Car-Sense: Vehicle Occupant Legacy Hazard Detection Method Based on DFWS

Zhanjun Hao, Guowei Wang * and Xiaochao Dang

Abstract: Casualties caused by people trapped in cars have been common in recent years. Despite a variety of solutions, complex detection devices need to be arranged, or privacy is poor. Since device-free Wi-Fi sensing has attracted much attention due to its simplicity, low cost, and no need for additional hardware, this paper proposes a contactless wireless Wi-Fi sensing-based method for detecting people left in cars: Car-Sense. The method uses ESP32 devices in the vehicle to build a wireless Wi-Fi network for low-cost, real-time, and accurate personnel awareness. By capturing and analyzing the CSI (Channel State Information) signal, extracting features, and building a machine-learning correlation model, the number and location of occupants can be estimated and further inferred in combination with sensing data such as vehicle temperature. Even better, with the computing power of the edge-side devices to process data in collaboration with the cloud, the computing process is partially localized to reduce the computing pressure and latency in the cloud. The approach has been experimentally verified to have more than 85% accuracy.

Keywords: CSI; ESP32; smart terminal; in-vehicle wireless sensing; personnel identification

1. Introduction

As car ownership has continued to grow in recent years, the intelligent development of the automotive industry has not completely solved some of the safety problems we are facing. According to published reports, safety accidents due to stranded children in cars are staged every year worldwide [1], and some parents continue to pay no attention to it. In fact, in the confined space of the car, high carbon dioxide concentration or high temperature, and in this environment for ten minutes child’s brain and kidneys will be damaged when the body temperature exceeds 40 °C, a threat to life [2]. Most existing solutions rely on a single IoT detection means: for example, the use of GPS, three-axis sensors, and other multi-sensors to determine whether the vehicle is completely parked, the use of cameras to detect whether there are children stranded in the car, and at the same time set the danger threshold through the temperature sensor to avoid danger. However, due to the face recognition approach of the camera, the perception performance is poor; it does not work in a lightless environment, and privacy cannot be guaranteed. Alternatively, infrared sensors, temperature sensors, carbon dioxide sensors, and other devices are used to determine whether a child has been left in the car; using more devices, deployment costs are higher, and infrared sensors have a large perceptive dead Angle.

Compared to traditional IoT sensors, sensing using CSI signals from Wi-Fi [3] overcomes the drawbacks of the above methods. CSI (Channel State Information), a channel property of a physical layer communication link, describes the signals on each signal transmission path. It has been used in many fields, such as headcount [4], action recognition [5], and gesture recognition [6]. Firstly, with the assembly of smart devices inside vehicles, Wi-Fi devices are widely deployed inside vehicles. In addition, the use of Wi-Fi for sensing is still essentially a non-contact identification method and does not directly capture the vehicle interior picture for identification: it offers good privacy and can operate normally.
in a lightless environment. Previous Wi-Fi-aware devices have used routers or devices equipped with fixed-model wireless cards to collect data that cannot be put inside real vehicles. This drawback is overcome by the ESP32 device, which allows direct access to a wealth of other RF (Radio Frequency) signals, such as CSI, from the microcontroller [7]. There is a high degree of versatility in IoT-related wireless sensing applications, which can be achieved using only battery power.

Based on the approach of using CSI sensing, we propose a contactless in-vehicle personnel sensing method, Car-Sense, based on ESP32. As shown in Figure 1, two ESP32 microcontrollers are installed in the front and rear of the interior space of a real vehicle. When both are active, the device can automatically initiate communication, and we can then collect CSI, RSSI, temperature, and other data for the detection and fully sense the various states of persons inside the vehicle in order to meet the low latency and high-agility real-time requirements needed for detection and also to reduce the computational pressure on the server, to allow the edge-end device at the source of data collection to integrate a series of core capabilities such as network computing and storage, to allow the device directly at the edge-end to undertake the task of data processing, and allow the cloud model to identify feedback more quickly. The main contributions of this paper are as follows.

Figure 1. Car-Sense, ESP32-based contactless in-vehicle occupant sensing.

- For dangerous events such as heat stroke, suffocation, or even death of persons left in the car, Car-Sense implements Device-free Wi-Fi Sensing with ESP32, which uses the characteristics of determining the number of people in the car as well as the location and temperature of the persons to derive the result of leaving/remaining. The module is very small compared to other commercial Wi-Fi devices, and the cost of the module is extremely low, requiring no additional hardware for the system to operate.
- The special model of Car-Sense Smart Side + Cloud localizes part of the computation process. The side end is not only used as a data collection device but also performs data cleaning and feature extraction locally; the cloud end performs machine-learning-related model calculations.
- The system utilizes self-built multi-person and multi-scene in-vehicle person perception data sets, sets up various evaluation metrics, and verifies the system's robust person location sensitivity and resolution through extensive experiments in different scenarios.

