

Article

A Comparison of Machine Learning Algorithms for Wi-Fi Sensing Using CSI Data

Muhammad Ali ^{1,*}, Paul Hendriks ¹, Nadine Popping ¹, Shaul Levi ¹ and Arjmand Naveed ²¹ Gamgee BV, Barbarossastraat 155, 6522 DK Nijmegen, The Netherlands² Faculty of Engineering and Informatics, City Campus, University of Bradford, Bradford BD7 1DP, UK

* Correspondence: muhammad.ali@gamgee.com

Abstract: In today's digital era, our lives are deeply intertwined with advancements in digital electronics and Radio Frequency (RF) communications. From cell phones to laptops, and from Wireless Fidelity (Wi-Fi) to Radio Frequency IDentification (RFID) technology, we rely on a range of electronic devices for everyday tasks. As technology continues to evolve, it presents innovative ways to harness existing resources more efficiently. One remarkable example of this adaptability is the utilization of Wi-Fi networks for Wi-Fi sensing. With Wi-Fi sensing, we can repurpose existing networking devices not only for connectivity but also for essential functions like motion detection for security systems, human motion tracking, fall detection, personal identification, and gesture recognition using Machine Learning (ML) techniques. Integrating Wi-Fi signals into sensing applications expands their potential across various domains. At the Gamgee, we are actively researching the utilization of Wi-Fi signals for Wi-Fi sensing, aiming to provide our clients with more valuable services alongside connectivity and control. This paper presents an orchestration of baseline experiments, analyzing a variety of machine learning algorithms to identify the most suitable one for Wi-Fi-based motion detection. We use a publicly available Wi-Fi dataset based on Channel State Information (CSI) for benchmarking and conduct a comprehensive comparison of different machine learning techniques in the classification domain. We evaluate nine distinct ML techniques, encompassing both shallow learning (SL) and deep learning (DL) methods, to determine the most effective approach for motion detection using Wi-Fi router CSI data. Our assessment involves six performance metrics to gauge the effectiveness of each machine learning technique.



Citation: Ali, M.; Hendriks, P.; Popping, N.; Levi, S.; Naveed, A. A Comparison of Machine Learning Algorithms for Wi-Fi Sensing Using CSI Data. *Electronics* **2023**, *12*, 3935. <https://doi.org/10.3390/electronics12183935>

Academic Editor: Miroslav Joler

Received: 28 July 2023

Revised: 12 September 2023

Accepted: 15 September 2023

Published: 18 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: motion detection; Wi-Fi sensing; machine learning; shallow learning; deep learning; channel state information

1. Introduction

Though Wi-Fi is not the only means of motion sensing, there are other means of detecting motion in any premises using Passive Infra-Red (PIR) sensors, vision sensors, ultrasound sensors, and other RF-based sensors such as RFID-based sensors. Alternatively, a lot of research has been carried out on improved and efficient solutions using low-energy communications devices like Bluetooth Low Energy (BLE) devices. The authors in [1] presented a scalable and non-intrusive method to detect occupancy in building zones, utilizing BLE technology in smartphones. Signal strength data collected by BLE beacons were processed through machine learning models to determine occupants' locations within zones. Both supervised ensemble and semi-supervised clustering models were assessed, with the latter showing efficient performance. The Singapore case study showcased up to 86% accuracy in locating occupants. Furthermore, this study identified distinct occupancy profiles based on movement patterns, offering insights for building management. The method's scalability suggested broader practicality. The downside of this approach is its dependence on the occupants carrying cell phones, plus the BLE technology of the cell phone needs to be active for the correct operation of this method. Furthermore, in

circumstances where a person is equipped with more than one BLE-enabled device, such as a cell phone, smart watch, or earphone, the estimated occupancy of the proposed method can end up with inaccuracies in the results.

As mentioned above, numerous ways of motion detection with high accuracies using various dedicated hardware technologies have been presented by researchers globally, but the goal of the research presented in this paper is to achieve motion detection using only the pre-existing Wi-Fi devices in the customers' premises with maximum precision, with minimum to no dependencies on other hardware like cell phones with BLE, and without introducing new hardware overhead to customers, keeping the privacy of the customers intact as there will be no images or video data being recorded or processed when applying Wi-Fi sensing. Specialized hardware, such as motion detection sensors with higher sensitivity, can detect motion more precisely, and Bluetooth low-energy devices can detect premises occupancy more efficiently with lower power consumption. Similarly, active and passive RFID tags can be used to identify a certain device and or person with which it has been associated. The actual percentage of precision achieved using different hardware and software strategies is highlighted in the literature review section. On a quick note, use of BLE can give a precision of up to 86% [1] in premises occupancy estimation; passive processing of Wi-Fi CSI data using AI has given 97% precision for motion detection; and a precision of 95% was achieved using Ultra Wide Band (UWB) technology for Human Activity Recognition (HAR). The use of these specialized hardware technologies for different goals will have a higher impact on the overall cost for the service provider as well as for the consumer. Therefore, the research work presented in this paper focuses on achieving Wi-Fi sensing using only the pre-existing Wi-Fi router in consumers' premises with improvements and add-ons on the router firmware and on customers' mobile apps to support features using Wi-Fi sensing.

The Wi-Fi routers in our daily lives are mainly used for internet and intranet connectivity, mostly for communications, entertainment, data exchange, etc. The channel state information (CSI) and Received Signal Strength Indicator (RSSI) statistics are utilized mainly to analyze Wi-Fi channel conditions and adjust router configuration as necessary. Because people inside a Wi-Fi router's range also cause radio waves used for Wi-Fi communication between devices to be distorted, examination of the distortion parameters allows one to infer information about nearby activity without actually using the visualization and PIR sensing devices. The Wi-Fi signal being transferred between Wi-Fi networking devices in the target premises is the main focus of Wi-Fi sensing. To detect distortions in the samples produced by movements in the target premises, we use the CSI of the Wi-Fi signal.

An indicator of signal intensity is the RSSI. Although it has been actively used for active localization based on the Wi-Fi fingerprinting technique or as a metric for passive tracking of mobile devices, it is in fact quite unstable and varies from vendor to vendor. Additionally, it cannot accurately capture changes in signal caused by human movements, especially if a person is not directly between an access point and a Wi-Fi router. The CSI approach offers more precise information on the state of the channel. At each sub-carrier frequency, it monitors the amplitude and phase distortions of the wireless signals that are in motion for each antenna pair of the transmitter and receiver. As a result, CSI variations in the time domain exhibit various patterns for various people, activities, etc., and this can be used for Wi-Fi sensing in intruder alarm systems, gesture recognition, and healthcare applications, particularly fall detection, etc.

Using orthogonal frequency division multiplexing (OFDM) and Multiple-Input Multiple-Output (MIMO) technology, the CSI saves the wireless signal amplitude and phase value information for each pair of transmit-receiver antennas in an OFDM subcarrier. You can imagine a 2.4-GHz band as a narrowband flat-fading channel, similar to wireless technology 802.11n, and it can be represented using the following straightforward equation:

$$Y = [H] \cdot [X] + [N] \quad (1)$$

Here, X and Y stand in for the transmitter and receiver's respective vectors, N for the Gaussian noise vector that is always present in the RF channel, and H for the channel matrix. Two TP-Link Archer C7 routers have been utilized in the experimental Wi-Fi sensing configuration, one as a transmitter and the other as a receiver. There are three antennas on both the CSI transmitter and receiver routers. As a result, our wireless communication system is 3×3 MIMO, and the CSI data are divided into nine streams in accordance with the nine pairs of transmitter-receiver links. For a device that distributes each stream's n subcarriers evenly among the channel's 56 subcarriers, a 20 MHz device. The size of the CSI data group matrix becomes nine rows and n columns, respectively, which leads to the $(9 \times n)$ data groups derived from each received CSI packet.

As CSI contains noises produced by indoor environments, the CSI data packets received at the receiver are quite noisy. Additionally, internal state transitions in the wireless signal transmitter and receiver devices are brought on by changes in transmission power, transmission rate adaptation, and internal reference level changes. Low-pass filters are used to first denoise the CSI data before the first data processing operations, such as data fusion employing correlation to reduce redundancy without losing important information, are implemented. CSI data can be utilized for motion detection by utilizing amplitude and or phase variance when the data processing is complete. It can also be used to train machine learning algorithms for motion detection, gesture recognition, and personal identification, among other things.

