

Unified Bundling of Wireless Sensing Models for Energy-Efficient Buildings Automation

Adolfo Bauchspiess^{a,1,*}, Tanmay Mane¹, Iris Walter¹, Jürgen Becker^b

^aLARA - Laboratório de Automação e Robótica ENE/FT/UnB - Universidade de Brasília, Brazil.

^bITIV - Institute for Information Processing, KIT - Karlsruhe Institute of Technology, Karlsruhe, Germany

Resumen

WiFi Sensing with CSI signals is wide known and can be successfully used for many indoor tasks, like crowd counting, localization, pose estimation, building automation and human activity recognition. WiFi CSI are all around through 802.11 OFDM devices while preserving the privacy of the occupants. Available IoT networks can be equipped with additional low cost sensing modules in order to provide additional services to environment users. Energy-saving in building automation is more than ever a relevant aspect. In this context we propose a bundled overlapped wireless sensor approach that supports and enhances many MIMO indoor HAR applications. We uses ESP32 with ESP-Now protocol. A network with 2 Senders and 6 receivers, each with a single antenna (12 links) was implemented and tested in different environments, locations and persons. With overlapped link bundles, the sharing of neighbor sensor information, allows higher precision with machine learning than the usual MIMO modelling. Accuracy greater than 60% could be reached with prototypical 3 bundles few-shot with gread domain diversity (environment, location, person).

Palabras Clave:

WiFi-Sensing, CSI, MIMO, Prototypical Few-Shot, Building Automation (from IFAC list).

1. Introduction

Ambient Intelligence delivers diverse services to the occupants of a given environment in an almost invisible way. Enabled by a wireless sensor-actuator network, access control, health monitoring, energy-saving through building automation, among others, can be offered. The IoT and Cloud integration expands greatly the possibilities, e.g., Guimaraes et al. (2019). The automatic power adjustment of heating and illumination with regard to the number of occupants, eventually zone-wise, and even accordingly to different activities, can save much energy. Taking faster knowledge of occupation/activity changes enables disturbance cancellation in a feed-forward thermal control loop, which further reduce power consumption. Reena et al. (2015) and Swaminathan et al. (2018).

It has been shown that occupancy estimation is very relevant in the context of energy-saving building automation. Silva, 2019 implemented a video curtain persons accounting system for a feed-forward HVAC thermal load adjustment, see

1.1. Feedforward thermal-load estimation

Feedforward control can be used to consider disturbances in advance so that their impact on the controlled variable is reduced. In the conventional feedback control strategy, disturbances $w(t)$ hit the process, deviating the controlled variable $c(t)$ from its set point. The negative feedback will bring $c(t)$ back to the desired value using the error information $e(t)$. The standard feedback topology implies a delay, because the disturbance has to walk through the process in order to be manifested by a significant error. In other words, the feedforward topology doesn't change reference response, determined by the feedback loop, but allows a much faster disturbance rejection. A feedback control loop enhanced by a feedforward block can be seen in figure 1.

1.2. ITIV/KIT Conference Room energy-saving potential

The present research aims the energy saving through occupancy-based building automation. To have a reference framework for the energy saving potential the Conference Room of ITIV/KIT will be considered.

Six fluorescent illumination zones can be switched individually. Heating is central commanded. Window can be opened manually for ventilation. Manual-driven blenders can adjust the desired natural lighting diffusion in the conference room. A Meeting Scheduler was installed recently.

*Contact Author

Correos electrónicos: adolfobs@ene.unb.br (Adolfo Bauchspiess), ulldt@studentkit.edu (Tanmay Mane), iris.walter@kit.edu (Iris Walter), juergen.becker@kit.edu (Jürgen Becker)

URL: <http://www.ene.unb.br/adolfo> (Adolfo Bauchspiess), [https://www.itiv.kit.edu/217587.php\(IrisWalter\)](https://www.itiv.kit.edu/217587.php(IrisWalter)),

¹Researcher on sabbatical year at ITIV/KIT, supported by the UnB, Brazil.

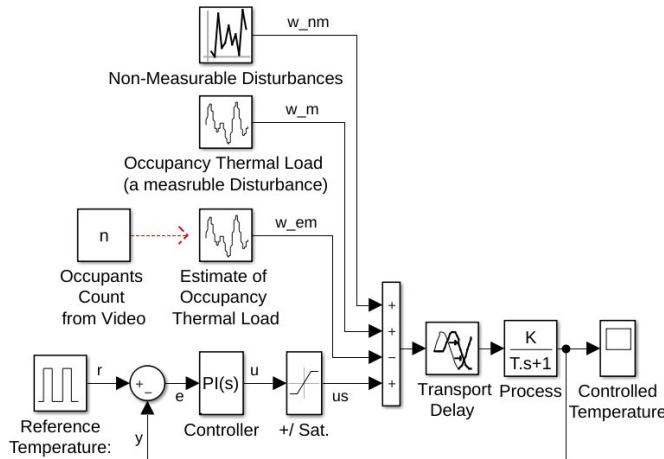


Figure 1: Feed-forward thermal Control with video occupants estimate. The thermal load changes are taken into account much earlier than by the feedback loop. The typical office thermal time constant is about 20 min while images are processed in milliseconds.



Figure 2: ITIV/KIT conference room as use case for Energy Efficiency. Large windows, south side, manual blinder, manual panel curtains. Central Heating. 4 switches for 4 zones with 2 lamps each.

Thermal and visual comfort should follow ergonomic standards. We will consider automation measures that favor energy saving.

Figure 3 illustrates a CSI coverage of the conference room Region of Interest with 8 ESP32 modules. Masters Tx1 and Tx2 ans 6 Rx sensing devices.

The main idea is to use low-cost single antenna ESP32 to reach a good spatial diversity (richer CSI channels, better Human Activity Recognition). A 4-antenna Sender and a 4-antenna Receiver cover with 16 links almost the same area, see Figure 8.

