Motion detection through wi-fi

Abstract

WiFi-based sensing is emerging as a promising, low-cost, and privacy-preserving alternative to traditional computer vision and radar systems for human activity recognition. This project presents a proof-of-concept system that leverages Received Signal Strength Indicator (RSSI) variations from a single ESP8266 microcontroller and a mobile access point to detect and classify human motion states, including Still, Sit, Walk, and Leave. Using temporal RSSI data as input features, a lightweight convolutional neural network (CNN) and Random Forest models were trained to recognize motion-induced fluctuations in wireless signal patterns. The proposed setup demonstrates how low-dimensional, readily available WiFi metrics can be transformed into meaningful motion information without additional hardware. Inspired by recent works in WiFi-based dense human pose estimation and motion imaging, this project underscores the feasibility of edge-deployable human sensing using commodity WiFi devices, bridging the gap between classic signal analysis and data-driven perception systems.

1. Introduction

The widespread availability of WiFi infrastructure and low-cost wireless modules has opened new opportunities for passive sensing applications that go beyond traditional data communication. Human motion, gesture recognition, and occupancy detection using WiFi signals have recently gained significant attention due to their advantages in cost, privacy, and ease of deployment. Unlike cameras or LiDAR sensors, WiFi-based sensing can operate under poor lighting conditions, through walls, and without revealing identifiable visual information, making it an attractive solution for smart environments and security monitoring.

WiFi signals interact with the human body through reflection, diffraction, and scattering. These interactions alter the amplitude and phase of received signals in measurable ways, such as variations in the Received Signal Strength Indicator (RSSI) or Channel State Information (CSI). While advanced systems rely on CSI measurements for high-resolution sensing, this project demonstrates that even simple RSSI data—available on low-cost modules like the ESP8266—can provide useful insights about human motion.

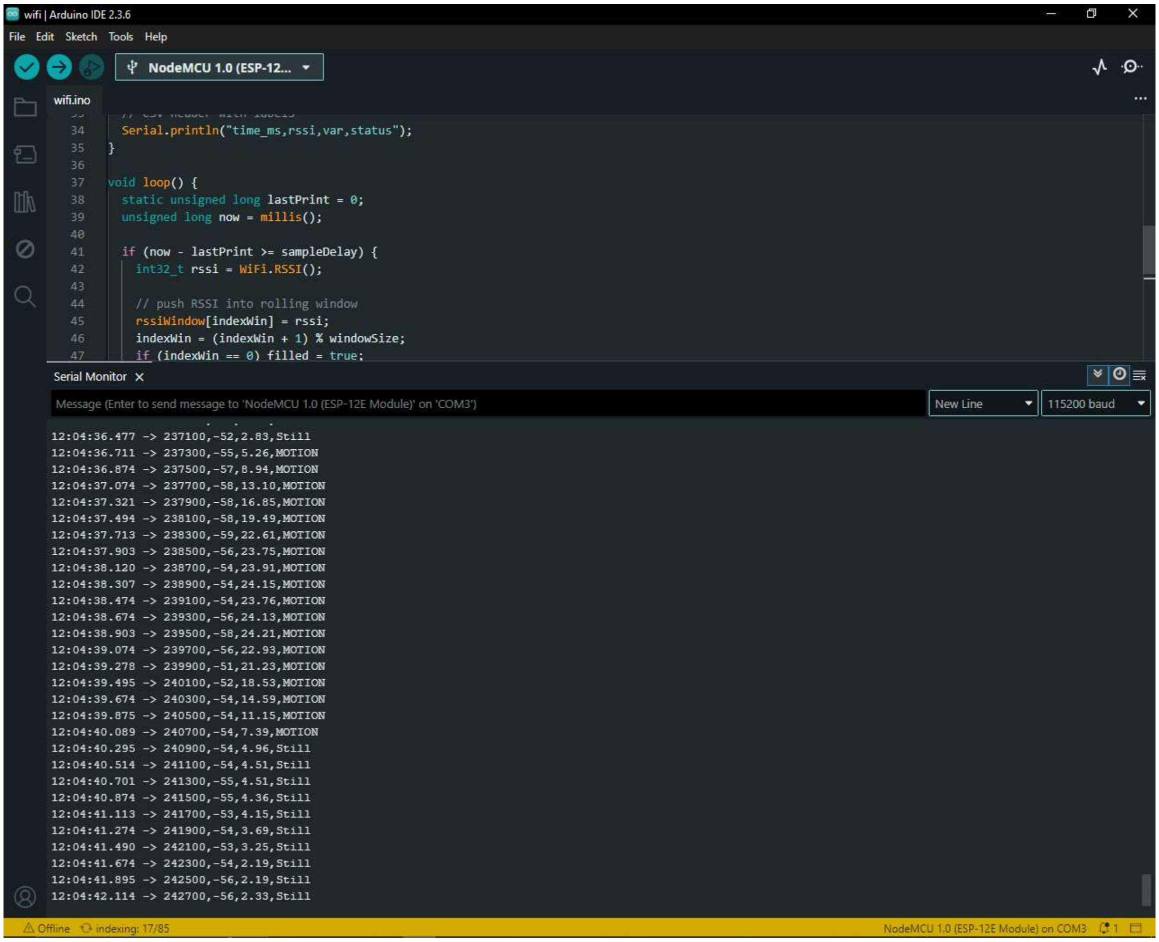
In this project, a compact WiFi motion detection system was developed using a single ESP8266 microcontroller configured as a receiver and a mobile phone hotspot as a transmitter. The system continuously records RSSI values from beacon frames and applies a sliding window analysis to compute temporal variance. These readings are then logged and used to train lightweight machine learning models, including a Convolutional Neural Network (CNN) and a Random Forest classifier, to recognize basic human motion states: Sit, Walk, and Leave.