In the orchestration process, we have considered several machine learning (ML) algorithms and applied both shallow learning (SL) and deep learning algorithms to the CSI data collected. We employed classification-based ML algorithms for operations like motion detection and clustering-based ML algorithms for unique personal identification. The list of SL algorithms considered includes SVM, naïve Bayes, decision tree, K-nearest neighbors, and K-means algorithms, and the list of DL algorithms considered includes recurrent neural networks (RNN), convolutional neural networks (CN), and deep neural networks (DNN), respectively. These algorithms were trained and validated using a harmonized set of samples from publicly available Wi-Fi CSI datasets from CRAWDAD and CSI datasets from IEEE DataPort. The performance metrics were analyzed from a set of results obtained using each of the ML algorithms considered to compare the performance and efficiency of each ML algorithm and select the most suitable for Wi-Fi sensing using the CSI data captured from TP-Link Archer C7 routers in the target premises.

The novelty and main contributions of this paper, which distinguish this work from pre-existing research work, are as follows:

Selection of Machine Learning Algorithms: The selection of machine learning algorithms is carried out for the comparison of performance when Wi-Fi sensing data, i.e., CSI data, is presented to these machine learning algorithms. The selected set of machine learning algorithms from both shallow learning (SL) and deep learning (DL) are chosen based on the type of solutions they provide. The presented research work in this paper only focuses on those machine learning algorithms that address the classification problem, i.e., classification of data for deciding whether motion has been detected or not.

Training and Evaluation of ML Models: All nine ML models, i.e., six SL and three DL models, are trained using the Wi-Fi CSI dataset from the IEEE data port repository, which contains a labeled dataset for humanoid motion detection. Hence, all the training carried out on these models was supervised machine learning. As mentioned earlier, the IEEE dataset for Wi-Fi sensing, i.e., the Widar dataset, was initially used for benchmarking and comparing the performance of ML models, which was then replaced by a locally captured CSI-based Wi-Fi dataset. In the locally captured dataset, two classes have been used, i.e., clean with no motion at all and another with a person walking. The number of training samples in the Widar dataset used is 35 k, and the number of samples in the testing data are 9 k. The number of training data samples used in locally captured CSI data are 8 k, and the number of testing data samples used in locally captured CSI data are 2 k.

MADM for ranking ML Algorithms: Last but not least, in this research work, after carrying out performance analysis of each machine learning algorithm, the multi-attribute decision-making algorithm has been employed to systematically rank each machine learning algorithm. The MADM is introduced because there are multiple performance metrics against each ML algorithm, which makes it very hard to select the most appropriate ML algorithms in this case. Therefore, MADM is utilized here as it best suits the problem when we have nine different ML algorithms, with each algorithm having six different performance attributes.

The structure of this paper is drafted in such a way that, followed by the brief introduction presented in this section, Section 2 presents the literature review, Section 3 presents the proposed work and methods, Section 4 presents the results, which are followed by discussion in Section 5, and finally the conclusion in Section 6 and references in the last section of this paper.

2. Literature Review

A vast number of researchers globally have performed research on Wi-Fi sensing using CSI, RSSI, and other methods, mostly focusing on CSI and RSSI methods. The following set of paragraphs sheds light on different research projects carried out globally by researchers to perform Wi-Fi sensing.

Different sensing technologies are used to examine diverse human actions and gestures to perform human activity recognition efficiently. These technologies include sensors for motion detection [2], sensors for vision-based detection [3], sensors for sound-based sensing [4], and pyroelectric infrared light-based sensors [5]. To measure body motions using motion sensor technology, people typically need to wear specialized devices, which is not always practical. Approaches using cameras and other devices or sensors based on the visual data can function effectively in specific lighting conditions, which can be easily obstructed by smoke, opaque objects, or low illumination conditions. Additionally, because acoustic signals attenuate quickly, acoustic-based techniques are unstable in the presence of background noise and outside sound interference, and their sensing range is constrained. Overall, using traditional approaches requires more work due to complex hardware installation and a variety of maintenance requirements. A low-cost, non-intrusive approach to recording human body motions associated with daily activities is desired to overcome the restrictions discussed in this article. Recently, an increasing amount of research has focused on radio frequency (RF)-based approaches for human activity detection, such as Wi-Fi. Nearly every electronic device in homes and offices, including smart speakers (like the Amazon Echo and Apple HomePod), smart TVs, smart thermostats, and home security systems, may now be connected wirelessly thanks to the widespread use of Wi-Fi technology. Indoor spaces typically allow Wi-Fi signals to spread out over tens of meters, and the wireless connections between these smart gadgets create an exhaustive combination of the reflected light rays that reaches each corner and narrow place. People's presence and associated body motion will have a significant impact on wireless signals, leading to significant variations in the amplitude and phase of received signals. These changes can be used to record human body movements associated with daily activities.

The research work presented in [6] aimed to tackle indoor occupancy estimation challenges using a combination of Bluetooth low energy (BLE) technology and machine learning. They developed a prototype system that comprises BLE beacons, a mobile application, and a remote server. By employing three distinct machine learning methods, they classified occupancy based on the data collected from these beacons. Their experimentation demonstrated the effectiveness of this approach in accurately estimating occupancy. The server handles data processing and training, eliminating the need for complex operations on the mobile application.

The authors in [7] presented "Plug-Mate", an IoT-based plug load management system that optimizes energy use and user comfort via intelligent automation leveraged high-resolution occupancy data, advanced plug load recognition, and personalized controls. In

a 5-month university office study, six strategies were evaluated, with the most successful achieving 51.7% energy savings across plug load types, a 7.5% reduction in building energy use, and high user satisfaction.

The paper [8] addressed the energy consumption in commercial buildings, focusing on heating, ventilation, and air conditioning (HVAC) systems. It introduced “Sentinel”, a system that utilizes existing Wi-Fi infrastructure and occupants’ smartphones for precise HVAC control based on occupancy. Unlike traditional sensor-based solutions, Sentinel reduces deployment costs. It achieved 86% accurate occupancy detection within office spaces, with minimal errors attributed to smartphone power management. In the real-world test, Sentinel controlled 23% of HVAC zones, resulting in a prominent 17.8% energy savings compared to static scheduling.

Research in [9] employed diverse sensor data for predicting occupancy in various room types. A new feature selection algorithm was introduced, surpassing the common approach by enhancing model performance with fewer sensors. Outcomes revealed that indoor CO₂ levels and Wi-Fi-connected devices are pivotal in predicting occupancy across offices, libraries, and lecture rooms. Optimal model performance was attained using distinct deep learning architectures for each room type. The algorithm’s usability was extended to other datasets, providing insights to curtail sensor needs and deployment expenses in building management.

In [10], a robust Wi-Fi-based passive sensing technique named CNN-ABLSTM was introduced, combining CNN and attention-based bi-directional LSTM to address challenges like low sensing accuracy and high computational complexity. By utilizing CSI for Wi-Fi passive sensing, it achieves precise human activity recognition. CNN extracts features, reducing redundancy, while the attention mechanism improves model robustness. Simulation results show that CNN-ABLSTM improves recognition accuracy by up to 4%, reduces computation significantly, and maintains 97% accuracy across different scenarios and objects. Compared to traditional approaches, this DL-based method outperforms them, making it promising for advanced wireless communication systems.

Also, the increasing elderly population and the strain on healthcare services due to the COVID-19 pandemic have led to a demand for technological solutions in elderly homes. Research [11] introduced a real-time, noninvasive sensing system that utilized radio frequency (RF) sensing and channel state information (CSI) reports to monitor activities of daily living (ADLs). Machine learning, specifically the random forest algorithm, was employed to accurately classify ADL categories like “movement”, “empty room”, and “no activity”, which achieved 100% accuracy on new testing data. The system detected movement using Wi-Fi signals without the need for wearables, and disruptions in CSI data indicate the presence of a person. This proposed real-time monitoring system enhances elderly care.

Another study [12] focused on ambient computing and used Wi-Fi channel state information (CSI) as a non-contact method for recognizing human activities indoors. LSTM outperformed CNN, and hybrid models achieved 95.3% accuracy in multi-activity classification. The research shows that RF sensing for indoor human activity recognition is feasible and offers privacy-friendly alternatives to vision-based systems. The study also suggested further investigation into the system’s resilience in diverse environments and its ability to recognize activities for multiple users. Overall, LSTM-based RF sensing proves effective for indoor activity recognition and holds significant potential in various applications.

A research paper [13] presented a sign language recognition system based on deep learning and Wi-Fi CSI data. The proposed model utilized CNN, LSTM, and ABLSTM with different optimizers and preprocessing methods. It achieved impressive recognition accuracy of 99.855%, 99.674%, 99.735, and 93.84% in various environments and multi-user scenarios. The study demonstrated the effectiveness of using Wi-Fi signals for gesture recognition, surpassing other deep learning approaches. Additionally, the researchers suggested considering transfer learning like ResNet for future improvements.

Another study [14] explored device-free human activity recognition (HAR) using Wi-Fi CSI signals. Two algorithms, SVM and LSTM, are proposed for classification, with SVM employing wavelet analysis for preprocessing and feature extraction, while LSTM processes raw data directly. The research achieved high accuracy in detecting various human activities, including falls and counting individuals in a room.