2. Related Work

The use of wireless devices to achieve vehicle interior sensing technology has a wide range of applications and can achieve a variety of states of identification of personnel. There are three specific types: infrared and other sensor-based methods, machine vision-based methods, and radio frequency sensing-based methods.

- Sensor-based approach: IoT sensors for detecting persons in vehicles generally include human infrared sensors, gas sensors, and temperature sensors. These methods generate sensory signals by sensing specific wavelengths of infrared light, specific
compositions of gases, and specific ambient temperatures to indirectly infer the status of people. Hairulnizam Mahdin [8] used Android apps with Bluetooth devices to warn parents or guardians that their children are being left in the car. Alternatively, using a microcontroller, Hall sensor, infrared, temperature, carbon dioxide sensor, and other modules, we have realized an alarm for children stranded in the car by setting the detection threshold. Alternatively, with the use of photoelectric sensors and other situation monitors inside the car, when stranded children are found in the car, the original car air conditioning system and windows can be alarmed and activated to protect the safety of the occupants. The above systems, because they rely on simple sensors with weak detection capabilities and are influenced by the environment, often achieve insufficient accuracy and, in practice, also depend on changes in the different states of the target.

- **Computer vision-based approach**: Based on image recognition methods arising from the development and application of neural networks in recent years, the technology of image recognition has crossed into a new stage. The camera takes photos of the interior of the car, extracts the characteristics of the person for analysis of the target state, and grasps the words and behaviors of the person inside the car. Using the age-recognition-based image recognition method, combined with machine learning, the Internet of things, cloud communication, and other technologies, around the face age recognition analysis of a variety of scenes and appropriate processing. The use of cameras and multiple IoT sensors combined to determine the true parking of vehicles and the presence of people. However, it is still not the best solution due to the lack of good privacy of image recognition and the fact that the recognition effect is greatly reduced when the photo is too dark in the absence of light or dark light.

- **RF signal-based methods**: wireless RF sensing technology has also been used as one of the sensing means due to its advantages, such as good privacy. The popularity of millimeter-wave radar and the maturity of the technology allows the millimeter-wave radar to achieve more fine-grained detection than Wi-Fi [9], whereas mm-Pose [10] uses millimeter-wave radar to achieve higher resolution accuracy; however, due to the high price of the equipment and the difficulty of deployment, it will not be available as a mainstream solution for some time in the future. The gradual maturity of Wi-Fi-CSI recognition technology has made it possible to use Wi-Fi signals for in-vehicle personnel detection. WiCAR [11] used commercial Wi-Fi to build a model to sense in-vehicle personnel movements with a recognition accuracy of about 95%. Although the resolution of Wi-Fi will never be as detailed as that of cameras and radar, there is greater scope for the devices to be deployed more easily.

Each of the above approaches has its own advantages and disadvantages. While cameras still seem to be the best solution, they are still not much used in the broad field of in-vehicle security because of privacy concerns. The means for wireless sensing becomes the optimal alternative to image sensing. In the Car-Sense mentioned in this paper, we use extremely inexpensive wireless Wi-Fi devices to achieve real-time data processing and presentation of results using locally trained models in advance while doing data collection. It also works well in the actual deployment stage.

### 3. Preliminaries

#### 3.1. CSI Signals from ESP32

CSI (Channel State Information) is a metric that describes signal propagation in a multipath environment and contains information about the changing channel environment in wireless signal propagation. The ESP32 Wi-Fi driver supports multiple 802.11 b/g/n protocols and can work in workstation mode or ap mode. When packets are transmitted from TX to RX, the channel frequency response contained in the CSI in ESP32 can be calculated. Each CFR register has two bytes of signed characters, the first part is the imaginary part, and the second part is the real part. Moreover, three types of CFR can be obtained from the CSI information, which are legacy long training field (LLTF), high-
throughput LTF (HT-LTF), and spacetime block code HT-LTF (STBC-HT-LTF) [12]. The ESP32 supports Wi-Fi bandwidth of either HT20 or HT40, but not both, in the default case of HT40. There are 128 subcarriers with complete signal gain, but the number of effective subcarrier bars is only 114, and the amplitude of each subcarrier is not the same. This is demonstrated by [13].

\[
H(f) = [H(f_1), H(f_2), \ldots , H(f_n)], n \in [1, N]
\]

(1)

where \(H(f_n)\) is the CSI on the \(n\)th subcarrier, and \(N\) is the number of valid subcarriers 114. The 114 CSI subcarrier variation curves caused by human activity within the ESP32 radio signal range are referred to in Figure 2.