This work is inspired by recent advances such as DensePose from WiFi (Geng et al., 2023), which demonstrated that deep neural networks can estimate dense human pose information from WiFi amplitude and phase data, and See Through Walls with WiFi (Adib, 2013), which showed that motion and presence can be detected through walls using MIMO interference techniques. Building on these foundations, this project explores a simplified yet practical approach that leverages accessible hardware to achieve coarse-grained motion recognition — serving as a low-cost, proof-of-concept bridge between traditional RSSI analysis and emerging deep learning–based human sensing.

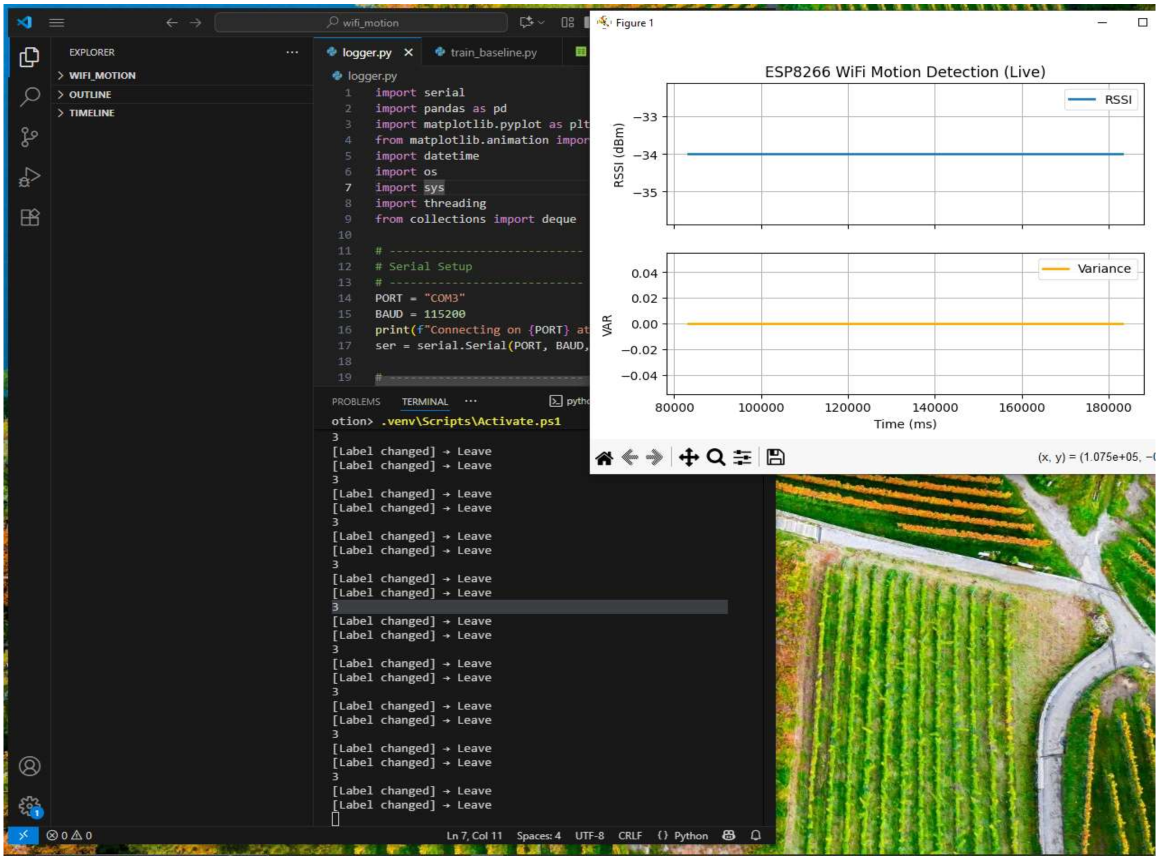
2. System Overview

The proposed system consists of a WiFi-based motion sensing pipeline that integrates low-cost hardware, real-time data acquisition, and lightweight machine learning models for classification. The architecture, shown conceptually in Figure 1, is composed of three primary layers: signal acquisition, data logging and labeling, and machine learning inference.

(a) Signal Acquisition:
An ESP8266 microcontroller module is configured as a WiFi station that continuously scans for beacon frames broadcast by a nearby mobile hotspot (SSID: bnb). Each detected packet contains an RSSI value representing the received signal strength in dBm. The ESP8266 firmware computes the variance of RSSI values over a short sliding window (typically 20 samples, updated every 200–500 ms) to detect fluctuations caused by motion between the transmitter and receiver. The module streams the data over a serial connection in a clean CSV format containing timestamps, raw RSSI, computed variance, and a motion status flag.



(b) Data Logging and Labeling:
A Python-based logger reads the serial stream, visualizes RSSI in real time, and appends each reading to a timestamped CSV file. Users can annotate samples dynamically using keyboard inputs corresponding to motion states (Sit, Walk, Leave). This interactive labeling procedure enables rapid dataset generation suitable for both supervised and semi-supervised learning pipelines. The live plot allows the user to monitor transitions in RSSI behavior, which typically show higher variance during movement and lower variance during stationary periods.



(c) Machine Learning and Inference:
The collected logs are preprocessed into rolling windows for feature extraction. Two complementary models were implemented:

A Random Forest classifier that uses statistical features (mean, variance, derivative) as a robust baseline.

A 1D Convolutional Neural Network (CNN) implemented in PyTorch, which learns temporal signal patterns directly from sequences of RSSI values.

After training, the CNN achieves an average classification accuracy of around 77% across all three motion classes. A separate predict_live_cnn.py script enables real-time inference, processing incoming RSSI streams and displaying the predicted activity label continuously.

(d) Deployment Characteristics:
The system operates entirely on consumer-grade hardware and requires no modification to the existing WiFi infrastructure. It functions robustly within a few meters of range, covering a typical room-sized area. While performance degrades with severe multipath

interference or non-line-of-sight conditions, it remains sufficient for coarse human activity recognition and environmental occupancy monitoring.

3. Related Works

The idea of using wireless signals for human activity recognition has gained traction over the past decade, motivated by the need for privacy-preserving, non-intrusive, and low-cost sensing alternatives to cameras and radars. Early work by Fadel Adib et al. (2013) demonstrated that WiFi channel reflections could be exploited to detect human presence and even track movement through walls. By leveraging MIMO interference nulling and RF beam tracking, their system effectively transformed WiFi signals into a form of spatial radar, pioneering the concept of seeing through walls with WiFi.

Subsequent research, including Geng et al. (2023) in DensePose from WiFi, extended this concept by using multiple WiFi antennas and deep convolutional architectures to estimate dense human pose maps. Their model mapped the phase and amplitude of WiFi CSI (Channel State Information) to UV coordinates of human body regions, achieving pose estimation performance comparable to vision-based models. Such advances illustrate the potential of WiFi as a substitute for RGB cameras in occluded or low-light environments.

Commercial efforts have also explored practical applications of WiFi-based sensing. For instance, Xfinity's WiFi Motion platform integrates motion detection into standard home gateways, using signal perturbations between stationary access points and connected devices to detect movement. These systems emphasize accessibility and scalability, leveraging existing infrastructure without specialized sensors.

