Human Activity Detector Using Wi-Fi With The Help of Deep Learning

Mohit Parikh
Student, Department of Computer Engineering, Devang Patel Institute of Advance Technology and Research, Charusat University, Anand, Gujarat, India

Keval Bhavsar
Student, Department of Computer Engineering, Devang Patel Institute of Advance Technology and Research, Charusat University, Anand, Gujarat, India

Prof. Urvashi Chaudhari
Faculty, Department of Computer Engineering, Devang Patel Institute of Advance Technology and Research, Charusat University, Anand, Gujarat, India

Abstract- We present a review of recent advances in passive human behavior recognition in indoor areas using commercial Wi-Fi systems' channel state information (CSI). Changes in wireless signal reflections are caused by human movement, resulting in CSI variations. Human behavior can be identified by analyzing CSI data streams for various activities and comparing them to stored models. This is accomplished by extracting features from CSI data streams and developing models and classifiers using machine learning techniques. The techniques presented herein have excellent results; however, instead of using machine learning techniques, we propose using deep learning techniques such as long-short term memory (LSTM) recurrent neural network (RNN) and demonstrating improved performance. We also discuss various challenges, such as changing the environment, choosing a frame rate, and dealing with a multi-user scenario, as well as possible future research directions.

Keywords: Behavior Recognition, channel state information, long-short term memory, machine learning, OFDM, Wi-Fi.

INTRODUCTION

Human Behavior and Activity Recognition has always been a popular research topic, owing to the wide range of applications in a variety of fields. The ability to monitor human behavior without physically attaching sensors to their bodies will have a significant impact on surveillance, healthcare, and a variety of other smart-home applications. In an ideal world, such monitoring would be device-free, requiring no special devices to be worn by the subjects being monitored. Due to their inherent violation of privacy, recording devices such as cameras may not be acceptable solutions as sensors for monitoring.

The use of Wi-Fi data for behavior recognition has gotten a lot of attention lately, thanks to the widespread availability of Wi-Fi-enabled devices. When a wireless access point and a connected device are present in the environment, analyzing changes in the received signal can provide enough information to recognize the activity. The signal transmitted by the access point travels across the room to the receiving device. As the signal travels through space, it interacts with its surroundings as it reflects off various surfaces and objects, including people. Different movements of body parts will result in recognizable variations in the received signal, according to WiFi-based behavior recognition.

In current systems, the user must wear a device with motion sensors such as a gyroscope and an accelerometer. Sensor data is processed locally on the wearable device or transmitted to a server for feature extraction, and then supervised learning algorithms are used for classification. This is referred to as active monitoring. The performance of such a system is around 90% for recognizing activities such as sleeping, sitting, standing, walking, and running.

However, always wearing a device is inconvenient and maybe impossible in many passive activity recognition applications where the person is not carrying any sensor or wireless device. While camera-based systems can be used to detect passive activity, the requirement for line-of-sight (LOS) is a major drawback. Additionally, camera-based approaches have privacy concerns and cannot be used in many situations. As a result, a passive monitoring system based on wireless signals that do not infringe on people's privacy is desired.

WiFi has recently been the subject of a lot of research for activity recognition because of its widespread availability in indoor areas. A WiFi access point (AP) and one or more Wi-Fi-enabled devices are used in such systems. Body movement affects wireless signals and changes the multi-path profile of the system when a person engages in an activity.

In this paper, we present a novel recurrent neural network-based classification system for CSI-based activity recognition that significantly improves on the state-of-the-art. The basic task provides several significant challenges. The data is skewed. It's difficult to pinpoint the beginning and end of behaviors in sensor-data streams. Behavior patterns can be quite complex in terms of time. All of these issues are addressed in our proposed solution. To remove the most corrupting components of noise, we introduce a learned subspace-projection de-noising algorithm. We solve the end pointing problem with a shift-invariant feature extractor that naturally accounts for timing uncertainty. A properly structured recurrent network model is used to address the complexity of temporal patterns.