As known from Malvar's² overlapped Wavelets Transform, Vargas y Malvar (1993), the superposition of transform windows give valuable information "over the fence". In this work we implemented the '**overlapped bundles**', as will be detailed in the next sections. Majority vote classifiers, often as CNN's in

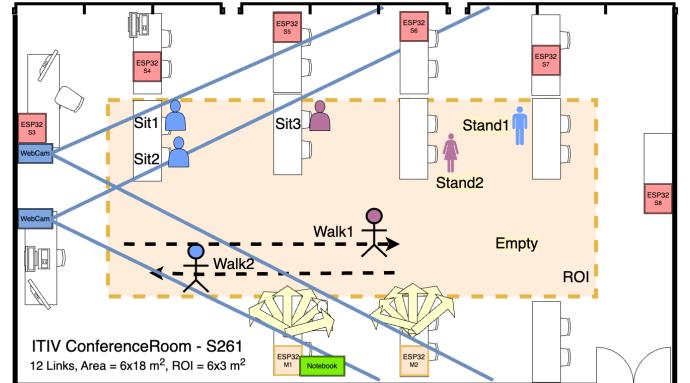


Figure 3: Energy efficiency WiFi-Sensing Sensor Placement at Conference Room ITIV/KIT. ROI can be covered by 12 links, ???. S261 - 2 Masters, 6 Sensors, 1 collecting device. 4K cameras for ground-truth and labeling.

machine learning, can benefit from overlapped bundles. Neighbor overlapped CSI links are shared by the CNN models enhancing accuracy.

Bundles CSI Taxonomy

The sensor physical network is defined as Smsc, with m-masters, s-sensors, c-collectors.

$$N = s.m \quad (1)$$

with *m*-masters, *s*-sensors, and *c*-collectors *N* links are build. The collector could be each sensing node, or one of the masters or one of the sensors. In the present work master M1, is also the collector (the module that is USB-connected to a notebook writes the CSI.csv file).

Most ESP32 papers use diverse CSI collectors, e.g., ?, use 4 Senders 4 Receiver and 4 collectors (S444, with *m*=4). Synchronization is a concern as man HAR captures human rapid movements. In this paper we use S261, *c* = 1 (1 collector starts the cycle with a broadcast and then collects CSI captured by the sensors. Synchronization is implicit). With 2 masters, *m*=2, two broadcast messages, from M1 and M2, are send in sequence. Only then the Sensors start sending both captured CSI carriers to the collector.

The bundles are links used in multi-model classifiers. In a given environment any available link can be associated to a bundle, and overlapping is welcome. For the 3 models majority vote we have links associated in Bundles by

$Bm_1m_2m_3 :$

$$\text{model } m_1 = [L_{11}, L_{12}, \dots L_{1m_1}],$$

$$\text{model } m_2 = [L_{21}, L_{22}, \dots L_{2m_2}],$$

$$\text{model } m_3 = [L_{31}, L_{32}, \dots L_{3m_3}].$$

were L_{ij} , $j = 1 : m_i$ are the links used in Bundle for model *i*. All *N* links available can compose a Bundle, be absent or appear more than once.

²Henrique "Rico" S. Malvar is a distinguished Brazilian engineer and a signal processing researcher at Microsoft Research. Malvar, UnB alumni, faculty member from 1979 to 1993, received 2022 the UnB Dr.-h.c. for his 'contributions to global multimedia technology'.

Table 1: Typical examples of "Task-Oriented" sensing Bundles. They enable many different HAR/CC/TLC/Zonal-based building automation strategies. Sensing, Actuators and Supervisory System should be integrated as IoT devices.

HAR			
Tx	Rx	TLC	HAR
1:2	3:4	M2	Walk
1:2	5:6	F1	Walk
1:2	6:7	M1	F2 Stand
1:2	7:8	M1	Empty
1:2	3:5	P1	M1 Sit M2 Sit Sit
1:2	4:6	P1	F3 Sit

CC			CC
Tx	Rx	TLC	
1:2	3:8	P1	P1 5 Persons
1:2	3:4	P1	P1 2 Persons Front
1:2	7:8	P1	P1 2 Persons Back

Zonal Automation				
Illumination (HCL)				
1:2	1:6	Loc2 Emppy	Light2 off	
Heating				
1:2	1:6	4 Pers. Loc1	Heating1 on	

Figure 4 illustrate the use of Bundles, B321 and B333 in a S161 physical sensor configuration (M1, L = S3,S9,S11,S12,S13,S15). Sample Bundles in S161:

- B321 (non-homogeneous):

$$m_1: 3 \text{ red links (S12.S13.S15)}$$

$$m_2: 2 \text{ green links (S9.S11)}$$

$$m_3: 1 \text{ blue link (S3)}$$

- B333 (with superposition):

$$m_1: 3 \text{ red Bundle (S12.S13.S15)}$$

$$m_2: 3 \text{ green Bundle (S11.S12.S13)}$$

$$m_3: 3 \text{ blue Bundle (S3.S9.S15)}$$

A great advantage of Bundles approach is that once the data is acquired or separable by links, the Bundles for different HAR experiments can chosen *a posteriori*.

The cross-domain scenario in 3 illustrates S261 physical sensor network for HAR sit, stand, walk, empty, Thermal Load Class medium(male), small (female) with CC, crowd-counting 0,1,2,3,4,5,6,7,8,9,10,11,12

In Figure 7 a) shows different occupation of 4 comfort zones. (iii) no occupancy, no light. (i) and (iii) at the window, need less light than (iv). On the right-side of the room, due to 6 people less heating is needed than on the left side, with 2 people. The specific activity estimation can further refine the thermal load estimation. With single Tx-Rx links, b), 6 pairs would cover the 4 comfort zones.

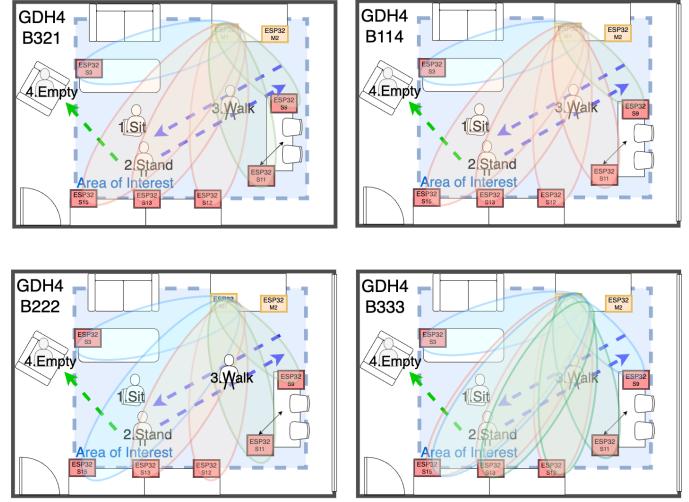


Figure 4: Bundles and Domain Diversity GDH/KIT.



Figure 5: Large researchers Room ITIV/KIT. Different furniture, but, "by construction", many similar energy-saving aspects.