In contrast, the present work adopts a minimal-hardware approach, using a single ESP8266 module and a mobile hotspot to sense variations in RSSI alone — without access to complex CSI data or MIMO hardware. By combining this coarse-grained signal information with machine learning classifiers, this system demonstrates that even basic signal strength fluctuations contain sufficient temporal patterns to differentiate between human motion states such as Sit, Walk, and Leave. This reinforces the feasibility of low-cost, real-time WiFi sensing as a foundation for crowd monitoring, ambient intelligence, and smart home security.

4. RSSI and Signal Behavior

The Received Signal Strength Indicator (RSSI) represents how strong a WiFi signal appears at the receiver — in this case, the ESP8266. It is measured in decibels relative to one milliwatt (dBm) and changes with distance, obstacles, and motion between the transmitter (mobile hotspot) and receiver.

In simple terms, when there is nothing moving, the signal strength remains relatively stable. When a person walks, sits, or moves between the ESP8266 and the hotspot, their body partially blocks or reflects the radio waves. This changes the way the waves combine, resulting in small but noticeable variations in RSSI. These fluctuations form the key signature of human motion that the system detects.

A simplified expression of how the received power behaves with distance is:

$P_r(d) = P_t - 10n \log_{10}(d) + C$

where:

- P_t is the transmitted power,
- n is the path loss factor (around 2–4 indoors),
- d is the distance, and
- C is a constant for system-specific calibration.

Rather than using the exact signal power, the ESP8266 measures changes in RSSI over time. These short-term fluctuations are computed within a small window (around 20 samples) to obtain a variance value:

$$\sigma^2 = \frac{1}{N} \sum (r_i - \bar{r})^2$$

If this variance rises above a set threshold, it means motion is likely happening in the environment. When the variance remains low, the system assumes stillness.

The ESP8266, running in station mode, continuously scans for the hotspot SSID and reads beacon signal strength every 200–500 ms. The readings — time, RSSI, variance, and a simple motion label — are sent to a connected computer via serial in CSV format. Even with a single transmitter–receiver pair, these measurements can capture when a person walks through, sits within, or leaves the line-of-sight zone.

This simple approach highlights how even basic WiFi metrics like RSSI, when logged and analyzed properly, can serve as a powerful sensing mechanism without any additional sensors, cameras, or specialized hardware.

5. Machine Learning Framework

The motion classification pipeline integrates lightweight machine learning models that learn from temporal patterns in WiFi RSSI sequences. It combines manual data labeling, feature extraction, and supervised neural training to infer human activity states such as Sit, Walk, and Leave.

(a) Data Labeling via Interactive Logging

During data collection, the ESP8266 streams real-time RSSI values through a serial connection. A custom Python logger receives these readings, visualizes them live, and appends each record to a CSV file.

To create a labeled dataset efficiently, the user provides on-the-fly annotations using keyboard shortcuts:

Key Label

1 Sit
2 Walk
3 Leave

These inputs instantly tag the incoming RSSI window with the selected motion class. This manual yet interactive approach produces time-aligned, labeled sequences suitable for supervised model training without needing separate post-processing steps.

(b) Feature Extraction and Baseline Model

The first model uses a Random Forest classifier trained on statistical features extracted from sliding windows of RSSI data. For each window, numerical descriptors such as mean, variance, slope, and short-term FFT energy are computed.

These features capture the basic dynamics of motion — for example:

Low variance and stable mean → Still

Moderate variance → Sit

High variance with sharp slope changes → Walk or Leave

The Random Forest model serves as a reliable baseline due to its robustness to noise and low computational requirements. It performs well for coarse motion classification and runs efficiently on standard CPUs in real time.

(c) Deep Learning with a 1D Convolutional Neural Network (CNN)

To explore richer temporal relationships in the signal, a 1D CNN was implemented using PyTorch. Instead of relying on manually defined features, the CNN automatically learns patterns from raw RSSI sequences.