1.1 RELATED WORK

The data from Wi-Fi Received Signal Strength (RSS) has been used in a variety of activity and behavior recognition tasks. The Wi-Fi signal strength varies as a human body passes between the transmitter and the receiver. For classification, Wi-Fi RSS behavior recognition techniques rely on this variation in signal strength. RSS data has been successfully used for tasks such as gesture recognition, which are similar to activity recognition. Software Defined Radio and modified wireless hardware are used for RSS-based behavior and activity recognition. RSS data has been shown to provide sufficient information for locating activities within the environment, but it does not capture fine-grained changes in the signal caused by human body reflections. Channel State Information has been demonstrated to work better than RSS approaches because it captures these tiny changes.

For behavior recognition, systems that leverage WiFi Channel State Information (CSI) employ a variety of categorization techniques. On pre-processed CSI data, a histogram-based classifier was utilized to categorize activity using signal modulation. The histogram comparison method is likely to be susceptible to any little changes in the environment because this system does not extract any high-level features. Furthermore, because of burst and impulse noise, a low-pass filter applied to CSI data may not be able to remove static multipath noise created by the environment.

CARM is another behavior recognition system that makes use of CSI data. Static-multipath noise is reduced in this system by using Principal Component Analysis (PCA) on the CSI data, removing the first component, and using the remaining five vectors. The remaining key components are expected to capture the dynamic multi-path information created by reflections of human body parts. Using a Hidden Markov Model classifier, this strategy has been shown to produce high classification accuracy, and the de-noising methods utilized make it significantly less sensitive to changes in the environment.

For classification, the most recent method for behavior identification use a deep learning model. For categorization, this study employs a Long Short-Term Memory (LSTM) network that has been trained using unprocessed CSI data. When this system was compared to the approaches proposed by the CARM model, it was discovered that the employment of LSTMs outperformed the HMM-based strategy by a large margin. This technique is susceptible to changes in the environment because of the lack of signal pre-processing and de-noising.

In this research, we apply a PCA model that is similar to CARM's de-noising techniques. To classify behavior, we also use feature extractors based on neural networks. We assess our system's performance and compare it to that of other methods.

1.2 METHODOLOGY

1. Wi-Fi Channel State Information: Commonly, we use Wi-Fi for communication purposes, but it could also be used for solving the HARD problem. Human movement within the range of a Wi-Fi network affects multipath propagation.

Why this happens: a Wi-Fi router works with radio waves that propagate through its location environment, and when a person is on the way of the radio wave, it gets around this person and its characteristics will change. Moreover, the Wi-Fi channel also consists of a signal reflected by static objects in some environments, such as furniture or others. As the Channel State Information represents the signal propagation effect in the channel, these additional reflections caused by different activities are observed.

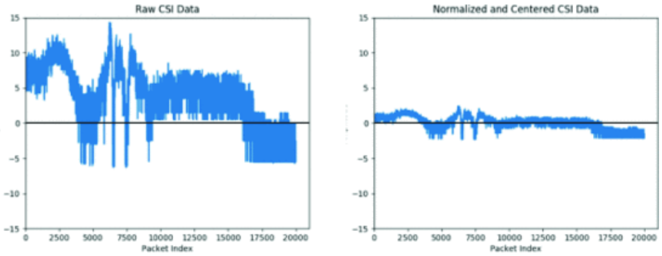
More precisely, CSI represents the change of the signal from the transmitter (denote it as x) to the receiver (denote it as y) and the Wi-Fi channel in the frequency domain can be represented as:

$$\mathbf{y = Hx + n}$$

where H is a complex matrix consisting of CSI values and n is the channel noise. The CSI is estimated for each Orthogonal Frequency Division Multiplexing (OFDM) subcarrier. OFDM splits the total frequency spectrum into 56 or 114 frequency subcarriers for a channel bandwidth of 20 and 40 MHz respectively. The CSI for each subcarrier is:

$$\mathbf{h = |h|e\ j\theta}$$

where |h| represents the amplitude and θ the phase. To measure the CSI, the transmitter sends Long Training Symbol (LTS), which contains predefined information for each subcarrier. When the receiver receives the LTS, it estimates the CSI having the difference between the original and received LTS. However, in real-world systems, the CSI is affected by a multi-path channel, receiver/transmit processing, hardware, and software errors.