A similar survey [15] investigated device-free human gesture recognition using Wi-Fi channel state information (CSI). It categorized recognition into device-based and device-free sensing methods and highlighted advancements in Wi-Fi CSI. The study examined model-based and learning-based approaches, discussing their recognition performance and signal processing techniques. Deep learning methods showed promise with large datasets, while model-based approaches performed well with a single participant. Challenges included handling non-Gaussian signal distributions and capturing fine-grained information.

Another article [16] presented EfficientFi, a new wireless sensing framework for large-scale Wi-Fi applications in smart homes. By overcoming existing limitations, EfficientFi used quantized representation learning with joint recognition, enabling efficient compression of Wi-Fi CSI data at the edge and accurate sensing tasks. It achieved remarkable data compression and high accuracy in human activity recognition and identification. Compared to classic methods, EfficientFi outperformed in compressive sensing and deep compression, demonstrating its potential for IoT-cloud-enabled Wi-Fi sensing applications.

The study in [17] also focused on human activity recognition (HAR) using ultra-wideband (UWB) technology and Wi-Fi CSI. Through experiments, the UWB CIR data achieved a remarkable F1-score of 95.53% in activity classification. In comparison, Wi-Fi CSI data achieved F1-scores of 92.24% and 80.89% with denoised amplitude values and spectrograms, respectively, for the same activities. The research highlighted UWB's superiority over Wi-Fi for HAR and offered advantages like a smaller data dimension and lower signal processing requirements. UWB technology proved valuable not only for localization/tracking but also for device-free HAR.

Researchers in [18] focused on a contactless respiration detection system using Wi-Fi CSI. The ResFi system achieved a remarkable 96.05% accuracy in detecting human respiration, outperforming traditional machine learning methods. The study emphasized the potential of learning-based approaches for non-contact vital signal detection.

A similar study [19] concentrated on detecting human presence in rooms without the need for devices using Wi-Fi CSI data. The proposed approach employed the dynamic time wrapping (DTW) algorithm to compare empty and filled rooms, achieving accuracy comparable to existing methods. Experimental results demonstrated a 99.21% accuracy comparable to a 99.98 accuracy with the RF algorithm.

The RSSI CSI, which is readily available on many commercial network interface cards with modified driver software, allowed the researchers to measure the physical layer parameters of the wireless channel and carry out motion detection using the Wi-Fi signals. Wi-Fi signals can be modified to transmit wireless signals on a radio platform defined by a universal software radio peripheral (USRP), such as frequency modulated carrier wave (FMCW), to determine the frequency shift of the signal brought on by human motion in the target premises [20]. The following Table 1 presents the overall comparison between different strategies in the literature, considering three different attributes: methodology considered, application of the methodology, and key findings of the corresponding strategies.

The research methods reviewed above target different domains, i.e., starting from premises occupancy estimation, smart energy management, HVAC, HAR, respiration detection, and motion detection, using a variety of approaches with different hardware and software assistance. The research work performed so far in the literature has mostly applied some analytical or artificial intelligence (AI) or machine learning (ML)-based methods, with some support from theoretical arguments in the literature. The lack of comparison between different ML methods, particularly a comparison between shallow learning (SL) and deep learning (DL) models for motion detection using a Wi-Fi-CSI-based dataset, has been identified and explored in the research work presented in this paper.

Furthermore, the selection of the most efficient ML algorithm has been carried out using the systematic approach of a multi-attribute decision-making algorithm, which was not seen in the literature. The work presented in this paper contributes to the validation of the process for selecting the best ML techniques in motion detection using Wi-Fi sensing. It also explores the behavior of various ML algorithms, i.e., SL and DL, when a CSI-based dataset is presented to these ML algorithms for training and testing.

Table 1. Comparison of approaches in literature.

Research Work	Methodology	Application	Key Findings
Filippoupolitis et al. [6]	Bluetooth Low Energy	Emergency Management	Implementation of BLE-based occupancy detection.
Tekler et al. [7]	IoT-based Plug Load	Plug load management	IoT-driven system for smart buildings.
Balaji et al. [8]	Wi-Fi Infrastructure	HVAC Actuation	Utilizing existing Wi-Fi infrastructure for HVAC control.
Tekler and Chong [9]	Deep Learning	Occupancy Prediction	Minimum sensing strategy using deep learning.
He et al. [10]	CSI-based Wi-Fi Sensing	Passive Sensing	Robust Wi-Fi passive sensing using deep learning.
Taylor et al. [11]	Real-Time Activity Sensing	Activity Sensing	Identification of optimal machine learning techniques.
Khan et al. [12]	Flexible SDR	Human Activity Detection	Contactless human activity detection using deep learning.
Bastwesly et al. [13]	Wi-Fi CSI	Sign Language Recognition	Deep learning for sign language recognition.
Damodaran et al. [14]	Wi-Fi CSI	Activity and Fall Recognition	Device-free human activity and fall detection.
Ahmed et al. [15]	Wi-Fi CSI	Gesture Recognition	Survey of device-free gesture recognition.
Yang et al. [16]	Efficient Wi-Fi Sensing	Wi-Fi Sensing	Large-scale lightweight Wi-Fi sensing via CSI compression.
Bocus et al. [17]	UWB CIR vs. Wi-Fi CSI	Activity Recognition	Comparison of UWB CIR and Wi-Fi CSI for activity recognition.
Hu et al. [18]	Wi-Fi-Enabled Respiration Detection	Respiration Detection	Wi-Fi-based device-free respiration detection using deep learning.
Soto et al. [19]	Wi-Fi CSI-based Presence Detection	Presence Detection	Human presence detection using Wi-Fi CSI and DTW features.
Adib et al. [20]	3D Tracking via Radio Reflections	3D Tracking	3D tracking using body radio reflections.
Proposed work	CSI based W-Fi sensing	Wi-Fi sending motion detection	Comparison of ML models for Wi-Fi sensing using CSI data

All the research work done so far has focused solely on methods, tuning, and utilization of machine learning (ML) techniques to achieve the goal of Wi-Fi sensing to detect humanoid motion in the coverage area of the Wi-Fi network. In this article, a broader aspect of Wi-Fi sensing has been addressed, which is to analyze a set of machine learning algorithms to find out which ML methods are more suitable for the problem of Wi-Fi sensing when using the CSI data for training and detecting motion in the Wi-Fi coverage area. For this purpose, a number of shallow learning (SL) and deep learning (DL) algorithms were selected based on their characteristics, such as suitability for tabular data and classification capabilities, to suit our requirements for motion detection using Wi-Fi CSI data.

2.1. Wi-Fi Sensing Techniques

Various types of techniques have been explored by researchers globally when implying Wi-Fi sensing for motion detection purposes. Here we have classified these techniques based on the hardware deployed for Wi-Fi sensing, i.e., using commercial off-the-shelf (COTS) hardware such as Wi-Fi routers used at home for Wi-Fi access and using customized hardware such as software-defined radio, e.g., URSP, FPGA boards, etc.

2.1.1. The COTS Hardware-Based Wi-Fi Sensing Techniques

Techniques using COTS routers involve the use of the received signal strength indicator (RSSI) and channel state information (CSI).

RSSI: data is available in most Wi-Fi devices, which indicates the path loss of wireless signals with respect to a certain distance and can be derived following the log-normal distance path loss (LDPL) model.

The CSI: To detect human activity with accuracy and dependability, Wi-Fi signal data are used. In order to accurately reflect the combined effect of, for instance, scattering, fading, and power decline with distance, more fine-grained CSI must be captured. Since wireless signals in an indoor setting could practically travel through any corner, the presence or movement of a human body would affect wireless signal propagation, leading to minute variations in numerous reflected rays. The measurable CSI values are created by all of these multi-path rays, which can also be utilized to identify and monitor human body movements. In contrast to RSSI, CSI is a set of complex values for several orthogonal frequency-division multiplexing (OFDM) subcarriers that include both amplitude and phase information. The effects of multi-path fading vary for every channel using a little variance in the center frequency, and all the subcarriers collectively represent the wireless channel in a fine-grained way. With customized drivers, any device with commercial Wi-Fi interfaces may measure CSI, just like RSSI. Researchers are now using it often to accomplish tasks including human intrusion detection, walking speed/direction estimation, and human activity recognition [21,22].

2.1.2. The Customized Hardware-Based Wi-Fi Sensing Techniques

Similar to the COTS device-based Wi-Fi sensing techniques, two main approaches to the customized hardware-based Wi-Fi sensing techniques are described in this article. These two techniques are frequency modulated carrier wave (FMCW) and Doppler shift methods.