![Figure 2. Human Activity CSI from ESP32.](image)

In the presence of influencing object activities within the Wi-Fi signal coverage, the Channel Impulse Response (CIR) [14] on different paths in the wireless channel will be forced to change accordingly. CIR is an extension of CSI in the time domain, and assuming linear time invariance, the channel impact response can be expressed as:

\[
h(\omega) = \sum_{i=1}^{N} a_i e^{-\theta_i} \sigma(\omega - \omega_i)
\]

(2)

where \(a_i\) is the amplitude decay of the \(i\)th path, \(\theta_i\) is the phase shift of the \(i\)th path, \(\omega_i\) is the time delay of the \(i\)th path, \(N\) is the total number of propagation paths, and \(\sigma(\omega)\) is the Dirichlet pulse function.

In the IEEE 802.11 n protocol, CSI can be extracted from all carrier bands of each channel in the Wi-Fi signal using OFDM techniques. Therefore, under the IEEE 802.11 n wireless protocol, CSI can be modeled as [15].

\[
\vec{Y} = H\vec{X} + Noise
\]

(3)

where \(\vec{Y}\) denotes the received signal, \(\vec{X}\) denotes the transmitter signal, \(Noise\) is the ambient noise. \(H\) is the channel matrix, which can be estimated as CSI. When a human body moves within the Wi-Fi signal, it affects the CSI, causing the CSI amplitude to change accordingly, and different numbers of humans in different locations will cause different CSI amplitude changes. In addition, CSI can capture the wireless characteristics of the nearby environment and is highly resistant to environmental noise.
3.2. AIOT of ESP32

AIoT [16,17] is a fusion application of AI and IoT, and the two technologies benefit from convergence, and their ubiquitous sensors and end devices provide a large number of analyzable data objects for AI, enabling AI research to be landed. IDC predicts that by 2022, more than 50 billion endpoints and devices will be connected, and more than 75% of future data will need to be analyzed, processed, and stored at the edge of the network [18]. Moreover, in a large number of vehicles relying on the distribution of edge nodes, the increase in computing power can take on more computing tasks. As shown in Figure 3, by arranging ESP32 nodes inside the vehicle, the ability to acquire data from common sensing devices can be accomplished, and some data can be processed directly and locally at the data edge side. Such a collaborative cloud-side approach will greatly reduce the time overhead.

![Figure 3. AIOT of ESP32.](image)

In traditional wireless sensing, data collectors are used only as a means of data collection and need to acquire physical world information from a single point, which needs to be collected manually and analyzed for data meaning. In the future, the core function of the edge-end node equipment is to realize the integration of “sense” and “know” through fusion-oriented sensing data collection. The ESP32 has a dual-core chip design, and its powerful dual-core performance is suitable for memory-hungry application scenarios, such as diverse AIoT applications. It is proven to provide intelligent services to users as a smart terminal node. In Car-Sense, the ESP32 wireless sensing node, through the Wi-Fi module integrated with the chip, acquires CSI data while using the data analysis module written in advance in the chip and sends it to the cloud after data cleaning and feature extraction processing at the side end. The cloud uses a supervised learning model to determine the final class of unknown samples using “labeled” samples. Specifically, machine learning algorithms are used in the cloud to determine data results and provide long-term, stable state awareness.

4. Method Design

In order to prevent suffocation events, ESP32 is deployed in the vehicle to obtain CSI to sense the behavior of the occupants, to determine whether there is such an event based on the location of the occupants and the temperature inside the vehicle, and to give timely feedback. The structure of this system is shown in Figure 4, which mainly includes the in-vehicle crowd counting method and in-vehicle person location method. Both use the CSI data and RSSI data collected by the ESP32 device, respectively, and use the same platform and different algorithms to synthesize the integrated in-vehicle people sensing.
The role of each of the modules included in Figure 4 is as follows.

Data acquisition stage: A transmit-and-receive Wi-Fi module is established in the vehicle to obtain data such as CSI generated by a different number of people staying in the vehicle several times.