In our research, the CSI matrix consists of 114 complex H matrices for the 5 GHz frequency band or 56 for 2.4 GHz of dimension NTx xNRx. 56 and 114 are defined by the number of subcarriers the CSI extraction tool can handle (on 2.4 and 5 GHz).

Dataset Description

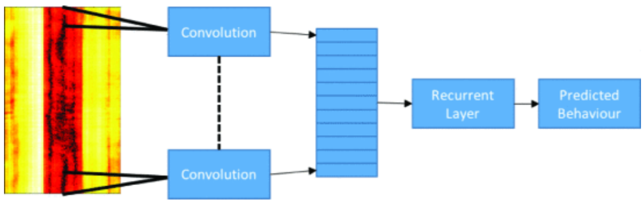
1. Neural Network: the mathematical model inspired by a study of the biological structure and organization of biological neural networks in the human brain. More precisely, it is a sequence of neurons, connected via synapses. Based on such architecture, the model can analyze, remember or even create new information as it works with the human brain, which consists of millions of neurons that exchange information based on electrical impulses.

The training consists of two steps:

1) The trained data is forwarded through the network. The output and error between the network prediction and the target values from the dataset are calculated.

2) This error backpropagates through the network and weight is updated concerning the error gradients.

This cycle is repeated many times until the network is stable and outputs accurate results.



2. Recurrent Neural Networks: The standard multilayer perceptron approach would not work well in sequence data. Let's imagine, that we want to summarize the text.

• Firstly, the inputs and outputs will not have the same length, as the article and summary lengths can vary greatly.

• Secondly, the studied functions will not be distributed at different positions of the text, which will lead to a decrease in performance.

This is why Recurrent Neural Networks (RNNs) are very popular for serial data such as sentences or music.

Bidirectional RNNs

In certain types of tasks, it is essential to make predictions taking into account both past and future contexts. Those tasks include speech and handwriting recognition. Bidirectional Recursive Neural Networks were invented to deal precisely with this problem. The idea is to feed data to the RNN network simultaneously in a positive and negative time direction.

3. Gated RNNs: Human thoughts have a beneficial property – persistence. It means that we do not start thinking from zero after reading a new article or paragraph, you understand each word based on the previous words no matter how far they were before. However, vanilla RRNs cannot handle it, and when we are interested in understanding long-term dependencies, it becomes unfeasible due to the vanishing gradient problem.

4. Results & Discussion: We evaluate our model for behavior recognition using CSI by using a publicly available dataset, containing 6 labeled activities, with 120 instances of each. The dataset was built using 6 subjects, performing each activity within a period of 20 seconds. The CSI data was collected using a commercial Intel IWC-5300 WiFi receiver, placed 3 meters away from the transmitter, within line of sight. The sampling rate of the data was 1kHz.

Using the LSTM RNN architecture described in as our baseline system, we first try to implement this system. The LSTM network consists of 200 units and tanh activation. We train the system on raw CSI values. After 10-fold cross-validation, the baseline classifier attained a baseline accuracy of 75%.

5. Evaluation of Preprocessing Techniques: Firstly, we evaluate our de-noised data effectiveness in model training. The model trained with normalized and de-noised CSI data lead to improvement in time and accuracy as compared to the model trained with raw CSI.

Secondly, we evaluate the introduction of 1-dimensional (1D) CNN feature extractor into the baseline model. Before passing the denoised CSI to the LSTM layer, features are extracted from the CNN layer. This model achieved an accuracy of 86% on the benchmark data-set. The significant improvement in classification accuracy strongly indicates that the shift-intolerant features we observed do encode significant information about the behaviour.

6. Evaluation of Multi-Layer RNNs: In order to evaluate our proposed LSTM classifier, we train a 3-layer LSTM on raw CSI data, resulting in an accuracy of 84%. This is a significant improvement from the baseline model. The improvement in performance validates our hypothesis that the representation of hidden states at different temporal time-scales is a better representation of human behaviour.