FMCW technique: The measurement of human motion based on radio reflection from the human body, particularly by calculating the amount of time needed for the Wi-Fi signal to go from the transmitter to the reflecting body and back to the receiver. Given that wireless transmissions often move at the speed of light, determining the time of flight for the Wi-Fi signal is not a simple operation. To calculate the radio signal's time of flight, the FMCW can map the difference in time to a carrier frequency shift. It is crucial to remember that FMCW technology relies on specialized equipment (such as USRP) to generate the signal that sweeps the frequency across time, in contrast to conventional Wi-Fi that employs OFDM. The writers of the references [23–26] have shown how to estimate motion detection using FMCW for a variety of uses.

Doppler Shift technique: Another physical layer characteristic of wireless transmissions that can be utilized to detect human activity is Doppler shift effects. It specifically monitors the frequency shift in the received signal of Wi-Fi as the transmitting and receiving devices change positions in close proximity to one another. Any movement of the human body would cause a Doppler shift if the wireless signal received and reflected from it were regarded as the signal sent out by the wireless transmitter. In particular, moving towards the receiver causes a positive frequency change (also known as a Doppler shift), but moving away from the receiver causes a negative frequency change. The authors in the cited publications [27–31] have suggested their work utilizing the doppler shift effects

with software-defined radio (SDR) for recognition of human movements such as walking and running.

3. Proposed Work

We harnessed classification-based ML algorithms for motion detection. Our roster of SL algorithms encompassed SVM, naïve Bayes, decision tree, K-nearest neighbors, and K-means. In the realm of DL algorithms, we delved into recurrent neural networks (RNN), convolutional neural networks (CNN), and deep neural networks (DNN). These algorithms underwent rigorous training and validation using a harmonized dataset sourced from publicly available Wi-Fi CSI datasets. To evaluate their effectiveness, we scrutinized performance metrics derived from each ML algorithm's results. This comprehensive analysis allowed us to gauge the efficiency of each ML algorithm and identify the most suitable candidate for Wi-Fi sensing with CSI data sourced from TP-Link Archer C7 routers within the designated premises. This paper's primary contributions and differentiating factors from existing research are as follows:

- ML algorithm selection: We meticulously selected a diverse set of ML algorithms tailored to our specific tasks of classification.
- Training and Evaluation of ML Models: Our models underwent rigorous training and evaluation processes to ensure their reliability and effectiveness.
- Systematic Model Ranking: We introduced a systematic approach for ranking the considered ML models based on statistical assessments of performance metrics, thereby enhancing decision-making in selecting the most efficient ML model.

3.1. Experimentation Setup

In this work, an indoor motion detection-based testbed has been configured with two TP-Link Archer C7 Wi-Fi routers with a CSI-enabled OpenWRT image flashed on both Wi-Fi routers. The routers are placed in such a way that any movements between the routers and within the premises within their range can be captured with the help of CSI data from the received Wi-Fi signals on the receiver router. One router becomes the access point, and the other becomes the client and CSI receiver, i.e., the `recvCSI` program runs on the receiver and the `sendData` program runs on the sender router. The motion detection is estimated with the help of the deviation in the CSI data received at the receiver router. The Figure 1 shown below depicts the general context considered for the experiments in Wi-Fi sensing. It shows our experiment setup where two TP-Link archer C7 routers are placed in a room with some furniture, and a person is moving from one point to another. The dataset constructed for locally generated data samples using no occupancy in the room is labeled no movement, and when a person is present in the room with continuous movements, it is labeled movement. This data has been used in the training, testing, and validation of ML techniques.

3.2. Experimentation Procedure

The motion detection is carried out using the difference in the CSI data whenever the user moves in the target environment. The difference is analyzed from the perspective of the signal variance magnitude caused by human movement direction. CSI data are captured from the target environment for both training and testing the efficiency of machine learning algorithms. The machine learning models were trained using the benchmarking dataset, i.e., the Wi-Fi sensing data from IEEE data portal called IEEE DataPort [32] plus the locally captured dataset, and then validated using data samples from the datasets that were never used for training. The Wi-Fi CSI dataset from the IEEE data port repository, which contains a labeled dataset for humanoid motion detection, is used to train machine learning models. Thus, supervised machine learning was used for all of the training that was done on these models. As was previously noted, the Widar dataset from the IEEE for Wi-Fi sensing was initially utilized for benchmarking and comparing the performance of ML models before being replaced by the locally collected CSI-based Wi-Fi dataset. Two classes

have been employed in the locally collected dataset: one is clean with no motion at all, and the other is with a human walking. There are 35 k training samples and 9 k testing samples in the Widar dataset that was used. Eight thousand samples of training data and two thousand samples of testing data were used in the locally collected CSI dataset, which was then used for experimentation. The list of machine learning models selected for comparison is naive Bayes, support vector machine, decision tree, linear regression, K-nearest neighbor, ensemble, convolutional neural network, recurrent neural network, and deep neural network, respectively.

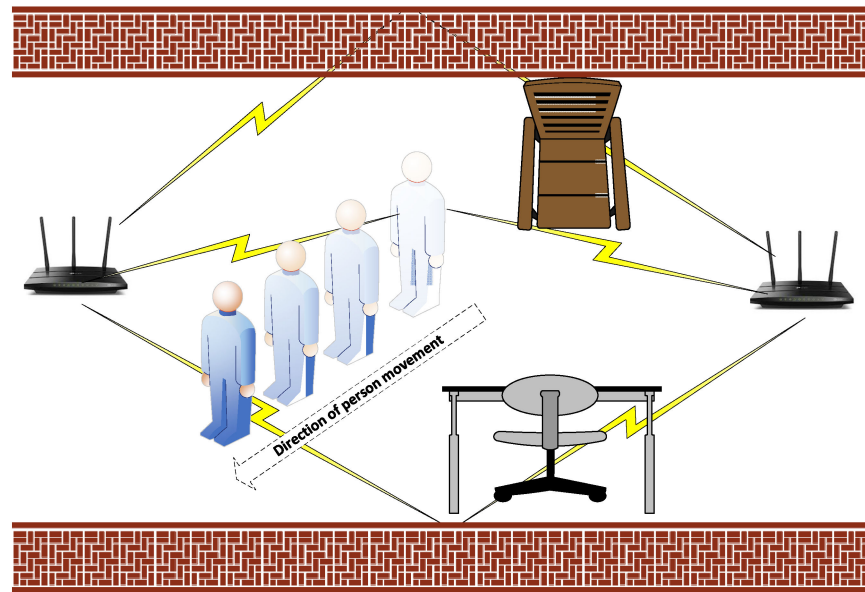


Figure 1. Wi-Fi sensing testbed setup.

3.3. Target Machine Learning Algorithms

A short description of each of the considered ML algorithms is given in the following subsections.

3.3.1. Naïve Bayes

Naive Bayes [33–38] is a probabilistic classification algorithm that has been adapted here for Wi-Fi sensing using CSI datasets for motion detection. By treating CSI measurements as features and motion/no-motion as classes, Naive Bayes has been utilized to estimate the conditional probabilities of motion given CSI values. Despite its “naive” assumption of feature independence, naive Bayes can perform well for motion detection as it works effectively with high-dimensional data like CSI. It is particularly suitable for real-time applications due to its computational efficiency and ability to handle continuous features.

3.3.2. Support Vector Machine (SVM)

SVM [39–42] is a powerful classification algorithm that has been employed here for Wi-Fi sensing with the CSI dataset for motion detection. SVM seeks to find a hyperplane that best separates instances of different classes in the feature space. In this context, SVM has been trained to classify instances based on the patterns and variations in CSI data that correspond to motion. By selecting an appropriate kernel function, SVM can effectively capture complex relationships within the dataset, aiding accurate motion detection from CSI information.

3.3.3. Decision Tree

Decision trees [43,44] are versatile machine learning models that can be used to classify instances based on a sequence of hierarchical decisions. In this context of Wi-Fi sensing, a decision tree technique has been trained using CSI data to determine the presence or absence of motion. Each decision node represents a specific feature threshold, such as changes in signal strength or frequency shifts, and the resulting branches lead to the final classification. Decision trees are interpretable and can capture non-linear relationships, making them suitable for motion detection tasks.

3.3.4. Linear Regression

While linear regression [45] is traditionally used for regression tasks, it can also be applied in a binary classification setup for motion detection, which is our target problem in motion detection using Wi-Fi sensing. By modeling the relationship between CSI features and the likelihood of motion, linear regression can provide a continuous output that represents the degree of motion. By setting a threshold on the predicted values, instances have been classified as motion or non-motion. However, linear regression might not capture complex patterns in the CSI data as effectively as other methods mentioned here.

3.3.5. K-Nearest Neighbor (KNN)

KNN [46,47] is a simple yet effective algorithm for classification tasks. It operates by assigning a class label to an instance based on the majority class of its k-nearest neighbors in the feature space. For our problem of Wi-Fi sensing with CSI data, KNN determines whether a new instance corresponds to motion based on the similarity of its CSI values to those of previously observed instances. KNN can handle non-linear relationships and is robust to noise, making it a viable option for motion detection tasks.