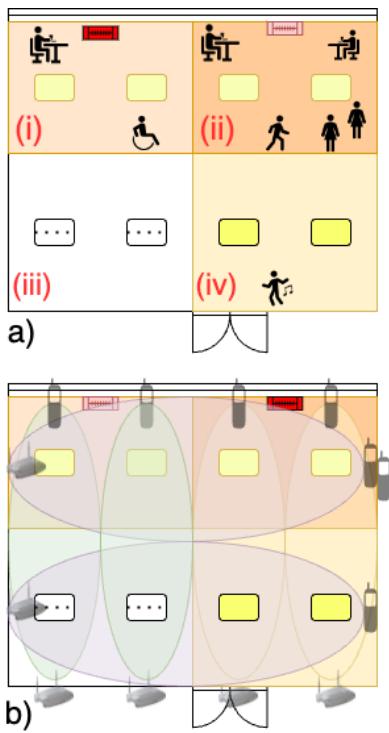


Figure 6: Large researchers room ITIV/KIT as use case for Energy Efficiency.

Possible automation enhancements:

- Movement/Presence detector
- Radiators with individual thermostat
- Radiators with IoT connection
- Counting people at the door
- Counting people by zones
- HAR by zones
- Identify individual thermal load classes

2. Related Work

2.1. WiFi Sensing

2.2. Few-Shot Transfer Learning

Occupancy

Aftab et al. (2017) shows the importance of using model predictive control with real-time occupancy recognition for HVAC. Video was used with machine learning. Raspberry Pi 3 implementation. Dong et al. (2018) also deals with occupancy for HVAC energy-efficient builds. A Short-term prediction algorithms are proposed using "classical" approaches: expectation maximization model, finite state automata and stochastic model based on uncertain basis functions. Nweye y Nagy (2021) propose the use of a "smart meter estimation" with WiFi Sensing and Clustering for HVAC Schedules and Energy Saving.

CSI HAR

Zhou et al. (2020) uses an attention model to measure Human Activity from WiFi. As interesting use case, thes phases of "drinking water" (Standing, Lifting, Holding, Pouring, Holding, Dropping, Standing) are recognized. LSTMs with attention are trained. Ma et al. (2021) present Wi-Fi Location-and Person-Independent HAR. Deep Neural Networks, and Reinforcement Learning are employed. Here, (table 7) a list of algorithms used recently for Activity Recognition with CSI is very helpful: kNN, SVM, Sparse Auto-Encoder, CNN, SOM, HMM DTW, Decision Tree, Logistic Regression, Finite Automata, Expert Models (Naive Bayes, Random Forest), Sparse Representation Classifier. Yang et al. (2021) propose CSI Signal Enhancement for Human Activity Recognition. This is indeed a critical aspect in the CSI use for HAR. On one side is the End-To-End approach where the raw signals are presented and the machine learning is supposed to 'extract' the relevant information and on the other side is the idea the pre-processing the signal can enhance the classification result. In Choi et al. (2022), for example, only features extracted from CSI are presented to the machine learing system. Maybe a middle-term is the best, some noise-reduction is OK. But we like to preserve the cross-domain capability. To much pre-processing will "overfit" to the training environment. Bahadori et al. (????) developed ReWiS a Few-Shot for Multi-Antenna Multi-Receiver CSI learning system.

Prototypical CNN are trained in one environment (features) and then the classification is adapted with few shots to 2 different environments. Activities are empty, jump, stand, walk. Accuracy for environment E2 for varies between 59.75 % for 1Tx-1Rx 1 antenna, 20 MHz and 98.85 % for 1Tx-3Rx

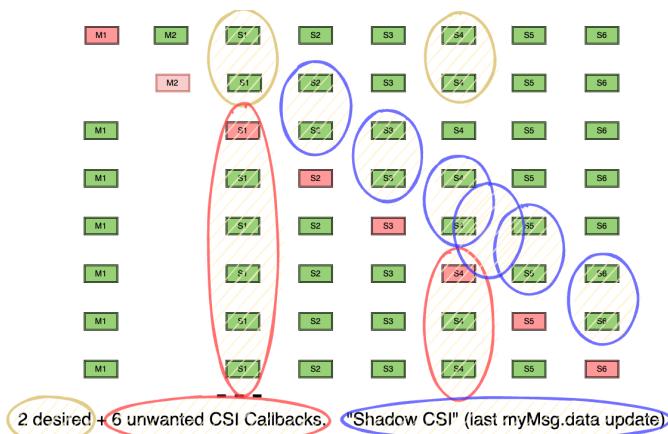


Figure 7: Room 1.26 ITIV/KIT as use case for Energy Efficiency.

3 antennas, 80 MHz. (1 link and 48 links, respectively). 242 subcarriers are collected by the NIC cards for each link used. (Tx = Netgear R7800 Router; 3x Rx = Asus RT-AC86U,5 GHz, Nexmon tool CSI collection).

2.3. Prototypical Few-Shot HAR Transfer Learning

Few-Shot Glossary, based on "Tutorial - Everything you need to know about Few-Shot Learning" (blog.paperspace.com/few-shot-learning by Rohit Kundu, accessed 6th March 2023)

Support Set: The support set consists of the few labeled samples per novel category of data, which a pre-trained model will use to generalize on these new classes.

Query Set: The query set consists of the samples from the new and old categories of data on which the model needs to generalize using previous knowledge and information gained from the support set.

Embedding Space The space in which the features are represented. The classification can be made by a similarity function (e.g., the euclidean distance) to the support novel classes.

Metric-Level FSL approaches aim at learning a distance function between data points. Features are extracted from images, and the distance between the images is computed in the embedding space. This distance function can be Euclidean distance, Earth Mover Distance, Cosine Similarity-based distance, etc. This is something we have covered while discussing Siamese networks.

One-Shot Learning. The face recognition technology used in modern smartphones uses One-Shot Learning.

Surveys are always a good starting point to start a research. In this work three surveys were considered valuable, in chronological order: i) Yousefi et al. (????), an early (2017) survey on Behavior Recognition Using WiFi Channel State Information. ii) Chen et al. (????), (2022) a survey on Cross-Domain WiFi Sensing with CSI and iii) Xiao et al. (2016), a thoroughly compilation of Application Scenarios, Current Solutions, and Open Issues.

Figure ?? illustrates the prototypical few-shot methodology. A machine learning model, e.g. CNN, is used to capture the non-linear feature map using a large data set (small labeled/colored circles), as illustrated in a). The prototypes (larger circles) represent the n-way resulting classifier. b) For a new Environment E2, few-shots (5 labeled circles for each class here), are used to adapt the prototypes to the new environment. c) Queries are classified by the shortest distance to the available prototypes. d) A new environment E3, brings renewed prototypes and could give even a different classification. In figure d) the different color used for the left most prototype, instead of green, should illustrate a completely new few-shot class, never trained, 'adopted' by transfer learning mechanism.