Data preprocessing stage: The original ESP32-acquired CSI data are preprocessed with a total of 114 subcarriers, and the data are Hampel filtered as well as the minimum mean square error filtered.

Feature extraction stage: The preprocessed data extracts the features of signal amplitude fluctuation as well as the corresponding features through the change of signal distribution, including mean, variance, and standard deviation.

Identification phase: Establish different data models for the number of people identification and location identification, use the data collected in the early stage to train the model to adjust the appropriate parameters, and, combined with data from the ESP32 external temperature sensor, real-time clustering analysis is performed in the cloud.

4.1. Data Cleaning

In order to ensure that the data given to the model is fully reflective of the vehicle interior, but due to the multipath effect of the complex environment inside the vehicle, there is a large amount of noise in the original CSI data, as shown in Figure 5a, such that the information fed back from the time domain signal and the real data content will have a large deviation, so it is very necessary to choose certain noise reduction means to clean the data. In this paper, the Hampel filter and the minimum mean square error filter are selected for filtering, and the filtered data are shown in Figure 5b. It can be clearly found that the outliers of the data are reduced, and the characteristics of the data are more obvious.
For the non-real-time data used in the pre-training model offline, we select all valid 114 subcarriers in the data processing stage and perform noise reduction and smoothing accordingly for the integrity of the model. For online real-time data, after the corresponding noise reduction and smoothing process for the original CSI data, in order to balance the arithmetic power of the nodes, PCA is used to select the optimal subcarriers in the time domain signal, extract the eigenvalues of the optimal subcarriers and invoke the model for analysis, as shown in Figure as Figure 5c.

4.2. Feature Extraction

Considering that the amplitude of the CSI time domain signal fluctuates with the increase in the number of people in the room, along with the increase in the number of people, the fluctuation range of CSI amplitude becomes larger, and the change effect becomes more and more significant. As shown in Figure 6a, the figure corroborates well with the working principle of using CSI amplitude fluctuation for number identification. The mean, variance, and standard deviation values of subcarrier CSI amplitude values are used as features to characterize the fluctuation of CSI amplitude values according to the variation of CSI amplitude fluctuations. In this paper, the data features are extracted according to each fixed-size time window and constitute the feature vector as follows:

\[
    \text{feature} = [m, v, s]
\]

where \( m \) is the mean of all 114 subcarrier data, \( v \) is the variance of all subcarrier data, and \( s \) is the standard deviation of all subcarrier data.

![Figure 6. Comparison of subcarrier signals and signal eigenvalues: (a) CSI signals for different numbers of people; (b) Magnitude of change of characteristics.](image)

The feature extraction is to extract effective information from the CSI data after noise reduction. Firstly, the optimal subcarrier is selected by the PCA algorithm, and then the multidimensional data features of this subcarrier are extracted. The feature information of the time-domain signal under different numbers of people in the vehicle, including the variation magnitude of the values of mean, variance, and standard deviation, is shown in Figure 6b.

From the figure, it can be seen that as the number of people increases, the mean, variance, and standard deviation show an increase, with the variance showing the most significant performance, which indicates that there is a monotonic relationship between the data of multiple features in dimensions such as the number of people and the variance of CSI magnitude and the number of people in the vehicle. This further proves that the above feature values can be used as training features for the model.
4.3. Personnel Status Identification

For the detection of the problem of stranded people in the car mentioned in this paper, and considering the limited computing power of the side device ESP32 itself, we choose to determine whether there are children or other people stranded in the car from the number of people, their location, and temperature sensing. Specifically, in the first part, the data collected in the offline stage is used to train the model with the labeled data after data cleaning and feature extraction and sent to the cloud platform. In the second part, when the device detects the CSI data signal, it first performs the subcarrier processing locally and then sends the data to the cloud platform, and uses the powerful computing power of the cloud platform to derive the recognition results and return them in time, as shown in Figure 7.

**Figure 7.** Edge-side computing + cloud co-computing.

With the above steps, Car-Sense can use the collected wireless data to determine the status of the vehicle occupants. After inputting the necessary parameters into the program, the results of the two tests are obtained separately, and the final determination of the personnel status is made in a comprehensive manner. The pseudo-code is shown in Algorithm 1.