3.3.6. Ensemble Methods

Ensemble methods [48,49], such as random forest and gradient boosting, combine the strengths of multiple models to improve overall classification accuracy. For Wi-Fi sensing, these methods can integrate information from various CSI features to enhance the motion detection process. Random forest creates multiple decision trees and aggregates their outputs, while gradient boosting builds trees sequentially, focusing on instances that were misclassified by previous trees. These techniques can effectively capture complex patterns and variations in CSI data. However, the complexity of implementation and high computational requirements make ensemble a less popular option here.

3.3.7. Convolutional Neural Network (CNN)

CNNs [50,51] are a class of deep learning models designed to capture spatial patterns in data, particularly images. In the context of Wi-Fi sensing, CSI data has been treated as a “sequence” of signal strength values. By using 1D convolutions, CNNs learned to extract relevant features from these sequences for motion detection. This approach is effective when dealing with patterns that evolve over time, allowing the network to identify motion-related changes in the CSI dataset.

3.3.8. Recurrent Neural Network (RNN)

RNNs [52,53] are specialized for sequences and time-series data. RNN long-short-term memory (LSTM) has been employed for Wi-Fi sensing by treating the CSI dataset as a sequence of values collected over time. RNNs can learn to capture temporal dependencies and patterns in the data, making them well-suited for detecting motion. LSTM and gated recurrent unit (GRU) variants of RNNs are often used to mitigate the vanishing gradient problem and capture longer-term dependencies, but in our comparison of ML techniques, only LSTM has been considered due to the complexity and processing overhead of the GRU technique.

3.3.9. Deep Neural Network (DNN)

The fully connected deep neural network (DNN) [54] architecture has been applied to Wi-Fi sensing by directly processing CSI features to classify instances as motion or non-motion. DNNs are capable of learning intricate relationships within the data, especially where a larger amount of labeled data is available for training. The large amount of training-labeled data prevents overfitting in the case of DNN. Using appropriate activation functions, regularization techniques, and optimization algorithms, DNNs can effectively handle motion detection tasks using CSI data.

In summary, each of these machine learning techniques has its strengths and limitations when applied to Wi-Fi sensing with the CSI dataset for motion detection. The choice of technique depends on the complexity of the patterns present in the CSI data, the amount of labeled data available, and the desired trade-off between interpretability and predictive performance. Experimentation and thorough evaluation are crucial to determining the most suitable approach for a specific motion detection application. This is the central goal of our research in this article: to train, validate, and compare the selected machine learning techniques, which are designed primarily to efficiently perform classification operations. Nine different ML techniques were presented with the Wi-Fi CSI dataset, and six different performance metrics, i.e., accuracy, precision, F1-score, true positive rate (TPR), true negative rate (TNR), and false positive rate (FPR), have been observed. Now this situation raises another issue of effective comparison and systematic selection of the most suitable ML techniques, considering six different attributes. The multi-attribute decision-making (MADM) technique has been employed to solve this problem. Here we have accuracy, precision, F1-score, true positive rate (TPR), true negative rate (TNR) as positive attributes, and false positive rate (FPR) as a negative rate. The weight assigned to these performance parameters is as follows: The accuracy has been assigned the highest weight as it is the most important performance parameter; FPR is assigned as 2nd most important parameter as more false positive occurrences can lead the model to higher inaccuracies; precision is followed by the F1-score; TPR is next; and TNR, which is a positive parameter, is the least important in the attributes list. The following section analyzes the results obtained using each of the considered ML techniques when employed on the same Wi-Fi CSI dataset for motion detection.

4. Results Analysis

This section presents the results obtained for motion detection using Wi-Fi sensing when a set of different machine learning models were exposed to the dataset. Analyzing the performance of ML models for Wi-Fi sensing typically involves a combination of some standard metrics [55,56] and evaluation methods, which include confusion matrix and its derived metrics, receiver operating characteristics curve (ROC), cross validation, etc. For performance comparison in this paper, the following set of performance metrics have been considered, which have been derived from the confusion matrix:

Accuracy, false positive rate, precision, F1-score, true positive rate, and true negative rate. Each of these performance metrics has been compared when the same set of datasets is applied to the trained machine learning models.

Figure 2, shown below, presents the accuracy rate values for each of the ML algorithms when these algorithms were presented with the testing segment of the dataset. It shows that deep learning algorithms, i.e., DNN, RNN, and CNN, are performing distinctively well as compared to shallow learning algorithms.

Figure 3, shown below, presents the precision rate results from each ML algorithm when presented with the sampling dataset segment for motion detection. The precision values for deep neural network models outperform the precision values of shallow learning, except for the RNN in deep learning, which shows very low precision values.

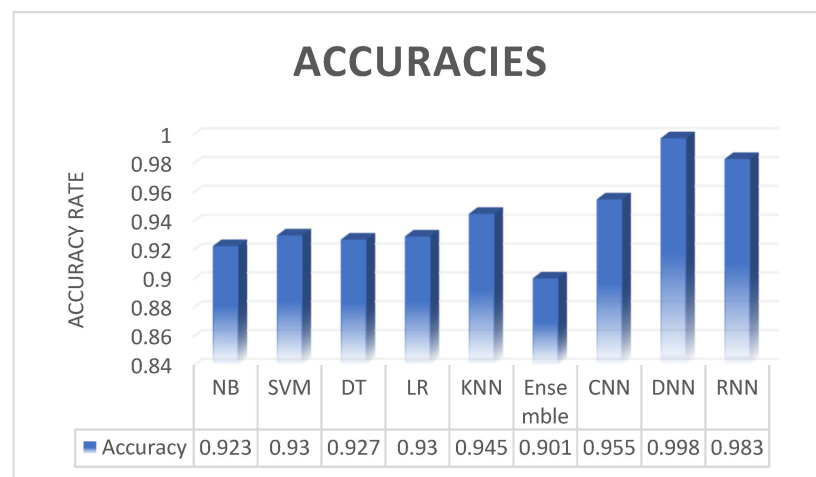


Figure 2. Rate of Accuracies for different ML algorithms.

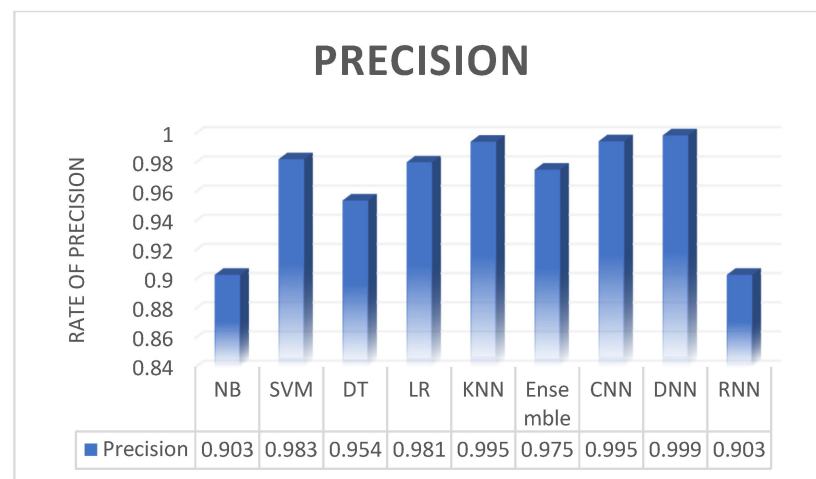


Figure 3. Precision rate for all ML algorithms.

The true positive rate values for each of the ML algorithms are depicted here in Figure 4 below. It shows that the DNN outperforms not only the other deep learning algorithms but also all the shallow learning models.

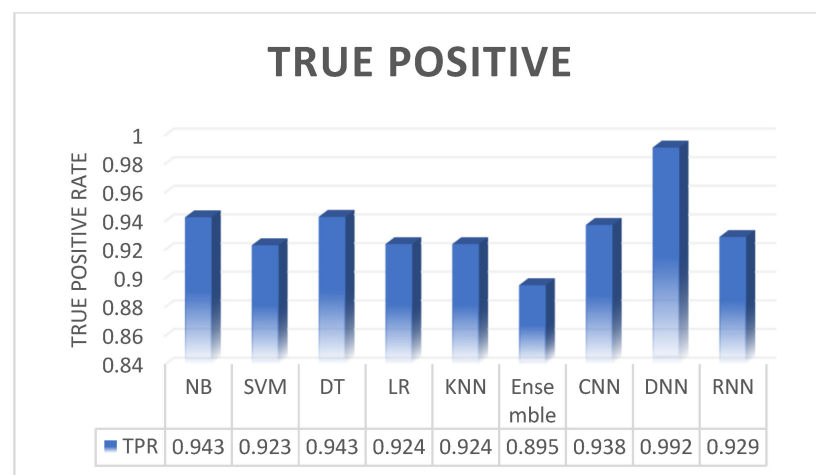


Figure 4. True positive rate for all ML algorithms.