Few-Shot CSI HAR State of the Art (March, 2023)

The great success of Few-Shot Machine Learning among other meta-learning algorithms (see Li et al. (2023) for a wide-used taxonomy of the field) have inspired many researchers. In

particular different ways to combine prototypes. Yang et al. (2020) propose Prototype Mixture Models for Few-shot Semantic Segmentation. Jiang y Cheng (2021) propose a Mixture of Gaussian Prototypes, while Allen et al. (2019) had proposed Infinite Mixture of Prototypes for Few-Shot Learning. And Ding et al. (????a) advocates for Few-Shot learning with Big Prototypes. Moreover Ding et al. (2022) propose the use of Hypersphere Modelling of Prototypes while Cui et al. (????) propose Continual Learning with Contrastive Mixture of Adapter for Few-Shot learning. A good recently published review paper on Deep metric learning is given by Li et al. (2023).

3. Bundled CSI Sensing

Using various CSI sensors a larger area can be covered (spatial diversity), considering the Fresnel zones as shown in Figure ?? . The technological challenge is to synchronize the acquisitions. By rapid HAR activities, as walking or jumping, it is crucial that the CSI capture almost the same moment. There are many good results about machine learning, however, for the practical use in building automation it is necessary that a deep trained model be quickly adapted to different rooms with different furniture, Aftab et al. (2017), Nweye y Nagy (2021). In this context meta-learning and in particular few shot learning seems to be the most promising approach, Jiang et al. (2018), Ding et al. (????b), Bahadori et al. (????),

Domain Diversity: Learn Once - Works (almost everywhere). Use it in a different Environment, different Location, different Person, etc. The configuration of the wireless sensors and the localization of the activities have a great impact on the captured CSI signals. In the first image of Fig. ?? we see red link (M1-S9) and blue link (M2-S15) would be greatly affected by the 'stand' activity. Standing on other positions, middle figure, will not be good represented by link (M1-S9). Walking, as seen in last figure of Fig. ?? is also strongly dependent of the ESP32 coverage.

3.1. Covering the Area of Interest - Fresnel Zones

As mentioned device-free WiFi-Sensing is very dependent on the spatial distribution of Transmitters, Receivers the location/orientation of the human subject.

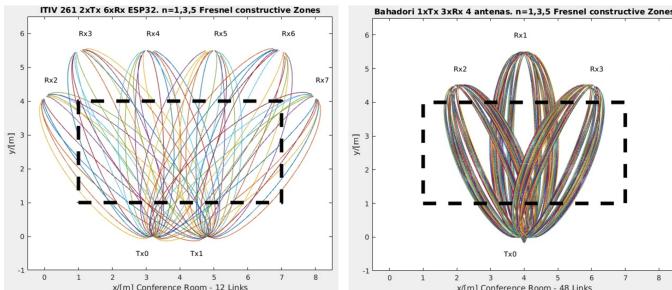
In Bahadori et al. (????) the antennas are attached to the access point and to the stations, giving less spatial possibilities. With Bundled CSI we have less links but a better spatial coverage, as can be seen in Fig. 8.

The approach adopted by Bahadori et al. (????) and here are quite different. Some keystones are listed behind the Fresnel images.

Inspired by Bahadori et al. (????) that uses 3 different "Monitors" to combine the 4 antennas AP and 4 Antennas of each STA in three Few-Shot models and then uses majority voting to decide the activity, we propose a "bundling" strategy for the ESP32 modules. See Figs. ?? and ??.

Training [GDH, L1, Pa, S221, B321].

Bundles 321, 114, 222 and 333. Were 321 stands here for a Bundle of 3 red links (S12.S13.S15) 2 green links (S9.S11)



12 Links M1:M2 S3:S8
1 Antenna
52 subcarriers
 $34 \times 264 \Rightarrow$ PCA 34×34 (Amp/Pha)
40 MHz
2x Tx ESP32 2.4GHz
6x Rx ESP32

48 Links. Tx0. Rx1:3
4 Antennas
242 subcarriers
 \Rightarrow PCA 242×242 (Amp)
80MHz BW
1x Tx Netgear R78006 STA
3x Tx Asus RT-AC86U WiFi routers

Figure 8: Floor plans for the different location of the activities

Tabla 2: Labeled activities of the CSI data acquisitions.

t s	$0 \rightarrow 10s$	$\rightarrow 20s$	$\rightarrow 30s$	$\rightarrow 40s$	$\rightarrow 50s$	$\rightarrow 60s$
0 min	Sit	Sit	Sit	Stand	Stand	Stand
1 min	Walk	Walk	Walk	Empty	Empty	Empty
2 min	Empty	Empty	Empty	Walk	Walk	Walk
3 min	Stand	Stand	Stand	Sit	Sit	Sit
4 min	Stand	Stand	Stand	Walk	Walk	Walk
5 min	Empty	Empty	Empty	Empty	Empty	Empty
6 min	Walk	Walk	Walk	Stand	Stand	Stand
7 min	Sit	Sit	Sit	Sit	Sit	Sit

and 1 blue link (S3). 333 stands for a bundle configuration with superposition: red Bundle (S12.S13.S15), green Bundle (S11.S12.S13), and blue Bundle (S3.S9.S15).

The six links to M1 are show. The six links to M2 are omitted, for clarity.

Summarizing, during the 8 minutes of data acquisition we have 48 Frames of 10 s comprising 17 CSI lines per Frame. The activities are balanced: 12 Sit, 12 Empty, 12 Stand and 12 Stand labels. Figure 9 show 4 pictures captured from the ground truth video at GDH 1 (Gastdozentenhaus location 1 - data training acquisition).

Figure ?? shows the 221 sensor arrangement at the ITIV, Engesserstr.5, room 1.26, used as test scenario.

4. ESP-Now MIMO CSI Sensing

ESP-Now, a Espressif developed MAC connection protocol use the PHY and MAC layer from 802.11n simplifying the setup when compared with the standard WiFi connection AP-STA with user and password checking. With ESP-Now the connection is immediately available after power-on and was so chosen for the experiments. The ESP-Now protocol allows the transmission of 250 useful bytes and so the core idea was so use two masters sending a broadcast signal in a short time interval. The CSI from each master will be stored in each sensor node. In the CSI261 configuration (2 Masters, 6



Figure 9: Activities Captured from Ground Truth Labeling Video: [GDH, L1, Pa, S261] - Sit, Stand, Walk, Empty.

