Human Activity Recognition using ESP32 and Channel State Information (CSI)

Author : Divyank Kushwaha

1. Introduction

- Human Activity Recognition (HAR) refers to the method of detecting and recognizing human movements or activities using Channel State Information data.
- In this project, we use **Wi-Fi Channel State Information (CSI)** for Human activity recognition. Unlike old camera-based systems (which causes privacy concerns) or wearable devices (which require the user to carry hardware all the time), CSI-based HAR works **contactlessly** by analyzing changes in wireless signal properties caused by human movement.
- CSI provides fine-grained measurements of the wireless channel, including amplitude and phase for multiple subcarriers. When a person moves, multipath propagation changes, which alters these values. These are the variations which can be recorded, processed, and can be classified into different activities.

- **Applications**:
  - ❖ Smart homes (automated lighting, gesture control)
  - ❖ Healthcare (elderly monitoring, fall detection)
  - ❖ Security (intrusion detection)
  - ❖ Workplace monitoring (employee activity analysis)

2. Methodology

- **Data Collection**
  - ❖ **Receiver**: ESP32-WROOM-32 board connected to Wi-Fi network (2.4GHz).
  - ❖ **Transmitter**: Router sending packets (2.4GHz).
  - ❖ CSI data captured via **UDP** on a laptop.

- **Data Pre-processing**
  - ❖ **CSI values are complex numbers:** $H = Real + j\,Imag$
  - ❖ **Convert CSI complex values to amplitude and Phase:** $Amplitude = \sqrt{Real^2 + Imag^2}$

$$Phase = \tan^{-1}\left(\frac{Imag}{Real}\right)$$

- ❖ Noise Removal & Signal smoothing using **Hampel & Savitzky-Golay filter** respectively.

- **Feature Extraction**
    - ❖ Extract the **Amplitude** and **Phase** data from the raw CSI data received at ESP32 receiver.

- **Model Training**
    - ❖ **Algorithm:** Random Forest Classification.
    - ❖ **Parameters:**
        - ➢ `n_estimator = 200`
        - ➢ `criterion = gini`
        - ➢ `max_depth = None`
        - ➢ `min_samples_split = 11`
        - ➢ `class_weight = balanced_subsamples`
        - ➢ `min_samples_leaf = 11`
        - ➢ `random_state = 42`
        - ➢ `max_features = log2`
    - ❖ **Dataset split:** 80% training data, 20% testing data

- **Evaluation Metrics**
    - ❖ Accuracy
    - ❖ Precision
    - ❖ Recall
    - ❖ F1-score
    - ❖ Confusion Matrix

3. Experimental Setup

- **Hardware**:
    - ❖ ESP32-WROOM-32 (Receiver – 2.4GHz)
    - ❖ Router (Transmitter – 2.4GHz)
    - ❖ Wi-Fi Router (SSID: “steeleye”, Password: “12345678”)
    - ❖ Laptop for data storage and processing.

- **Software**:
  - ESP-IDF framework for ESP32 firmware.
  - **Python libraries:** Pandas, NumPy, Scikit-learn, Matplotlib

- **Environment**:
  - **Room Length :** ~ 4.5m and **Room Width :** ~ 2.8m
  - **Distance between transmitter and receiver:** ~ 3m
  - Minimal interference from other objects.

- **Recorded Activities**:
  - Empty room
  - Sitting between transmitter and receiver
  - Sleeping between transmitter and receiver
  - Standing between transmitter and receiver

4. Results

- **Performance:**
  - Training Accuracy: 98.69%
  - Testing Accuracy: 97.12%
  - Mean Cross Validation Accuracy: 95.22%
  - Training Loss: 15.04%
  - Testing Loss: 18.56%

- **Observations:**
  - All activities achieved > 98% precision.
  - Learning curves indicate stable performance with very less overfitting.
  - Machine Learning model predicts the activities correctly, until there is any sudden change in activity detected.
  - Live Prediction is shown on terminal in every 0.5 second, which is the median of all the predicted activities in the time period of 0.5 second.

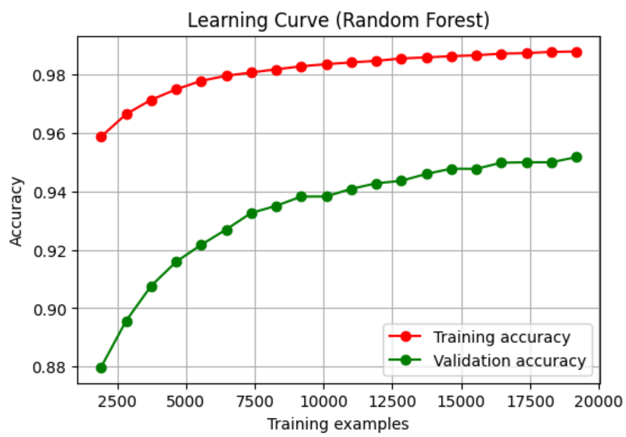

Fig-1: Accuracy Learning Curve

Fig-2: Loss Learning Curve

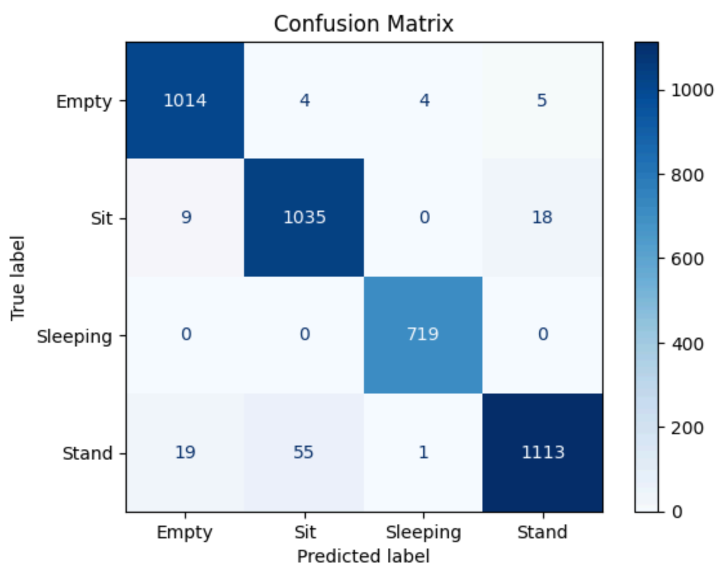

Fig-3: Confusion Matrix

5. Conclusion & Future Scope

- **Conclusion**:
  - ❖ The project describes that CSI-based HAR is an effective, privacy-friendly, and low-cost method for classifying human activities inside a room. This achieves higher accuracy while avoiding the limitations of camera or wearable-based systems.

- **Future Scope**:
  - ❖ Classification to more complex activities (e.g.- exercising, dancing).
  - ❖ Implement real-time processing by integrating cloud features.
  - ❖ Explore deep learning models (CNN, LSTM) for improved performance.
  - ❖ System for outdoor environments to recognize activities.