Once all the performance metrics have been recorded using all the target ML algorithms, then comes another challenge to compare the performance metrics of each ML algorithm to see which is the most optimal choice amongst the considered ML algorithms for motion detection using Wi-Fi CSI data. This is a multi-dimensional and multi-criteria problem that can be best resolved using the multiple-attribute decision-making (MADM) algorithm. Once all the performance metrics data from all the ML algorithms have been recorded, a score is added to each ML algorithm using the MADM algorithm to see which ML algorithm is performing better considering all the performance metrics at once.

The MADM [57,58] has been applied to evaluate and rank different machine learning algorithms based on their performance across various criteria (attributes). In this case, the decision matrix consists of rows representing different machine learning algorithms and columns representing different performance metrics (attributes) such as accuracy, FPR (false positive rate), precision, F1-score, TPR (true positive rate), and TNR (true negative rate). The goal of the MADM analysis is to rank these machine learning algorithms based on their overall performance across these attributes. The result of the MADM analysis is presented in the “Scores” column, and the algorithms are ranked based on these scores from low to high. Here are the general steps for applying MADM to the statistics in Table 2:

- Define the Decision Problem: The decision problem is to determine the best-performing machine learning algorithm among the given options based on multiple performance criteria.
- Data Collection: Data for each algorithm’s performance across different criteria is collected and presented in columns 2 to 7.
- Weight Assignment: The importance or weight of each performance metric (attribute) is determined. Depending on the specific MADM method used, these weights could be assigned subjectively based on domain knowledge or obtained through a formal weighting process.
- Normalization: If needed, data in columns 2 to 7 may be normalized to ensure that all metrics are on a comparable scale. Normalization is important when metrics have different units or scales.
- Decision Matrix: Create a decision matrix where each row represents an algorithm and each column represents a performance metric (attribute). The values in this matrix are the algorithm’s performance scores for each metric, possibly weighted.
- MADM Method: Choose an appropriate MADM method for ranking the alternatives. Common methods include the technique for order of preference by similarity to the ideal solution (TOPSIS), the analytic hierarchy process (AHP), and the weighted sum model, among others. In our case, we have selected the weighted sum method and assigned the weights to attributes such as accuracy as the highest and true negative rate as the lowest.
- Ranking or Scoring: Apply the weighted sum MADM method to the decision matrix to calculate an overall score or ranking for each algorithm. This score reflects the algorithm’s performance across all criteria, considering their weights.
- Result in Column 8: The “Scores” column (column 8) contains the results of the MADM analysis. Each algorithm is assigned a score based on its overall performance.
- Ranking: The algorithms are then ranked based on their scores in descending order. The algorithm with the highest score is typically considered the best-performing one.

Table 2. MADM scoring on ML algorithms performance scores.

	MDMA Scores						Scores	Rank
	Accuracy	FPR	Precision	F1 (+)	TPR	TNR (+)		
NB	0.9228	0.1103	0.9033	0.9227	0.9428	0.9293	21.3288	9
SVM	0.9302	0.1336	0.9826	0.9521	0.9233	0.8919	23.44	6
DT	0.9273	0.0842	0.9543	0.9487	0.9431	0.8267	24.1331	5
LR	0.9295	0.1269	0.9805	0.9516	0.9241	0.886	23.4146	7
KNN	0.9452	0.2825	0.9946	0.9582	0.9241	0.9503	24.6198	4
Ensemble	0.9005	0.0998	0.9754	0.9338	0.8954	0.8544	21.8578	8
CNN	0.9552	0.2825	0.9949	0.9655	0.9375	0.9503	25.4966	3
DNN	0.9976	0.0134	0.9989	0.9935	0.9915	0.8456	27.7903	1
RNN	0.9832	0.0128	0.9033	0.9555	0.9291	0.8811	26.0445	2

In the Table 2, the algorithms have been ranked based on their scores in the “Scores” column, from the highest score (rank 1) to the lowest score (rank 9). The Table 2 above clearly shows that deep learning algorithms have outperformed all the shallow learning algorithms collectively when the MADM algorithm has been applied to rank the best-performing ML algorithms.

5. Discussion

The current literature study represents a significant advancement in the field of Wi-Fi sensing for motion detection, particularly within the context of Gamgee BV in the Netherlands. The primary objective of this research was to explore the integration of ML techniques into Wi-Fi sensing technology for improved motion detection. This milestone was achieved through the careful consideration and evaluation of a diverse set of ML algorithms, encompassing both SL and DL approaches. The utilization of publicly available datasets, including CSI datasets from the IEEE data port and locally captured datasets, was integral to benchmarking the performance of various ML models. These datasets served as valuable resources for training, validating, and testing the developed ML algorithms. Our focus was specifically directed towards classification-based ML algorithms tailored for motion detection. The array of algorithms assessed in the study encompassed a range of SL and DL models. Among the SL algorithms, SVM, naïve Bayes, decision tree, K-NN, and K-means algorithms were systematically evaluated. Additionally, we delved into the realm of DL algorithms, considering RNN, CNN, and DNN. Through meticulous performance analysis, we compared the efficiency of each algorithm, eventually leading to the identification of the most suitable ML algorithms for motion detection via Wi-Fi sensing using CSI data captured from TP-Link Archer C7 routers deployed within the target premises. Our findings underscored the superiority of DL algorithms, specifically DNN and RNN, in scenarios where larger datasets were utilized for training and validation. These DL models exhibited remarkable performance gains when exposed to extensive datasets, outperforming their SL counterparts by a significant margin. This outcome emphasizes the potential of DL techniques to enhance the accuracy and efficacy of motion detection via Wi-Fi sensing.

While this study represents a substantial leap forward in identifying the most suitable ML technique for motion detection using the Wi-Fi CSI dataset and the integration of ML with Wi-Fi sensing, certain limitations warrant consideration. First, the effectiveness of the selected ML algorithms might be influenced by variations in environmental conditions, potentially impacting the consistency of motion detection results. Additionally, the generalization of the trained models to different premises and contexts remains an aspect that requires validation. These limitations can, of course, be tackled with countermeasures such as the deployment of a sufficient number of Wi-Fi devices in the target premises, which will eventually also improve the performance of Wi-Fi connectivity at the same time. The dataset selection, although carefully considered, might not encompass the full spectrum of real-world scenarios, leading to potential biases in the developed models. Furthermore, the

computational resources required for DL algorithms can be substantial, posing challenges for real-time implementation in resource-constrained environments. The solution to these limitations can be a more comprehensive dataset for training the ML model and the use of networking devices such as routers with higher specifications to handle the higher computational requirements, particularly in the case of ML models.

To contextualize our findings and highlight their relevance in the broader research landscape, it is essential to draw parallels with existing studies. Recent research in the field of Wi-Fi sensing and related domains has showcased a similar trend favoring deep learning approaches. Prominent works by Yongsan et al. [59] and Atzeni et al. [60] have reported remarkable success in employing deep neural networks for Wi-Fi-based applications. These studies have emphasized the ability of deep learning models to extract intricate patterns and representations from CSI data, leading to enhanced accuracy and reliability in Wi-Fi sensing tasks. Among the DL algorithms, DNN excelled with a remarkable accuracy of 0.9976. This performance surpasses recent work in [61] with an accuracy of 99.76%, establishing DNN as the leading choice for Wi-Fi sensing applications. The results in [61] show maximum accuracies of 99.38 for DL models such as RNN and CNN in different versions. Furthermore, with much lower accuracies for SL algorithms such as naive Bayes, SVM, and KNN, there is a similar trend to the results shown in this paper for SL techniques. In [62], the authors obtained a maximum accuracy of 98.2% using the DL technique for crowd estimation on CSI data obtained from Wi-Fi. Though the goal of that work was crowd estimation, Wi-Fi CSI data was utilized with ML techniques to achieve it. The accuracy achieved was closer to that presented in our research work, which still surpasses it with a difference of 1.56%.

In comparison to these contemporary research outcomes, our study corroborates the growing consensus that DL, particularly DNN and RNN architectures, represents a potent tool for Wi-Fi sensing applications. The exceptional accuracy and efficiency demonstrated by these DL algorithms in our experimentation underscore their viability in real-world scenarios, where robust Wi-Fi sensing is essential for diverse applications such as indoor localization, occupancy detection, and smart home automation. In conclusion, our study not only contributes valuable insights into the selection of suitable algorithms for Wi-Fi sensing using CSI data but also aligns with and reinforces the findings of recent research in the field. The superior performance of DL algorithms, as highlighted in our results, positions them as promising candidates for addressing the evolving challenges and opportunities in Wi-Fi sensing applications.