Sensors, 1 Collecting node) 12 CSI channels are almost synchronously captured. After this phase each sensor transmits its CSI (M1, M2 channel state information) to master 1, which save a new line in the CSI file. See Figure 10

In the practical realization it was noted that using a **telecommunication device** (data integrity is more important, than eventual data delays) as a **sensing device**, where the synchronous acquisition is critical for correct activity recognition, was a big issue. The ESP-32 modules do not expect ACK for broadcast signals. In the MAC layer the events occur with different priorities (e.g., Rx CSI \wedge CSI \wedge Tx \wedge CallBack Rx \wedge Call Back Tx \wedge Print Msg on Terminal) and enable corresponding tasks. Each task, when active, tries to process old data, re-sending data so many times, see Fig. ???. The consequence is that code gets many checkings for "old data". The CSI acquisition is thus only enabled for a short window. Unfortunately, some data is also missed. A fixing routine is needed to clean the acquired data.

The consequence from ?? is that a much greater Cycle is needed (≈ 500 ms). Without collisions 20 ms was expected at the project start (6×2.2 .ms by 1Mb/s for ESP-Now).

5. CSI signal fixing and interpretation

Each CSI line comprises Imaginary and Real part of 6 Sensors in response to 2 Masters. Thus we have 54 carriers leading to $54 \times 12 = 648$ samples. Removing the pilot("guard") frequencies we have 624 useful samples. With N10s = 102, 102 lines of the .csv file are collected. As the 6 sensor (M1,M2) are concatenated for a common broadcast/sampling time, we have images with $102/6 = 17$ lines. Imaginary and Real Parts are separated in two images (see Figure 13). They are then mapped in Amplitude and Phase and combined in a single image with 34 lines. Real part from 0 to 16 and imaginary part from 17 to 33 (lower, darker part in the 34 x 624 Amplitude/Phase image). To compress the information, reducing redundancy, a PCA routine is used with 34 components, resulting in the standard 34 x 34 image used for the classification task.

Table 14 summarizes the sizes of the images used in the few-shot task.

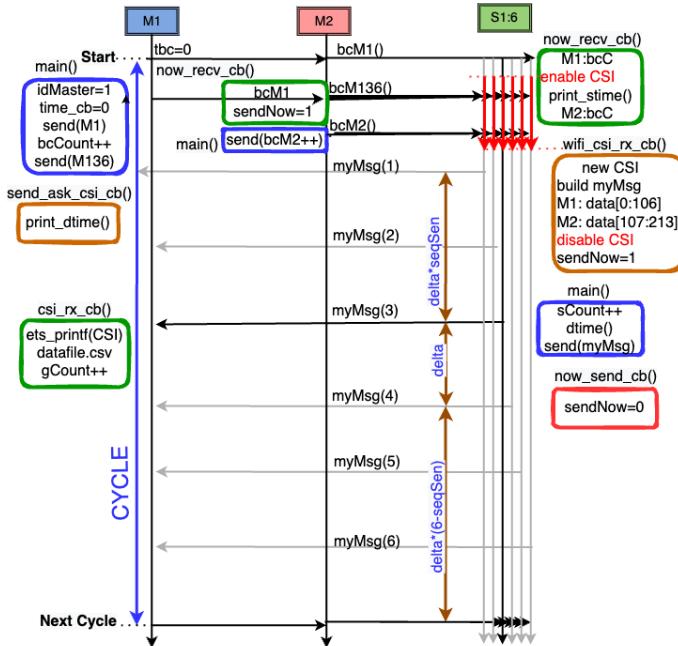


Figure 10: Time flow of the CSI261 acquisition. M1 starts the process and collects the CSI signals. M2 fires short after receiving the broadcast from M1. A Time Division Multiple Access gives each sensor a preferred window to transfer its data to M1. Call

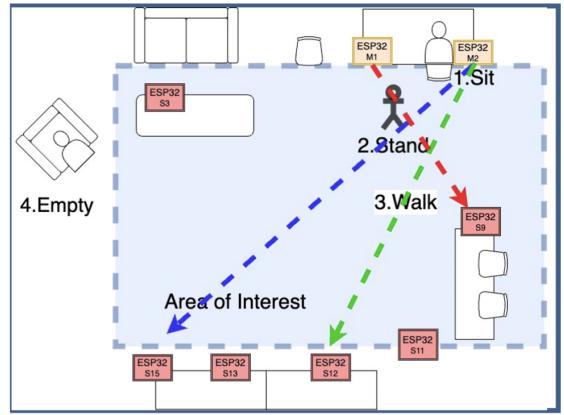


Figure 12: Environment GDH, Location 1 where the HAR training data were acquired. S261 sensor arrangements: 2 Masters (M1, M2) close to the collecting notebook. 6 sensors S3, S9, S11, S12, S13, S15. 1 Collecting Notebook (near 1.Sit). The CSI signal richness (HAR identifiable) in the Region of Interest depend on the activity location.

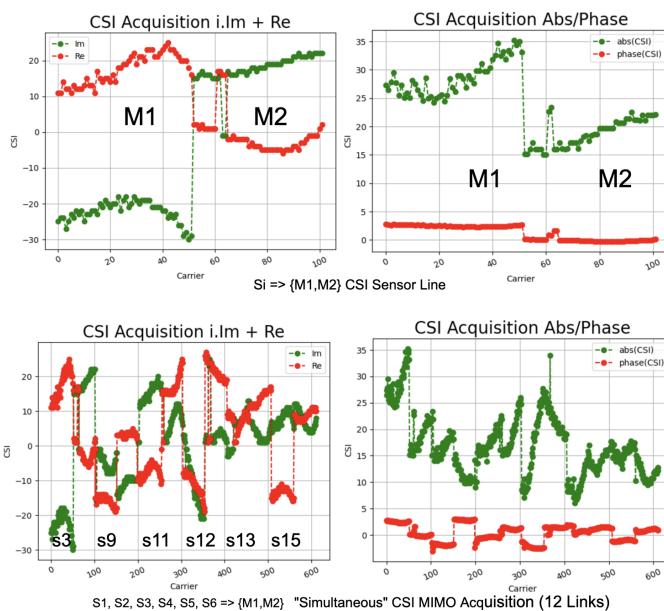


Figure 11: Acquisition signals in Sensor Configuration S261 (2 Masters, 6 Sensors, 1 collecting device.: Above CSI Real/Img and Amplitude/Phase. Below: one complete CSI cycle, correponding to the flowchart of figure 10