7. Evaluation of Combined Model: Building on the promising improvement in performance from pre-processing techniques, CNN feature extraction and stacked RNN architectures, we train a model combining all of these methods. This model outperforms the baseline classifier significantly, and after 10-fold cross validation, yields an accuracy of 95%, as shown in Table 1.

Table 1. Table of Results for CSI Classification Models: Baseline (2017), De-noised, CNN, Stacked-LSTM, and Combined Model

Model	Accuracy
Baseline LSTM	75%
De-noised LSTM	86%
CNN-LSTM	84%
Our System	**95%**

The confusion matrix of the combined model trained on 70% of the data-set, and evaluated on the remaining 30%, is presented in Figure 5. The model is able to classify Lie Down and Fall perfectly, but does not perform well when classifying Walk, Sit Down or Stand Up. The combined model is our best performing model with an accuracy of 96.7%.

CONCLUSION AND FUTURE WORKS

Evidence from our experimental evaluation suggests that the combination of de-noising, Convolutional Neural Networks, and deep stacked Recurrent Neural Networks performs much better than traditional approaches.

De-noising techniques, such as those used in the CARM model, utilize Principal Component Analysis to eliminate noise in the CSI data. Our approach performed PCA on data collected during periods of inactivity, specifically to combat noise due to the static multi-path effect caused by the signal reflecting off objects in the environment. The use of this de-noising technique has been shown to result in an increase in the convergence rate of our model and results in a slight improvement in classification accuracy when compared to a model trained on raw Channel State Information.

The use of Convolutional Neural Networks as signal feature extractors has also been shown to perform better than using the RNN layers as the sole feature extraction mechanism. After analyzing heat-map visualizations of CSI data, we hypothesized that there existed a set of shift-invariant features responsible for encoding the movement of different body parts. Our experimental analysis showed that there was a significant improvement in classification accuracy when using CNN feature extractors. This improvement came at the cost of increased training time, as the CNN model had a slower convergence than the baseline classifier.

We also see that the use of stacked multi-layer RNNs outperforms current single-layer RNN solutions. We believe that this is because stacking RNN layers allows the model to learn hidden-state representation at multiple time scales, providing a better representation of human behavior than a single time-scale model. The combined model that utilizes de-noising, convolutional feature extraction, and stacked RNNs, outperformed all of the individual components of the system, suggesting that their individual contributions are somewhat additive.

From the confusion matrix of the combined model illustrated in Figure 5, we see that the model often mistakes sitting down for lying down, and vice versa. The activities are similar in terms of the motion of body parts, albeit in reverse order. For future research, the use of bi-directional LSTMs might perform better as they might be capable of discerning between similar activities like sitting and standing. We have seen that models are capable of performing well when evaluated in the same room they are trained in, but they do not tend to generalize well to new environments. Further improving de-noising techniques for CSI data may not only improve the classification accuracy of a model, but also the ability for a model to be generalized to a new environment.

REFERENCES:

[1] Claudine Badue et al. Self-Driving Cars: A Survey. 2019. arXiv: 1901 . 04407 [cs.RO].

[2] R. W. Wolcott and R. M. Eustice. “Visual localization within LIDAR maps for automated urban driving”. In: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2014, pp. 176–183.

[3] Tobias Grosse-Puppendahl et al. “Finding Common Ground: A Survey of Capacitive Sensing in Human-Computer Interaction”. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI ’17. Denver, Colorado, USA: Association for Computing Machinery, 2017, 3293–3315. ISBN: 9781450346559. DOI: 10 . 1145 / 3025453 . 3025808. URL: https : / / doi . org / 10.1145/3025453.3025808.

[4] A. Tahat et al. “A Look at the Recent Wireless Positioning Techniques With a Focus on Algorithms for Moving Receivers”. In: IEEE Access 4 (2016), pp. 6652– 6680.

[5] J. Wilson and N. Patwari. “Radio Tomographic Imaging with Wireless Networks”. In: IEEE Transactions on Mobile Computing 9.5 (2010), pp. 621–632.