The future trajectory of this research is marked by several compelling avenues. Expanding our focus on localization holds great promise, as the ability to precisely identify the location of detected motion could significantly enhance security and monitoring applications. The automation of model learning within the target premises is a critical step towards achieving seamless and adaptable motion detection systems. The integration of Wi-Fi sensing with home automation and healthcare represents a paradigm shift with immense potential. Exploring the feasibility of leveraging Wi-Fi CSI data and AI for enhanced automation, ambient intelligence, and personalized healthcare interventions is an exciting direction for future investigation. In conclusion, the current work not only sets a foundation for ML-driven Wi-Fi sensing but also opens doors to a plethora of innovative applications. The journey from motion detection to localization, automation, and healthcare integration underscores the dynamic and transformative nature of this research trajectory.

6. Conclusions

The work performed for this article was the first milestone to introduce ML in Wi-Fi sensing for motion detection at Gamgee BV in the Netherlands. We have considered several ML algorithms composed of both shallow learning and deep learning algorithms. The publicly available datasets, i.e., CSI datasets from the IEEE data port and locally captured datasets, have been utilized for benchmarking the ML models before applying the testing segment of the datasets to look for the most suitable ML algorithms for motion detection

using Wi-Fi sensing. We employed classification-based ML algorithms for operations like motion detection, which is part of the research work presented in this article, and clustering-based ML algorithms for unique personal identification in other subsequent research work being carried out at Gamgee BV. The list of SL algorithms considered includes SVM, naïve Bayes, decision tree, K-nearest neighbors, and K-means algorithms, and the list of DL algorithms considered includes recurrent neural network (RNN), convolutional neural networks (CN), and deep neural networks (DNN), respectively. The performance metrics were analyzed from a set of results obtained using each of the ML algorithms considered to compare the performance and efficiency of each ML algorithm and select the most suitable for Wi-Fi sensing using the CSI data captured from TP-Link Archer C7 routers in the target premises. Our results showed that DL algorithms, i.e., DNN and RNN, performed much better as compared to the SL algorithms when larger datasets were exposed to the ML models for training and validation purposes. Our research has already been extended to further include localization to identify the exact zone where motion was detected and automation of model learning in target premises. The research work will be further extended to include home automation and healthcare applications using Wi-Fi CSI data and artificial intelligence (AI)-augmented Wi-Fi sensing.

Author Contributions: Conceptualization, M.A., P.H., N.P. and S.L.; Software, M.A.; Validation, M.A. and A.N.; Investigation, M.A.; Writing—original draft, M.A.; Writing—review & editing, P.H. and A.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to company data protection and privacy policies.

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

References

1. Tekler, Z.D.; Low, R.; Gunay, B.; Andersen, R.K.; Blessing, L. A scalable Bluetooth Low Energy approach to identify occupancy patterns and profiles in office spaces. *Build. Environ.* **2020**, *171*, 106681. [\[CrossRef\]](#)
2. Mukhopadhyay, S.C. Wearable Sensors for Human Activity Monitoring: A Review. *IEEE Sens. J.* **2015**, *15*, 1321–1330. [\[CrossRef\]](#)
3. Moeslund, T.B.; Hilton, A.; Krüger, V. A survey of advances in vision-based human motion capture and analysis. *Comput. Vis. Image Underst.* **2006**, *104*, 90–126. [\[CrossRef\]](#)
4. Zieger, C.; Brutti, A.; Svaizer, P. Acoustic Based Surveillance System for Intrusion Detection. In Proceedings of the 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, Genova, Italy, 2–4 September 2009; pp. 314–319. [\[CrossRef\]](#)
5. Sahoo, K.C.; Pati, U.C. IoT based intrusion detection system using PIR sensor. In Proceedings of the 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 19–20 May 2017; pp. 1641–1645. [\[CrossRef\]](#)
6. Filippoupolitis, A.; Oliff, W.; Loukas, G. Bluetooth Low Energy Based Occupancy Detection for Emergency Management. In Proceedings of the 2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS), Granada, Spain, 14–16 December 2016; pp. 31–38. [\[CrossRef\]](#)
7. Tekler, Z.D.; Low, R.; Yuen, C.; Blessing, L. Plug-Mate: An IoT-based occupancy-driven plug load management system in smart buildings. *Build. Environ.* **2022**, *223*, 109472. [\[CrossRef\]](#)
8. Balaji, B.; Xu, J.; Nwokafor, A.; Gupta, R.; Agarwal, Y. Sentinel: Occupancy based HVAC actuation using existing wifi infrastructure within commercial buildings. In Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems, Rome, Italy, 11–15 November 2013; Association for Computing Machinery: New York, NY, USA, 2013. [\[CrossRef\]](#)
9. Tekler, Z.D.; Chong, A. Occupancy prediction using deep learning approaches across multiple space types: A minimum sensing strategy. *Build. Environ.* **2022**, *226*, 109689. [\[CrossRef\]](#)
10. He, Z.; Zhang, X.; Wang, Y.; Lin, Y.; Gui, G.; Gacanin, H. A Robust CSI-based Wi-Fi Passive Sensing Method Using Attention Mechanism Deep Learning. *IEEE Internet Things J.* **2023**, *1*, 1. [\[CrossRef\]](#)
11. Taylor, W.; Khan, M.Z.; Tahir, A.; Taha, A.; Abbasi, Q.H.; Imran, M.A. An Implementation of Real-Time Activity Sensing Using Wi-Fi: Identifying Optimal Machine-Learning Techniques for Performance Evaluation. *IEEE Sens. J.* **2022**, *22*, 21127–21134. [\[CrossRef\]](#)
12. Khan, M.Z.; Ahmad, J.; Boulila, W.; Broadbent, M.; Shah, S.A.; Koubaa, A. Contactless Human Activity Detection using Deep Learning with Flexible and Scalable Software Define Radio. *arXiv* **2023**, arXiv:2304.09756.