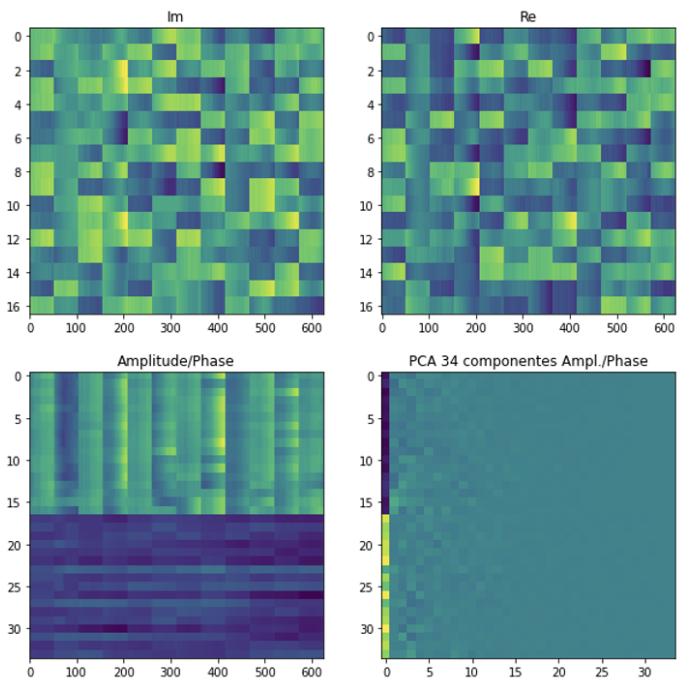


Figure 13: One sample "CSI image" used as input to the training/testing few-shot classification.

PCA Frame 34x34	Train	Test
Masters	2	2
Sensors	6	6
Amp/Phase	2	2
Carriers w/o Pilots	52	52
CSI Lines	17	17
Experiments	9	1
Frames/Exp.	47	42
Total Frames	423	42
Points/Carrier	7191	714
Points/Link	373932	37128
Total A/Ph points	8974368	891072
Total PCA points	1061208	117912

Figure 14: PCA Images are 34x34. 423 Frames are used for training and 42 for testing in the different scenarios. A total of 1.061.208 points are used in the training. 117.912 for testing.

2 Bundles (S3:S9)	S3(M1:M2)	S9(M1:M2)	PCA comp:34	carriers	Compression
			6 Bundles	624	94.55%
17Amp	104 Carriers	104 Carriers	4 Bundles	416	91.83%
			2 Bundles	208	83.65%
			1 Bundle	104	67.31%
6 Bundles (S3:...:S9)	S3(M1:M2)	S9(M1:M2)	S11(M1:M2)	S12(M1:M2)	S13(M1:M2)
			104 Carriers	104 Carriers	104 Carriers
17Phase	104 Carriers	104 Carriers	S11(M1:M2)	S12(M1:M2)	S13(M1:M2)
			104 Carriers	104 Carriers	104 Carriers
S15(M1:M2)	104 Carriers	104 Carriers	104 Carriers	104 Carriers	104 Carriers
			104 Carriers	104 Carriers	104 Carriers

Figure 15: Images used in the experimental validation.

Table 15 illustrates the CSI image sizes with 2 and with 5 Bundles. The PCA compressed images always have 34 x 34, so the compression rate varies, depending on the bundles used. All carriers from M1 to S3 are depicted in 16 for full acquisition cycle (sit, stand, walk, empty). To enhance the readability the first 5 carriers amplitude and phase are shown in separate. It is interesting to note that the first 5 carriers (shown separately) present a similar behaviour while considering all 52 carriers the amplitude show a considerable spread over the carriers. The phase spread is significantly smaller.

In Figure 16 5 carriers and all 52 useful carriers are depicted, to facilitate the understanding. Neighbor carriers show more similarity. Over the 52 carriers a higher spread can be recognized (which is positive for activity characterization as different carriers "see" the environment by its own "lenses"). In Figure ?? we can recognize "indicators" of a relation between CSI behavior and Human Activity:

"Expert" CSI Knowledge:

- Rule 1 - "sit" has higher CSI values
- Rule 2 - "stand" has lower CSI values

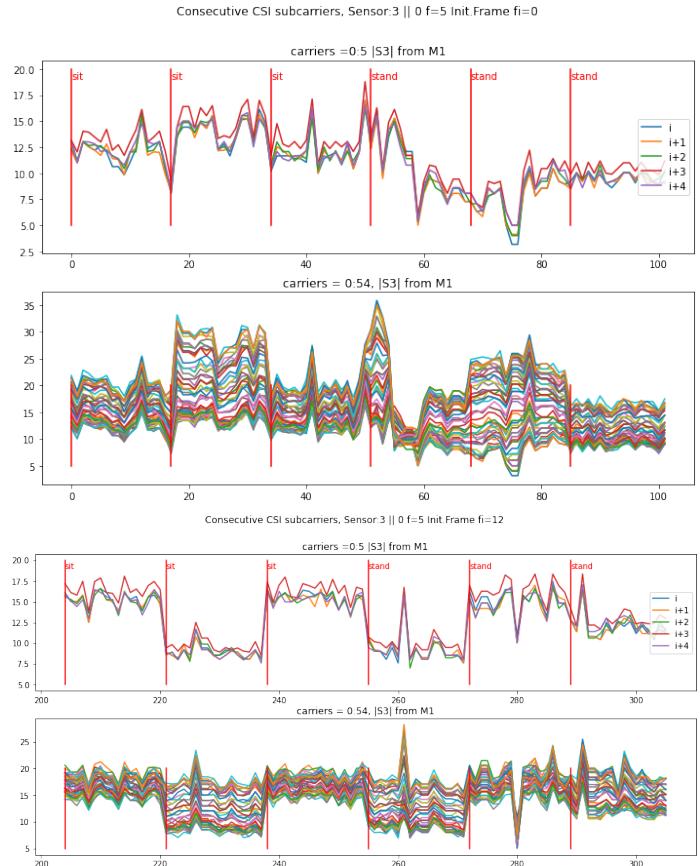


Figure 16: Typical time-domain CSI amplitude signals for link S3-M1, showing a 2 min acquisition. Activity collected over 10 s produce an CSI Image (Im, Re). 30 s HAR activity are labeled as (sit, stand, walk, empty...)