**Algorithm 1: Person status recognition**

**Input:** CSI Feature Vectors feature = [m, v, s], RSSI s, Temperature t

**Output:** Personnel Status δ

1. NOP = NumberOfPeople (feature) // Determine the number of personnel
2. FO = FingerprintOrientation (s) // Determine the location of personnel
3. The use of three indicators to determine the combination of people left in car δ
4. function NumberOfPeople (feature)
5. group, labels, k // Test sample, test sample label, K-value
6. result = knn_classify (feature, group, labels, k)
7. return result
8. end function
9. function FingerprintOrientation (s)
10. S // Off-line Fingerprint Library
11. for j = 1, 2, . . . , m do
12. Iterate over all data objects S = {RSSI1, RSSI2, . . . , RSSIm};
13. if |RSSIj – s| Smallest of all results
14. end if
15. end for
16. return j
17. end function
4.3.1. Personnel Quantity Detection

The number of people identified using CSI is based on the following principle: the activities of people in the car will have a certain impact on the propagation of Wi-Fi signals, and when the number of people in the car increases, the impact on Wi-Fi signals becomes more obvious, which in turn generates a specific pattern in the channel state information. Therefore, the key to using CSI for headcount identification is to find the relationship between CSI amplitude fluctuations and the number of active people in the vehicle.

In the offline model training phase, 500 sets of data are collected for each type of vehicle occupant, and different CSI magnitudes for different numbers of people are used to train the number of people recognition model and the location fingerprint detection model using machine learning methods, respectively, SVM, KNN, and multilayer perceptron MLP. The parameters of the three models are shown in Table 1:

Table 1. Parameters of the three classification algorithms.

| Classification Algorithm | Parameters |
|--------------------------|------------|
| KNN                      | K = 3      |
| SVM                      | C = 1, kernel = 'RBF', degree = 3, gamma = 'auto' |
| MLP                      | Number of hidden layer nodes [32, 16, 4] |

In the online phase, after the data is obtained locally, the optimal subcarrier is first selected according to the PCA principal component analysis algorithm mentioned above in the feature extraction phase, and simple operations such as filtering of the data are performed, and then the processed data feature values are sent to the cloud for the relatively more computationally intensive model identification. This allows for a timelier response.

4.3.2. Wi-Fi Fingerprint Positioning

The core idea of Wi-Fi fingerprinting is to associate a location in the physical environment with a certain “fingerprint”, a location that corresponds to a unique fingerprint. This fingerprint can be a feature or multiple features of this signal (most often the signal intensity). The establishment of the correspondence between location and fingerprint is usually performed in the offline phase. The target geographic area is usually divided into several rectangular grids, and at each grid point, the data fingerprint is obtained by sampling the data over a period of time.

Based on the distribution characteristics of Wi-Fi signals in a certain area, the location in the actual environment is associated with the signal “fingerprint” of the current location, and a location corresponds to a signal characteristic. The next step is to determine the current position based on the signal characteristics. In this paper, we divide the interior space of the vehicle into a 2×2 grid, as shown on the left side of Figure 8. Using the Received Signal Strength Indicator (RSSI) [5] as a feature of the signal and constructing a location fingerprint library $S$ in the offline phase, $S$ can be expressed as:

$$S = [RSSI_{11}, RSSI_{12}, RSSI_{21}, RSSI_{22}]$$  \hspace{1cm} (5)

where $RSSI_{ij}$ denotes the RSSI signal data sampled and processed by the person at row $i$, $i \in \{1, 2\}$, on the grid at column $j$, $j \in \{1, 2\}$.

The difference in the waveforms of the RSSI data can be clearly seen on the right side of Figure 8. We use this difference in features of different time-domain signals to build a fingerprint library and store it on the cloud platform. The system also uses a clustering algorithm to implement a classifier and arrange it on the cloud and uses the computed probability that the signal features belong to a certain distribution (stored in the fingerprint library) to estimate the location of the person. Next, after the wireless signal of the location information of the personnel inside the vehicle is collected at the side end of the online
stage, the target location is located in real-time with the help of the model of the cloud platform, i.e., the classifier is used to classify the target location and obtain the result.

![Figure 8. RSSI waveforms for different personnel positions in the vehicle.](image)

5. Experimentation and Evaluation

5.1. Experimental Setup

**Prototype.** Two ESP32 development boards are used as the main devices in this experiment. The integrated Wi-Fi modules support 802.11 b/g/n protocols, operate at 2.4 GHz with a bandwidth of up to 40 MHz, and can support external antennas to provide higher signal gain. The devices are equipped with ESP32-C3I-Tool [19] tools as RX and TX in the sensing environment, respectively. Moreover, they are equipped with independent power supply modules, so they can normally work for a long time even when the power supply is not available, and they are well prepared for possible power outages when the system is in operation. When the system is running, firstly, RX establishes WLAN LAN, TX connects to the LAN and sends packets to the gateway (RX) at a certain rate, RX parses the required wireless signal data such as CSI and RSSI in the received packets, performs noise reduction in real-time, extracts the amplitude characteristics of the optimal subcarriers to the cloud platform equipped with machine learning models for content recognition, and finally, feeds the recognition results back to TX for real-time results.