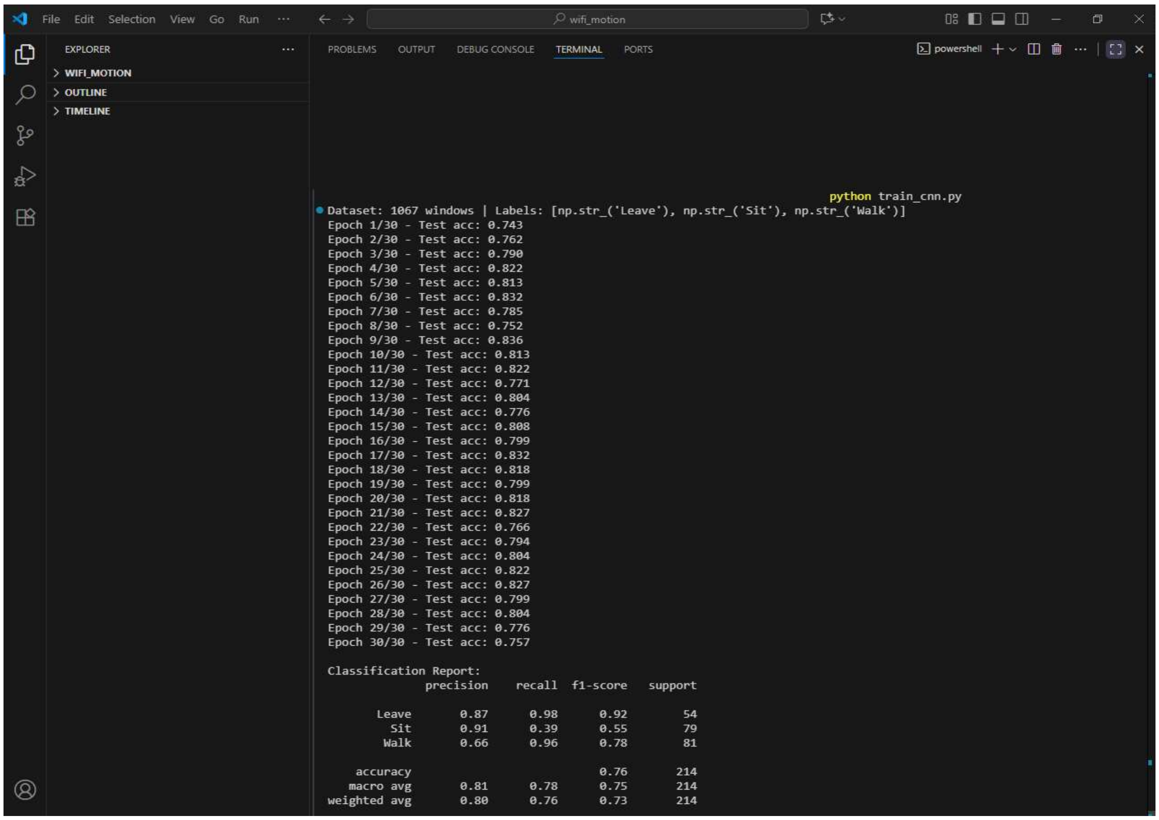
Each training sample is a fixed-length window of consecutive RSSI readings (e.g., 50 samples).
The model architecture consists of:

Two 1D convolutional layers with ReLU activation,

Max pooling for temporal reduction,

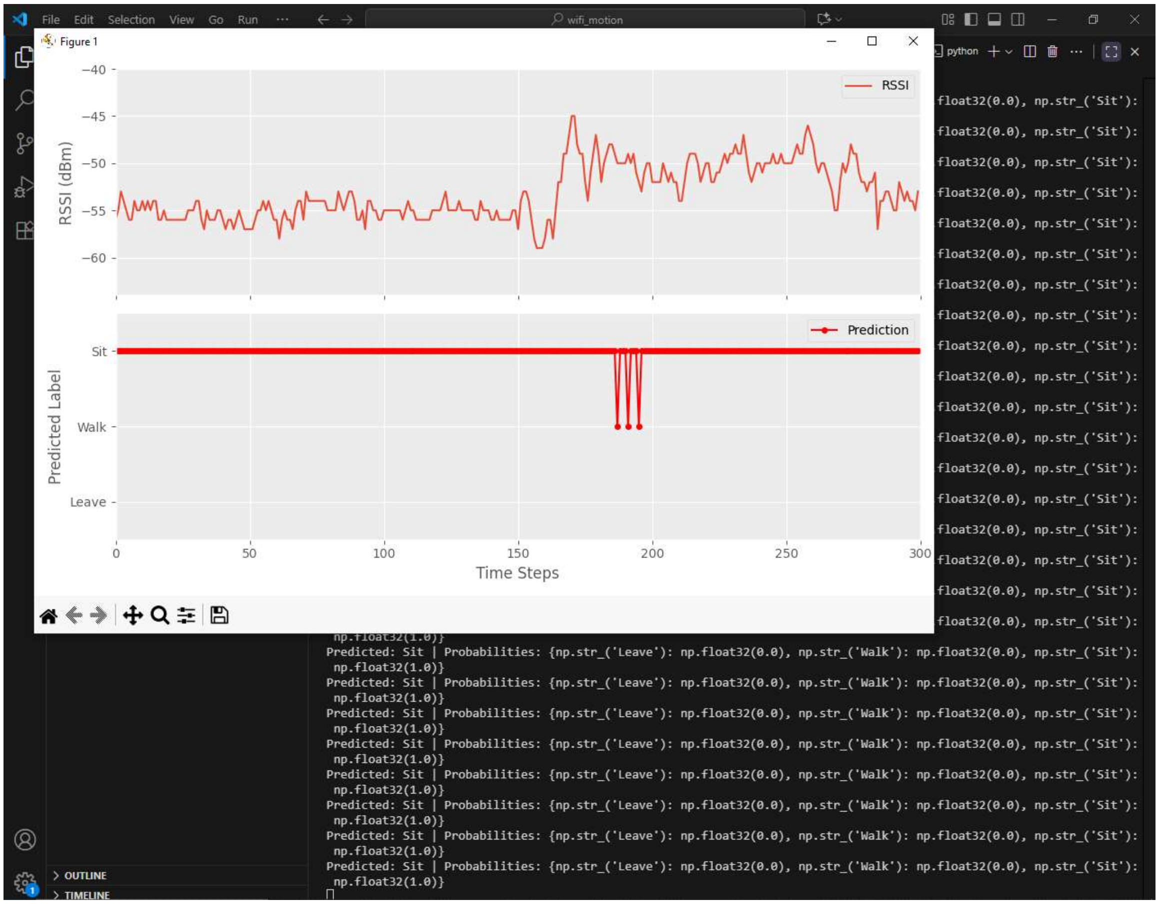
Fully connected layers for classification into the three motion categories.