Rule 3 - "walking" has higher CSI variance
 Rule 4 - "empty" has the highest CSI - "free sight"

6. Bundled Esp-Now Few-Shot HAR Evaluation

In this section the evaluation methodology of "Few-Shot CSI Bundles with Superposition HAR" project at ITIV will be presented. The python implementation is available at the ITIV/KIT Karlsruhe gitlab: redFev23_1_SBundle_FewS.ipynb. (Last version: Februar 28th, 2023). The purpose is to test in different environments (spatial diversity) with different bundles.

ACRONYMS user for few-shot evaluation:

Few-Shot Domain Diversity:

Cross-Environment (GDH, ITIV)
 Cross-Location (1,2,3,4, ITIV_AB/IT, ITIV_2AB),
 Cross-Person (a,e,i),
 Run (i,ii,iii).

HAR, 4 Activities:

Sit
 Stand
 Walk
 Empty

Sensor Conf.

CSI S261 acquisitions: 2xTx - 6xRx, TDMA: 100ms, CSI acq.

Cycle: 750 ms.

CSI S221 acquisition: 2xTx - 2xRx, TDMA: 60ms, CSI acq.

Cycle: 600 ms

Bundles:

B221 = (S9).(S11.S13)(S3)
 B321 = (S12.S13.S15)(S9.S11)(S3)
 B222 = (S3.S9)(S11.S12)(S13.S15)
 B411 = (S3.S9.S11.S12)(S13)(S15)
 B332 = (S12.S13.S15)(S11.S12.S13)(S3.S9)
 B333 = (S12.S13.S15)(S11.S12.S13)(S3.S9.S15)
 B432 = (S11.S12.S13.S15)(S11.S12.S13)(S3.S9)

Few-shot configuration:

n_way = 4, n_support = 5, n_query = 1, test_episode = 300

x_dim = (1, 34, 34), hid_dim = 32, z_dim = 64

ProtoNet: CNN network to learn the prototypes.

```
def conv_block(in_channels, out_channels):
    return nn.Sequential(
        nn.Conv2d(in_channels, out_channels, 3, padding=1),
        nn.BatchNorm2d(out_channels),
        nn.ReLU(),
        nn.MaxPool2d(2)
    )
encoder = nn.Sequential(
    conv_block(x_dim[0], hid_dim),
    conv_block(hid_dim, hid_dim),
    conv_block(hid_dim, hid_dim),
    conv_block(hid_dim, z_dim),
    Flatten()
)
```

Training:

Table 3: Test accuracy of BundledS3_KIT GDH over 3 runs(i,ii,iii).

	S261 GDH-a			
Run/	B321	B222	B411	B333
2-i	0.611	0.522	0.573	0.620
2-ii	0.498	0.528	0.516	0.576
2-iii	0.545	0.588	0.618	0.610
avg	0.551	0.546	0.569	0.609
σ	0.057	0.036	0.051	0.023
3-i	0.526	0.569	0.537	0.639
3-ii	0.593	0.594	0.491	0.588
3-iii	0.623	0.598	0.539	0.681
avg	0.580	0.587	0.522	0.636
stdev	0.050	0.016	0.027	0.046
4-i	0.523	0.608	0.566	0.560
4-ii	0.580	0.508	0.577	0.545
4-iii	0.564	0.585	0.656	0.558
avg	0.556	0.567	0.599	0.554
σ	0.029	0.052	0.049	0.008

	S261 GDH-e			
Run/	4-i	4-ii	4-iii	avg
4-i	0.536	0.585	0.509	0.655
4-ii	0.494	0.607	0.608	0.615
4-iii	0.606	0.668	0.675	0.595
avg	0.545	0.620	0.597	0.626
σ	0.056	0.043	0.083	0.031

```
max_epoch = 2, epoch_size = 1000
test_envs=['3a','4a','4e','ITIV','ITIV_IT','ITIV_AB','ITIV_2AB']
bundles = ['121','321','411','332','333','432']
```

Train: 9 csv files: each 7m 50s, 47 Frames
 Environment/Location/Person/Sensors:

'261_10_75': GDH 1a
 collected Feb. 4th-5th, 2023

Test: 1 csv file: 7 m, 42 Frames
 Environment/Location/Person/Sensors:
 '261_10_75': '1a','2a','3a','4a','4e',
 'ITIV_AB','ITIV_IT','ITIV_2AB',
 '261_6_60' : 'ITIV'

Main Loop:

```
for run in runs:
    for test_env in test_envs:
        for bundle in bundles:
            read (train_data) for given [bundle, test_env]
            train (train_data) for given [n_way,n_support,n_query, epoch]
            read (test_data) for given [bundle, test_env]
            CF, acc = test (bundle1, bundle2, bundle3)
```

Tabla 4: Test accuracy of BundledS3_KIT ITIV over 3 runs(i,ii,iii).

S261 ITIV-a				
Run/	B321	B222	B411	B333
1-i	0.323	0.361	0.285	0.375
1-ii	0.328	0.358	0.318	0.433
1-iii	0.392	0.366	0.303	0.378
avg	0.348	0.361	0.302	0.395
σ	0.038	0.004	0.016	0.033
2-i	0.576	0.585	0.566	0.545
2-ii	0.580	0.584	0.602	0.602
2-iii	0.645	0.5658	0.6108	0.6133
avg	0.600	0.578	0.593	0.587
σ	0.039	0.011	0.024	0.037
S261 ITIV-i				
1-i	0.294	0.323	0.359	0.342
1-ii	0.244	0.408	0.388	0.326
1-iii	0.254	0.353	0.293	0.372
avg	0.264	0.361	0.347	0.347
σ	0.026	0.043	0.048	0.023
S221 ITIV-a				
1-i	0.411	0.350	0.314	0.409
1-ii	0.378	0.390	0.194	0.468
1-iii	0.309	0.381	0.281	0.376
avg	0.366	0.374	0.263	0.418
σ	0.052	0.021	0.062	0.047

7. Results

Three complete runs (-1,-2 and -3) with Locations GDH (1a,2a,3a,4a,4e) and ITIV in Bundle 321, 222,411 and 333 have stressed the proposed few-shot approach. From the 4 table some interesting conclusions can be reached.