**Data collection.** The size layout of the interior space varies depending on the vehicle model and is affected by the multipath effect [20]. So, for this experiment, we conducted it in a general sedan (five seats), as shown in Figure 9. In the above environment, five different personnel situations are set up: no one, one person, two persons, three persons, and four persons. Moreover, the population of people included both large adults and small children. Data were collected continuously for 20 min in each case as offline training data. The mean, variance, and standard deviation are calculated separately to form the feature vector, and the feature vector is normalized to form the dataset, of which 80% is used for training and 20% for offline testing.

**Tooling Definition.** In order to determine the best results of the system under which time slice, we set different time slice lengths and experimented with three different models, KNN, SVM, and MLP, respectively, and concluded after extensive experimental analysis that the best results were achieved when using the KNN algorithm. Moreover, the classifier has a relatively good performance when $t = 20$ s. This is due to the fact that it measures the absolute distance between points in the multidimensional space. Reflected in this experiment is the measurement of the mean and variance of the CSI time domain signal caused by the difference between multiple dimensional data.

We define and train the model using the Keras framework. The collected data are also trained, and the weights and biases of the parameters are adjusted according to the degree of deviation of the output values from the expected values while a portion of the data is selected for validation. After having a more accurate model, we try to deploy the model to the cloud platform, which will save the model file in a special format to ensure the speed of identification, so that it can perform timely and accurate data analysis for identification.
5.2. Impact of User Diversity

Due to the existence of different reminders for different persons, adults and children have especially huge differences in body size and have different effects on wireless signals. In order to verify the recognition effect of the model on people with different body sizes, we set up a child group and an adult group for data comparison and compared the recognition accuracy of different people under the premise of ensuring the consistency of other indicators. In addition, we also explored the discrimination accuracy of general objects (using larger plush toys) instead of real people in a vehicle environment. The correct test rates are shown in Figure 10, respectively.

![Figures 9 and 10](image.png)

**Figure 9.** Vehicle interior equipment deployment.

**Figure 10.** Accuracy of recognition of different categories: (a) different number of adults; (b) different number of children; (c) position when a single person; (d) comparison of persons and dolls.

Several results can be drawn from the figure: (1) There is a 10% improvement in recognition of adults compared to children, probably due to the higher amplitude of movement, which has a significant effect on the signal resulting in higher accuracy. (2) The 100% recognition result at 0 can be attributed to the absence of human activity in the LOS area, where the signal fluctuation is essentially zero. (3) Children are smaller than adults, so there is a greater discount in recognition, but despite this, the system is still able to maintain an accuracy rate of over 65%. (4) The recognition of individual person locations is consistently above 80%, an advantage that is very relevant in the case of dangerous situations where individual children are stranded in the vehicle. (5) When a larger plush toy was introduced instead of a person for comparison, it was found that although the doll was larger, it had almost no effect on the recognition effect, and the difference in recognition effect between the presence or absence of the toy was not significant, this is because the toy does not move in the LOS area, which is then equivalent to “invisibility” for the CSI signal”.

Although the performance of Car-Sense is not very bright in terms of the multi-person recognition effect, but oriented to perceive the general situation where a single
person is enclosed in a car in a dangerous situation, Car-Sense can achieve satisfactory experimental results.

5.3. Impact of Location Diversity

In order to investigate the effect of different installation positions of Car-Sense’s wireless devices on the accuracy of detection, three different positions were set up in this paper, namely, the rear side and front side of the car, right B-pillar and the front side of the car, right B-pillar and left B-pillar, and packets were sent to RX via TX, and Figure 11a–c shows these three different device positions. Next, the relationship between accuracy and device deployment location is explored by determining the location information of personnel under different locations and the number of personnel information under different locations. The comparison results are shown in Figure 11d.

As can be seen from Figure 11d above, when the transmitting and receiving position is Location 1, when the device is in the installation position shown in Figure 11a above, the average accuracy of the number of people and the average accuracy of the location of a single person both reach more than 80%, which has high robustness. When the device is at Locations 2 and 3, analysis of Figure 11b, c above shows that due to the complex environment with more interfering objects in the signal propagation path, the recognition effect is around 70%, which is lower than that of Location 1. The reason for this is that when the device is placed in front and behind, the person is more fully exposed to the coverage of the wireless signal, and the time domain of the CSI signal is more robust. The time domain of the CSI signal carries more information, and the classification accuracy is higher. In general, the device has the best recognition effect under Location 1, which on the other hand, also shows that Car-Sense still requires a certain level of location for the device.