[6] Fadel Adib et al. “3D Tracking via Body Radio Reflections”. In: Usenix NSDI (Dec. 2013).

[7] Yongsen Ma, Gang Zhou, and Shuangquan Wang. “WiFi Sensing with Channel State Information: A Survey”. In: ACM Comput. Surv. 52.3 (June 2019). ISSN: 0360-0300. DOI: 10.1145/3310194. URL: https://doi.org/10.1145/3310194.

[8] Charmi Jobanputra, Jatna Bavishi, and Nishant Doshi. “Human Activity Recognition: A Survey”. In: Procedia Computer Science 155 (2019). The 16th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2019),The 14th International Conference on Future Networks and Communications (FNC-2019),The 9th International Conference on Sustainable Energy Information Technology, pp. 698 –703. ISSN: 1877-0509. DOI: https : / / doi . org/10.1016/j.procs.2019.08.100. URL: http://www.sciencedirect.com/ science/article/pii/S1877050919310166.

[9] Zawar Hussain, Quan Sheng, and Wei Emma Zhang. "Different Approaches for Human Activity Recognition: A Survey". In: (June 2019).

[10] Liming Chen et al. "Sensor-Based Activity Recognition". In: IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 42 (Nov. 2012), pp. 790–808. DOI: 10.1109/TSMCC.2012.2198883.

[11] Y. Chen and Y. Xue. "A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer". In: 2015 IEEE International Conference on Systems, Man, and Cybernetics. 2015, pp. 1488–1492

[12] Jian Bo Yang et al. "Deep Convolutional Neural Networks on Multichannel Time Series for Human Activity Recognition". In: Proceedings of the 24th International Conference on Artificial Intelligence. IJCAI'15. Buenos Aires, Argentina: AAAI Press, 2015, 3995–4001. ISBN: 9781577357384.

[13] Jindong Wang et al. "Deep learning for sensor-based activity recognition: A survey". In: Pattern Recognition Letters 119 (2019). Deep Learning for Pattern Recognition, pp. 3 –11. ISSN: 0167-8655. DOI: https : / / doi . org / 10 . 1016 / j.patrec.2018.02.010. URL: http://www.sciencedirect.com/science/article/pii/S016786551830045X.

[14] Michalis Vrigkas, Christophoros Nikou, and Ioannis Kakadiaris. "A Review of Human Activity Recognition Methods". In: Frontiers in Robotics and Artificial Intelligence 2 (Nov. 2015). DOI: 10.3389/frobt.2015.00028.

[15] Zheng Yang Yunhao Liu Kyle Jamieson Kun Qian Chenshu Wu. "Widar: DecimeterLevel Passive Tracking via Velocity Monitoring with Commodity Wi-Fi". In: Mobihoc '17: Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing (July 2017). URL: http://dx.doi.org/10.1145/ 3084041.3084067.

[16] Qin Lv Jie Xiong Shengjie Li Yue Zhang Hong Mei Xiang Li Daqing Zhang. "IndoTrack: Device-Free Indoor Human Tracking with Commodity Wi-Fi". In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (Sept. 2017). URL: https://dl.acm.org/doi/10.1145/3130940.

[17] Deepak Vasisht, Swarun Kumar, and Dina Katabi. "Decimeter-Level Localization with a Single WiFi Access Point". In: 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16). Santa Clara, CA: USENIX Association, Mar. 2016, pp. 165–178. ISBN: 978-1-931971-29-4. URL: https://www.usenix.org/conference/nsdi16/technical-sessions/presentation/vasisht.

[18] Zheng Yang Kun Qian Chenshu Wu and Yunhao Liu. In: ACM Transactions on Embedded Computing Systems (TECS) (Jan. 2018). URL: https://dl.acm.org/ doi/10.1145/3157677.

[19] IEEE Paper URL: https://ieeexplore.ieee.org/document/8682821/references

PLAGIARISM REPORT

We Check Plagiarism on Plagiarism checker through google suggested website. We provided Fig or screenshot of plagiarism below.