- **Bundle** - The Bundle with Superposition "333" brings the best test results. Only configuration GDH 4a is better with "411". See the blue highlighted values in the table.
- **Location** - ITIV is a the "very different" test environment. Different furniture in different positions. Sensors in a different arrangement. The greatest difference is however the sampling used for ITIV: 600 ms Cycle with 69 ms TDMA interval. All other signals were collected with 750 ms and 100 ms TDMA.
- **Gender** - Acquisition 4e was done by a woman. Despite the best evaluation for this location was reached with "411", with an accuracy of 0.675, the best average, with 0.6217 was reached with the 333 bundle. All other acquisitions were taken with the same male test person.
- **Best Result** - In an unexpected way the best result (0.6808) and the best average (0.6361) were obtained with bundle "333" at location GDH 3a. It is not clear why testing 1a gives not the best learning. The test data set was not used for training but was obtained in the very training position, 1a!

8. Conclusions & Perspectives

The summary of all tests can be seen in Figure ???. A collection of CSI HAR results obtained in many different configurations can be seen in first Appendix. Some remarks (most already expected):

- 5-Shot best result (0.642) is better than 1-shot best result (0.604).
- Highest accuracy was obtained in a different location (4a, 43) at the training environment GDH.
- Best result in a different environment was ITIV_2AB with 5-Shot 1-Query.
- A female test person gives a small accuracy drop in environment GDH. Another female test person at ITIV_IT is more difficult to classify by Few-Shot.

References

- Aftab, M., Chen, C., Chau, C. K., Rahwan, T., 2017. Automatic HVAC control with real-time occupancy recognition and simulation-guided model predictive control in low-cost embedded system. *Energy and Buildings* 154, 141–156.
URL: <https://doi.org/10.1016/j.enbuild.2017.07.077>
DOI: 10.1016/j.enbuild.2017.07.077
- Allen, K. R., Shelhamer, E., Shin, H., Tenenbaum, J. B., 2019. Infinite mixture prototypes for few-shot learning.
- Bahadori, N., Ashdown, J., Restuccia, F., 1 ???? Rewis: Reliable wi-fi sensing through few-shot multi-antenna multi-receiver csi learning.
- Chen, C., Zhou, G., Lin, Y., ????
- Choi, H., Fujimoto, M., Matsui, T., Misaki, S., Yasumoto, K., 2022. Wi-cal: Wifi sensing and machine learning based device-free crowd counting and localization. *IEEE Access* 10, 24395–24410.
DOI: 10.1109/ACCESS.2022.3155812
- Cui, Y., Yu, Z., Cai, R., Wang, X., Kot, A. C., Liu, L., 2 ???? Generalized few-shot continual learning with contrastive mixture of adapters.
- Ding, N., Chen, Y., Cui, G., Wang, X., Zheng, H.-T., Liu, Z., Xie, P., 11 2022. Few-shot classification with hypersphere modeling of prototypes.
URL: <http://arxiv.org/abs/2211.05319>
- Ding, S., Chen, Z., Zheng, T., Luo, J., ????.a
- Ding, X., Jiang, T., Zhong, Y., Huang, Y., Li, Z., ????.b
- Dong, J., Winstead, C., Nutaro, J., Kuruganti, T., 2018. Occupancy-based HVAC control with short-term occupancy prediction algorithms for energy-efficient buildings. *Energies* 11 (9), 1–20.
DOI: 10.3390/en11092427
- Guimaraes, V. G., de Moraes, R. M., Obraczka, K., Bauchspies, A., 2019. A novel iot protocol architecture: Efficiency through data and functionality sharing across layers. 28th International Conference on Computer Communication and Networks (ICCCN), Valencia, Spain, pp. 124–127.
- Jiang, R., Cheng, Z., 2021. Mixture gaussian prototypes for few-shot learning. Vol. 2021–December. IEEE Computer Society, pp. 902–908.
DOI: 10.1109/ICDMW53433.2021.00118
- Jiang, W., Miao, C., Ma, F., Yao, S., Wang, Y., Yuan, Y., Xue, H., Song, C., Ma, X., Koutsonikolas, D., Xu, W., Su, L., 10 2018. Towards environment independent device free human activity recognition. Association for Computing Machinery, pp. 289–304.
DOI: 10.1145/3241539.3241548
- Li, X., Yang, X., Ma, Z., Xue, J. H., 6 2023. Deep metric learning for few-shot image classification: A review of recent developments.
DOI: 10.1016/j.patcog.2023.109381
- Ma, Y., Arshad, S., Muniraju, S., Torkildson, E., Rantala, E., Doppler, K., Zhou, G., 2021. Location- and Person-Independent Activity Recognition with WiFi, Deep Neural Networks, and Reinforcement Learning. *ACM Transactions on Internet of Things* 2 (1), 1–25.
DOI: 10.1145/3424739

- Nweye, K., Nagy, Z., 2021. MARTINI: Smart Meter Driven Estimation of HVAC Schedules and Energy Savings Based on WiFi Sensing and Clustering, 0–13.
 URL: <http://arxiv.org/abs/2110.08927>
 DOI: 10.1016/j.apenergy.2022.118980
- Reena, K. E. M., Mathew, A. T., Jacob, L., 2015. An occupancy based cyber-physical system design for intelligent building automation 2015, 1–15.
- Swaminathan, S., Wang, X., Zhou, B., Baldi, S., 2018. A university building test case for occupancy-based building automation 11, 1–15.
- Vargas, L. F. C., Malvar, H. S., 1993. Elt-based wavelet coding of high-fidelity audio signals. IEEE International Symposium on Circuits and Systems (ISCAS), pp. 124–127.
- Xiao, J., Zhou, Z., Yi, Y., Ni, L. M., June 2016. A survey on wireless indoor localization from the device perspective. ACM Computing Surveys 49,
- 1–25.
- Yang, B., Liu, C., Li, B., Jiao, J., Ye, Q., 8 2020. Prototype mixture models for few-shot semantic segmentation.
 URL: <http://arxiv.org/abs/2008.03898>
- Yang, J., Liu, Y., Liu, Z., Wu, Y., Li, T., Yang, Y., 2021. A Framework for Human Activity Recognition Based on WiFi CSI Signal Enhancement. International Journal of Antennas and Propagation 2021.
- DOI: 10.1155/2021/6654752
- Yousefi, S., Narui, H., Dayal, S., Ermon, S., Valaee, S., ????
 Zhou, C., Liu, J., Sheng, M., Li, J., dec 2020. Hybrid RSS/CSI Fingerprint Aided Indoor Localization: A Deep Learning based Approach. En: 2020 IEEE Global Communications Conference, GLOBECOM 2020 - Proceedings. Vol. 2020-Janua. Institute of Electrical and Electronics Engineers Inc.
 DOI: 10.1109/GLOBECOM42002.2020.9348190