5.4. Impact of Different Classifiers

The performance of the classification algorithm is also an important indicator of the capability of the system. In order to verify the performance of the classification algorithm,
As can be seen from Figure 11d above, when the transmitting and receiving positions are fixed, the position accuracy is above 70%, and the accuracy rate is more than 85% for people with 0 or 1. Moreover, in the effect of personnel location recognition, the accuracy of MLP is only 58.4% which may be due to the problem of parameter setting at the beginning of training MLP, which does not achieve the optimal classification effect. Overall, due to the multipath effect, several experiments show higher recognition when the number of people is around zero, one, and the recognition decreases when the number of people is three and four. The comparison results prove that Car-Sense obtains the best results with the addition of the KNN algorithm.

5.5. Impact of Different Environments

In the real scenario, different vehicles will also have a greater impact on the recognition results because they have different cabin environments. Considering that most of the vehicles people use in daily life are cars and SUVs, we also set up two kinds of environments for testing in the environment settings, such as ordinary cars and SUVs, which keep other experimental parameters such as the location of the equipment unchanged and carry out comparison experiments of different vehicle internal environments under the same parameters, and the comparison results are shown in Figure 13.

As can be seen from the figure, in both scenarios, the basic recognition results are in line with those described in 6.1. The overall average recognition effect of the number of people is above 70%, and the accuracy rate is more than 85% for people with 0 or 1. The average result for a single-person location can be 88%. Due to the high similarity of the interior environments of the two scenarios selected, the experimental results can prove that Car-Sense can be more effective in producing correct results in general vehicles, especially for dangerous environments with a single person inside the vehicle, with better recognition accuracy and robustness in the different tested environments.
6. Conclusions

This paper presents Car-Sense, a lightweight, integrated ESP32-WiFi-based in-vehicle occupant sensing method that uses channel state information as the primary data source and is easy to deploy with good privacy compared to other sensing approaches, and addresses safety hazards caused by occupant retention through collaborative processing between local and cloud platforms. It has been shown to be unaffected by the presence of same-sized objects in the vehicle, to be more than 85% accurate for both adults and children, and to be robust across different vehicle types. However, the shortcomings of this system are that the recognition model is relatively simple, and the data set collected does not cover all vehicles. More efficient and simple solutions will be explored in the future.

Author Contributions: Conceptualization, G.W. and Z.H.; methodology, G.W. and Z.H.; software, data acquisition, and interpretation, G.W.; writing—original draft preparation, G.W.; critically revised the manuscript, Z.H., G.W. and X.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant 62262061, Grant 62162056), the Key Science and Technology Support Program of Gansu Province (Grant 20YF8GA048), 2019 Chinese Academy of Sciences “Light of the West” Talent Program, Science and Technology Innovation Project of Gansu Province (Grant CX2JA037, 17CX2JA039), 2019 Lanzhou City Science and Technology Plan Project (2019-4-44), 2020 Lanzhou City Talent Innovation and Entrepreneurship Project (2020-RC-116), and Gansu Provincial Department of Education: Industry Support Program Project (2022CYZC-12).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: All participants in the experiment were informed and consented to the publication of the paper.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors would like to thank all the participants and the school for their financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Barrera, J.P.S.; Sandoval, G.M.; Ortiz, G.C.; González, R.N.; Aguilar, E.R. A multi-agent system to avoid heatstroke in young children left in baby car seats inside vehicles. In Proceedings of the 2014 International Conference on Computational Science and Computational Intelligence, Las Vegas, NV, USA, 10–13 March 2014; IEEE: Piscataway, NJ, USA, 2014; Volume 2, pp. 245–248.
2. Adato, B.; Dubnov-Raz, G.; Gips, H.; Heled, Y.; Epstein, Y. Fatal heat stroke in children found in parked cars: Autopsy findings. *Eur. J. Pediatr.* 2016, 175, 1249–1252. [CrossRef] [PubMed]
3. Chetty, K.; Smith, G.E.; Woodbridge, K. Through-the-wall sensing of personnel using passive bistatic WiFi radar at standoff distances. *IEEE Trans. Geosci. Remote Sens.* 2011, 50, 1218–1226. [CrossRef]

4. Xi, W.; Zhao, J.; Li, X.Y.; Zhao, K.; Tang, S.; Liu, X.; Jiang, Z. Electronic frog eye: Counting crowd using WiFi. In Proceedings of the IEEE INFOCOM 2014 Conference on Computer Communications, Toronto, ON, Canada, 27 April–2 May 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 361–369.