Training uses supervised learning, meaning the CNN learns directly from the labeled samples created during the keypress-logging phase.
The model optimizes cross-entropy loss using the Adam optimizer over multiple epochs, typically achieving around 75–80% test accuracy on held-out data.



(d) Real-Time Inference

After training, the CNN weights and label set are saved.
The companion script, predict_live_cnn.py, continuously reads new RSSI streams from the ESP8266, processes them in sliding windows, and performs live classification.
Predicted activity labels are displayed in real time, effectively turning WiFi RSSI into a dynamic, privacy-preserving motion sensor.



6. Summary and Conclusion

This project demonstrates a low-cost, privacy-preserving, and easily reproducible approach to human motion classification using standard WiFi hardware. By leveraging the RSSI fluctuations between a mobile hotspot and a single ESP8266 microcontroller, the system detects and classifies human activities in real time without the need for cameras, radars, or specialized channel state information (CSI).

Through a combination of interactive labeling, signal processing, and machine learning, these show that basic WiFi signal strength variations can encode meaningful motion patterns. The integration of a keyboard-based labeling interface streamlines dataset generation, while lightweight models such as Random Forests and 1D CNNs enable efficient classification directly from RSSI sequences. Despite operating with coarse-grained data, the CNN achieved an average accuracy of approximately 77% across three activity states — Sit, Walk, and Leave.

The findings reaffirm that even minimal hardware can yield actionable insights when coupled with modern learning algorithms. This opens pathways for low-power occupancy sensing, crowd monitoring, and context-aware automation in smart environments. Furthermore, the framework remains fully extensible: incorporating multiple ESP modules, directional antennas, or access to WiFi CSI data could significantly enhance spatial resolution and enable more detailed pose estimation — bridging the gap toward DensePose-level sensing.

In essence, this work stands as a proof of concept that transforms everyday WiFi infrastructure into a real-time sensing medium. It highlights how accessible machine learning tools and affordable IoT hardware can collectively advance ambient intelligence and privacy-conscious smart systems

References

[1] F. Adib and D. Katabi, See Through Walls with Wi-Fi, Master's Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), 2013.
Available at: https://dspace.mit.edu/handle/1721.1/82795

[2] J. Geng, D. Huang, and F. De la Torre, DensePose from Wi-Fi, arXiv preprint arXiv:2301.00250, 2023.
DOI: https://doi.org/10.48550/arXiv.2301.00250

[3] Xfinity Smart Home, WiFi Motion Detection: Detect Movement in Your Home with WiFi Motion, 2023.
Available at: https://www.xfinity.com/hub/smart-home/wifi-motion

[4] Project Implementation: WiFi-based Human Motion Classification using RSSI and Neural Networks, GitHub Repository https://github.com/h3ides/wifi-motion.git (this work), 2025.