiFi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 3, pp. 1–29, 2020.

[17] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free Your CSI: A Channel State Information Extraction Platform for Modern Wi-Fi Chipsets," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, pp. 21–28, 2019.

[18] S. Di Domenico, M. De Sanctis, E. Cianca, F. Giuliano, and G. Bianchi, "Exploring Training Options for RF Sensing Using CSI," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 116–123, 2018.

[19] F. Wang, W. Gong, and J. Liu, "On Spatial Diversity in WiFi-based Human Activity Recognition: A Deep Learning-based Approach," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2035–2047, 2018.

[20] Y. Zhuo, H. Zhu, H. Xue, and S. Chang, "Perceiving Accurate CSI Phases with Commodity WiFi Devices," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, pp. 1–9, 2017.

[21] Y. Ma, S. Arshad, S. Muniraju, E. Torkildson, E. Rantala, K. Doppler, and G. Zhou, "Location-and Person-Independent Activity Recognition with WiFi, Deep Neural Networks, and Reinforcement Learning," *ACM Transactions on Internet of Things*, vol. 2, no. 1, pp. 1–25, 2021.

[22] C. Xiao, D. Han, Y. Ma, and Z. Qin, "CsiGAN: Robust Channel State Information-based Activity Recognition with GANs," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10191–10204, 2019.

[23] Y. Zheng, Y. Zhang, K. Qian, G. Zhang, Y. Liu, C. Wu, and Z. Yang, "Zero-effort cross-domain gesture recognition with wi-fi," in *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 313–325, 2019.

[24] B. Bloessl, M. Segata, C. Sommer, and F. Dressler, "An IEEE 802.11 a/g/p OFDM Receiver for GNU Radio," in *Proc. of the Second Workshop on Software Radio Implementation Forum*, pp. 9–16, 2013.

[25] Y. Xie, Z. Li, and M. Li, "Precise Power Delay Profiling with Commodity Wi-Fi," *IEEE Transactions on Mobile Computing*, vol. 18, no. 6, pp. 1342–1355, 2018.

[26] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "FallDeFi: Ubiquitous Fall Detection Using Commodity Wi-Fi Devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 1–25, 2018.

[27] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaee, "A Survey on Behavior Recognition Using WiFi Channel State Information," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 98–104, 2017.

[28] L. Guo, L. Wang, C. Lin, J. Liu, B. Lu, J. Fang, Z. Liu, Z. Shan, J. Yang, and S. Guo, "Wiar: A Public Dataset for WiFi-based Activity Recognition," *IEEE Access*, vol. 7, pp. 154935–154945, 2019.

[29] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge university press, 2005.

[30] J. Xiong, K. Sundaresan, and K. Jamieson, "Tonetrack: Leveraging Frequency-agile Radios for Time-based Indoor Wireless Localization," in *Proc. of the 21st Annual International Conference on Mobile Computing and Networking*, pp. 537–549, 2015.

[31] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Elsevier, 2011.

[32] S. Ravi and H. Larochelle, "Optimization as a Model for Few-Shot Learning," 2016.

[33] Y. Tian, Y. Wang, D. Krishnan, J. B. Tenenbaum, and P. Isola, "Rethinking Few-Shot Image Classification: A Good Embedding Is All You Need?," in *Proc. of European Conference on Computer Vision (ECCV)*, pp. 266–282, Springer, 2020.

[34] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *Proc. of International Conference on Machine Learning (ICML)*, pp. 448–456, PMLR, 2015.

[35] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[36] I. Karmanov, F. G. Zanjani, S. Merlin, I. Kadampot, and D. Dijkman, "WiCluster: Passive Indoor 2D/3D Positioning using WiFi without Precise Labels," *arXiv preprint arXiv:2107.01002*, 2021.

[37] R. Rosen, "Linux Wireless-Linux Kernel Networking (4) - Advanced Topics," 2009.

[38] M. Mortimer, "iperf3 Documentation," 2018.

[39] J. Zhang, F. Wu, B. Wei, Q. Zhang, H. Huang, S. W. Shah, and J. Cheng, "Data Augmentation and Dense-LSTM for Human Activity Recognition Using WiFi Signal," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4628–4641, 2020.