5. Jiang, W.; Xue, H.; Miao, C.; Wang, S.; Lin, S.; Tian, C.; Murali, S.; Hu, H.; Sun, Z.; Su, L. Towards 3D human pose construction using WiFi. In Proceedings of the 26th Annual International Conference on Mobile Computing and Networking, London, UK, 21–25 September 2020; pp. 1–14.

6. Gao, R.; Zhang, M.; Zhang, J.; Li, Y.; Yi, E.; Wu, D.; Wang, L.; Zhang, D. Towards position-independent sensing for gesture recognition with Wi-Fi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2021, 5, 1–28. [CrossRef]

7. Atif, M.; Muralidharan, S.; Ko, H.; Yoo, B. Wi-ESP—A tool for CSI-based device-free Wi-Fi sensing (DFWS). *J. Comput. Des. Eng.* 2020, 7, 644–656. [CrossRef]

8. Mahdin, H.; Omar, A.H.; Yaacob, S.S.; Kasim, S.; Fudzee, M.F.M. Minimizing heatstroke incidents for young children left inside vehicle. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2016; Volume 160, p. 012094.

9. Li, G.; Zhang, Z.; Yang, H.; Pan, J.; Chen, D.; Zhang, J. Capturing human pose using mmWave radar. In Proceedings of the 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Austin, TX, USA, 23–27 March 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.

10. Siengupta, A.; Jin, F.; Zhang, R.; Cao, S. mm-Pose: Real-time human skeletal posture estimation using mmWave radars and CNNs. *IEEE Sens. J.* 2020, 20, 10032–10044. [CrossRef]

11. Wang, F.; Liu, J.; Gong, W. WiCAR: WiFi-based in-car activity recognition with multi-adversarial domain adaptation. In Proceedings of the 2019 IEEE/ACM 27th International Symposium on Quality of Service (IWQoS), Phoenix, AZ, USA, 24–25 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–10.

12. ESP-IDF Programming Guide. 2022. Available online: https://docs.espressif.com/projects/esp-idf/zh_CN/latest/esp32/api-guides/wifi.html (accessed on 3 December 2021).

13. Hernandez, S.M.; Bulut, E. Lightweight and standalone IoT based WiFi sensing for active repositioning and mobility. In Proceedings of the 2020 IEEE 21st International Symposium on “A World of Wireless, Mobile and Multimedia Networks” (WoWMoM), Cork, Ireland, 31 August–3 September 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 277–286.

14. Hao, Z.; Zhang, D.; Duan, Y.; Dang, X. Indoor sneezing and coughing detection based on COTS wireless device. In Proceedings of the 2021 IEEE 27th International Conference on Parallel and Distributed Systems (ICPADS), Beijing, China, 14–16 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 763–772.

15. Niu, K.; Zhang, F.; Jiang, Y.; Xiong, J.; Lv, Q.; Zeng, Y.; Zhang, D. WiMorse: A Contactless Morse Code Text Input System Using Ambient Wi-Fi Signals. *IEEE Internet Things J.* 2019, 6, 9993–10008. [CrossRef]

16. Wu, J.; Li, W.; Cao, J.; Qian, S.; Zhang, Q. BUYYA Rajkumar. A Review of Intelligent Internet of Things (IoT) Research. *Telecommun. Sci.* 2021, 37, 1–17.

17. Li, T.C.; Lai, Z.; Chen, L.Q.; Li, M.L. White paper on Intelligent Internet of Things (AIoT) in China in 2020. *Internet Econ.* 2020, 90–97. [CrossRef]

18. IDC. *IDC Industry Data Deep Research Analysis and Development Strategic Planning Report*; IDC: Needham, MA, USA, 2021.

19. ESP32-C3I-Tool. 2022. Available online: https://github.com/StevenMHernandez/ESP32-C3I-Tool (accessed on 12 June 2022).

20. Wu, D.; Zeng, Y.; Zhang, F.; Zhang, D. WiFi CSI-based device-free sensing: From Fresnel zone model to CSI-ratio model. *CCF Trans. Pervasive Comput. Interact.* 2021, 4, 88–102. [CrossRef]