13. Bastwesly, M.R.M.; ElShennawy, N.M.; Saidahmed, M.T.F. Deep Learning Sign Language Recognition System Based on Wi-Fi CSI. *Int. J. Intell. Syst. Appl.* **2020**, *12*, 33–45. [\[CrossRef\]](#)
14. Damodaran, N.; Haruni, E.; Kokhkhharova, M.; Schäfer, J. Device free human activity and fall recognition using WiFi channel state information (CSI). *CCF Trans. Pervasive Comput. Interact.* **2020**, *2*, 1–17. [\[CrossRef\]](#)
15. Ahmed, H.F.T.; Ahmad, H.; Aravind, C.V. Device free human gesture recognition using Wi-Fi CSI: A survey. In Proceedings of the IEEE Latin-American Conference on Communications, Rio de Janeiro, Brazil, 30 November–2 December 2022; Volume 2022, pp. 13086–13095. [\[CrossRef\]](#)
16. Yang, J.; Chen, X.; Zou, H.; Wang, D.; Xu, Q.; Xie, L. EfficientFi: Toward Large-Scale Lightweight WiFi Sensing via CSI Compression. *IEEE Internet Things J.* **2022**, *9*, 13086–13095. [\[CrossRef\]](#)
17. Bocus, M.J.; Chetty, K.; Piechocki, R.J. A Comparison of UWB CIR and WiFi CSI for Human Activity Recognition. In Proceedings of the IEEE Radar Conference (RadarCon), Atlanta, GA, USA, 10–14 May 2021.
18. Hu, J.; Yang, J.; Ong, J.B.; Wang, D.; Xie, L. ResFi: WiFi-Enabled Device-Free Respiration Detection Based on Deep Learning. In Proceedings of the 2022 IEEE 17th International Conference on Control & Automation (ICCA), Naples, Italy, 27–30 June 2022; pp. 510–515. [\[CrossRef\]](#)
19. Soto, J.C.; Galdino, I.; Gouveia, B.G.; Caballero, E.; Ferreira, V.; Muchaluat-Saade, D.; Albuquerque, C. Wi-Fi CSI-based Human Presence Detection Using DTW Features and Machine Learning. In Proceedings of the 2022 IEEE Latin-American Conference on Communications (LATINCOM), Rio de Janeiro, Brazil, 30 November–2 December 2022; pp. 1–6. [\[CrossRef\]](#)
20. Adib, F.; Kabelac, Z.; Katabi, D.; Miller, R.C. 3D tracking via body radio reflections. In Proceedings of the 11th USENIX Conference on Networked Systems Design and Implementation, Seattle, WA, USA, 2–4 April 2014; pp. 317–329.
21. Wang, Y.; Liu, J.; Chen, Y.; Gruteser, M.; Yang, J.; Liu, H. E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures. In Proceedings of the MobiCom '14: 20th Annual International Conference on Mobile Computing and Networking, Maui, HI, USA, 7–11 September 2014; pp. 617–628. [\[CrossRef\]](#)
22. Wang, W.; Liu, A.X.; Shahzad, M.; Ling, K.; Lu, S. Understanding and modeling of WiFi signal based human activity recognition. In Proceedings of the 21st Annual International Conference on Mobile Computing and Networking, Paris, France, 7–11 September 2015; pp. 65–76. [\[CrossRef\]](#)
23. Adib, F.; Hsu, C.-Y.; Mao, H.; Katabi, D.; Durand, F. Capturing the human figure through a wall. *ACM Trans. Graph.* **2015**, *34*, 219. [\[CrossRef\]](#)
24. Hsu, C.-Y.; Liu, Y.; Kabelac, Z.; Hristov, R.; Katabi, D.; Liu, C. Extracting gait velocity and stride length from surrounding radio signals. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 2116–2126. [\[CrossRef\]](#)
25. Hsu, C.-Y.; Ahuja, A.; Yue, S.; Hristov, R.; Kabelac, Z.; Katabi, D. Zero-effort in-home sleep and insomnia monitoring using radio signals. *ACM Interact. Mobile Wearable Ubiquitous Technol.* **2017**, *1*, 59. [\[CrossRef\]](#)
26. Zhao, M.; Adib, F.; Katabi, D. Emotion recognition using wireless signals. *Commun. ACM* **2018**, *61*, 91–100. [\[CrossRef\]](#)
27. 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.* **2012**, *50*, 1218–1226. [\[CrossRef\]](#)
28. Adib, F.; Katabi, D. See through walls with WiFi! In Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM, Hong Kong, China, 12–16 August 2013; Association for Computing Machinery: New York, NY, USA, 2013; pp. 75–86.
29. Pu, Q.; Gupta, S.; Gollakota, S.; Patel, S. Whole-home gesture recognition using wireless signals. In Proceedings of the 19th Annual International Conference on Mobile Computing & Networking, Miami, FL, USA, 30 September–4 October 2013; pp. 27–38.
30. Tan, B.; Woodbridge, K.; Chetty, K. A wireless passive radar system for real-time through-wall movement detection. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 2596–2603. [\[CrossRef\]](#)
31. Tan, B.; Woodbridge, K.; Chetty, K. A real-time high resolution passive WiFi doppler-radar and its applications. In Proceedings of the 2014 International Radar Conference, Lille, France, 13–17 October 2014; pp. 1–6.
32. Meneghello, F.; Fabbro, N.D.; Garlisi, D.; Tinnirello, I.; Rossi, M. CSI Dataset for Wireless Human Sensing on 80 MHz Wi-Fi Channels. *IEEE Commun. Mag.* **2023**, 1–6. [\[CrossRef\]](#)
33. Lewis, D.D. Naïve (Bayes) at forty: The independence assumption in information retrieval. In Proceedings of the Machine Learning: ECML-98: 10th European Conference on Machine Learning, Chemnitz, Germany, 21–23 April 1998; Springer: Berlin/Heidelberg, Germany, 1998.
34. Sahami, M.; Dumais, S.; Heckerman, D.; Horvitz, E. A Bayesian approach to filtering junk e-mail. In *Learning for Text Categorization: Papers from the 1998 Workshop*; AAAI Technical Report WS-98-05; AAAI: Washington, DC, USA, 1998.
35. Rish, I. An empirical study of the Naïve Bayes classifier. In *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, Seattle, WA, USA, 4–6 August 2001; IBM: New York, NY, USA, 2001; pp. 41–46.
36. Jeevaraj, D.; Karthik, B.; Vijayan, T. Feature Selection Model using Naive Bayes ML Algorithm for WSN Intrusion Detection System. *Int. J. Electr. Comput. Eng. Syst.* **2023**, *14*, 179–185. [\[CrossRef\]](#)
37. Duda, R.; Hart, P.E.; Stork, D.G. *Pattern Classification*; The Naïve Bayes Classifier is Introduced in Chapter 2; Wiley: Hoboken, NJ, USA, 2000.
38. Hastie, T.; Tibshirani, R. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Chapter 6 Covers Naïve Bayes; Springer: Berlin/Heidelberg, Germany, 2009.

39. Guo, Y.; Fei, R.; Yang, L.; Bai, X.; Liu, Z.; Chen, X. A Passive Indoor Population Detection Method Based on WiFi Channel Status Information. In Proceedings of the 2022 6th Asian Conference on Artificial Intelligence Technology (ACAIT), Changzhou, China, 4–6 November 2022; pp. 1–6. [\[CrossRef\]](#)
40. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000.
41. Schölkopf, B. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, USA, 2002; ISBN 0-262-19475-9.
42. Vapnik, V. *The Nature of Statistical Learning Theory*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2013; ISBN 0-387-98780-0.
43. Castillo-Botón, C.; Casillas-Pérez, D.; Casanova-Mateo, C.; Ghimire, S.; Cerro-Prada, E.; Gutierrez, P.A.; Deo, R.C.; Salcedo-Sanz, S. Machine learning regression and classification methods for fog events prediction. *Atmos. Res.* **2022**, *272*, 106157. [\[CrossRef\]](#)
44. Chen, X.; Ishwaran, H. Random forests for genomic data analysis. *Genomics* **2018**, *110*, 119–129. [\[CrossRef\]](#)
45. Montgomery, D.C.; Peck, E.A.; Vining, G.G. *Introduction to Linear Regression Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2021; p. 704, ISBN 978-1-119-57875-8.
46. Bathija, R.; Bajaj, V.; Meghani, C.; Sawara, J.; Mirchandani, S. Revolutionizing Recruitment: A Comparative Study of KNN, Weighted KNN, and SVM—KNN for Resume Screening. In Proceedings of the 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 19–21 June 2023; pp. 834–840. [\[CrossRef\]](#)
47. Duda, R.O.; Hart, P.E. *Pattern Classification*; Wiley: Hoboken, NJ, USA, 2000; ISBN 978-0-471-05669-0.
48. Dietterich, T.G. Ensemble Methods in Machine Learning. In *Multiple Classifier Systems*; MCS 2000. Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2000; Volume 1857. [\[CrossRef\]](#)
49. Sagi, O.; Rokach, L. Ensemble Learning: A Survey. *WIREs Data Mining Knowl. Discov.* **2018**, *8*, e1249. [\[CrossRef\]](#)
50. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst. (NIPS)* **2012**, *25*. [\[CrossRef\]](#)
51. Sarikaya, R.; Hinton, G.E. A Survey of Deep Learning Architectures and Their Applications. *IEEE Signal Process. Mag.* **2019**, *234*, 11–26. [\[CrossRef\]](#)
52. Ghojogh, B.; Ghodsi, A. Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and Survey. *arXiv* **2023**, arXiv:2304.11461v1. [\[CrossRef\]](#)
53. Lipton, Z.C.; Berkowitz, J.; Elkan, C. A Critical Review of Recurrent Neural Networks for Sequence Learning. *arXiv* **2017**, arXiv:1506.00019.
54. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016. Available online: <http://www.deeplearningbook.org> (accessed on 27 July 2023).
55. Powers, D.M.W. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Int. J. Mach. Learn. Technol.* **2011**, *2*, 37–63.
56. Gong, M. A novel performance measure for machine learning classification. *Int. J. Managing Inf. Technol.* **2021**, *13*. [\[CrossRef\]](#)
57. Ishizaka, A.; Nemery, P. *Multi-Criteria Decision Analysis: Methods and Software*; Wiley: Hoboken, NJ, USA, 2013; p. 320, ISBN 978-1-119-97407-9.
58. Daugavietis, J.E.; Soloha, R.; Dace, E.; Ziemeles, J. A Comparison of Multi-Criteria Decision Analysis Methods for Sustainability Assessment of District Heating Systems. *Energies* **2022**, *15*, 2411. [\[CrossRef\]](#)
59. Ma, Y.; Zhou, G.; Wang, S. WiFi Sensing with Channel State Information: A Survey. *ACM Computer Survey* **2019**, *52*, 46. [\[CrossRef\]](#)
60. Atzeni, D.; Bacciu, D.; Mazzei, D.; Prencipe, G. A Systematic Review of Wi-Fi and Machine Learning Integration with Topic Modeling Techniques. *Sensors* **2022**, *22*, 4925. [\[CrossRef\]](#)
61. Yang, J.; Chen, X.; Zou, H.; Wang, D.; Xie, L. AutoFi: Toward Automatic Wi-Fi Human Sensing via Geometric Self-Supervised Learning. *IEEE Internet Things J.* **2023**, *10*, 7416–7425. [\[CrossRef\]](#)
62. Choi, H.; Fujimoto, M.; Matsui, T.; Misaki, S.; Yasumoto, K. Wi-CaL: WiFi Sensing and Machine Learning Based Device-Free Crowd Counting and Localization. *IEEE Access* **2022**, *10*, 24395–24410. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.