WI-FI-BASED HUMAN ACTIVITY RECOGNITION USING CSI AND MACHINE LEARNING TECHNIQUES

Sangeeta Rai,
Center of Energy and Environment, Dr B R Ambedkar National Institute of Technology Jalandhar-144008, Punjab, India
Email Id: rai.sangeetarai16@gmail.com
Naveen Kumar Gupta
Department of Information Technology, Dr B R Ambedkar National Institute of Technology Jalandhar-144008, Punjab, India
Email Id: guptank@nitj.ac.in
DOI: https://doi.org/10.64558/bllp/198089vqcvah

1. INTRODUCTION

Human Activity Recognition has become an essential technology for applications in smart homes, healthcare monitoring, and human computer interaction. Traditional approaches rely on vision-based systems or wearable sensors, which often raise privacy concerns and may require user compliance [1]. To overcome these limitations, recent research has turned towards Wi-Fi-based sensing, which leverages the widely available wireless infrastructure to achieve contactless and device-free recognition of human activities [2].

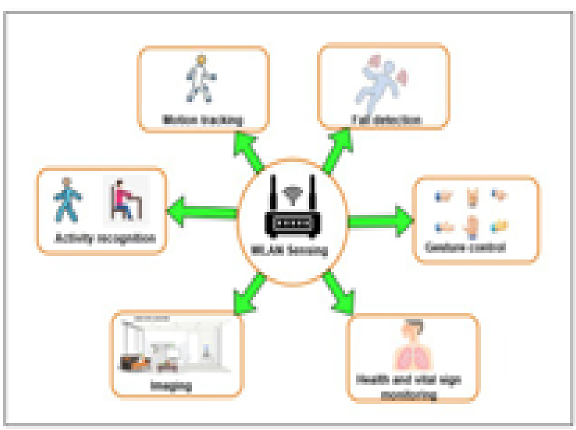

Fig. 1: Basic working representations of a WiFi-Based Human Sensing System

Fig. 1 A key enabler of Wi-Fi based HAR is Channel State Information (CSI), which captures fine grained signal variations induced by human motion in the wireless environment [3]. By analyzing these variations, it is possible to detect and classify activities such as walking, sitting, falling, and even respiration monitoring [4]. For example, Zhang et al. demonstrated the use of commodity Wi-Fi devices for non-intrusive respiration detection, showing the sensitivity of CSI to subtle body movements [4]. Similarly, Wu et al. showed that Wi-Fi can be used to detect both moving and stationary humans in indoor settings [2]. The integration of machine learning techniques has further enhanced HAR performance. Statistical modeling approaches, such as Wi Detect, provided robust motion detection under varying environments [5], while more recent advances have explored deep learning frameworks to extract spatial-temporal features directly from CSI streams [6]. Yang and Jiang proposed selective sharing modules for multitask sensing, demonstrating the adaptability of deep models for complex recognition scenarios [7]. Such progress highlights the growing synergy between wireless signal processing and artificial intelligence. In the study, we train and test the classical machine learning models such as SVM, Random Forest, and KNN and test them against the deep learning ones, a shallow neural network and a Bidirectional Long Short-Term Memory (Bi-LSTM) network. The primary contributions is a CSI-inspired modeling approach applied to the UCI HAR dataset, detailed comparison of classical and deep learning models for activity recognition and an experimental evaluation demonstrating that Bi-LSTM achieves the highest recognition accuracy.

The remainder of this paper includes section 2, which includes a brief study about Wi-Fi signal-based sensing and its applications and future. The next section is section 3, which includes the description of the dataset used and details about CSI signals. In section 4 we have the details about the methodology we have used for analysis and implementation of algorithms. The next section is section 5 with the more details of implementation and results we got after analysis. The last section 6 we have concluded this paper with discussion on future works.

2. RELATED WORK

The rapid growth of WiFi sensing has enabled non-intrusive HAR, leveraging the ubiquity of wireless networks and the fine-grained information available in CSI. Unlike early approaches relying on Received Signal Strength (RSS) [3], CSI captures multipath propagation details, enabling more accurate recognition of diverse human activities, including both subtle respiration signals [4] and gross body movements [2]. These developments have established WiFi sensing as a promising alternative to vision-based or wearable sensing approaches, particularly in healthcare and smart home contexts. Several surveys and overview studies have mapped the progress of WiFi-based HAR. Hernandez and Bulut [8] discussed signal processing techniques and challenges in real-world edge deployments, while Ahmad et al. [6] reviewed deep learning–driven advances in WiFi sensing, highlighting unresolved challenges such as environment dependence and domain adaptation. Similarly, Ahmed [9] provided a comprehensive survey of device-free gesture recognition using CSI, underscoring its potential as a versatile, low-cost sensing modality. Collectively, these works demonstrate both the maturity and the evolving research challenges of WiFi CSI sensing. Researchers have also explored statistical and model-based approaches. WiDetect [5] introduced a statistical electromagnetic model for robust motion detection under environmental variations, while Taylor et al. [1] developed an intelligent real-time HAR system tailored for healthcare. Song et al. [10] further incorporated AI methods into contactless sensing, demonstrating improved robustness in complex settings. These works show that statistical methods remain relevant, though increasingly complemented by AI techniques.

With the rise of machine learning and deep learning, performance improvements have become evident. Yang and Jiang [7] proposed a multitask WiFi sensing framework using selective-sharing deep modules, whereas Zhu et al. [11] developed a deep learning system to distinguish human and non-human motion. Meng et al. [12] introduced Secur-Fi, a secure CSI-based sensing system on commercial WiFi devices, emphasizing privacy and robustness. Abuhoureyah et al. [13] extended the field with through-wall HAR via deep learning, demonstrating feasibility even in non-line-of-sight scenarios.

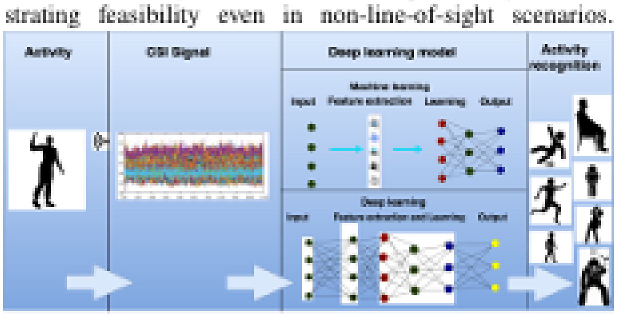

Fig. 2: Conceptual Representation of CSI-Based Activity Recognition

3. DATASET AND CSI SIGNAL OVERVIEW

The primary focus of this work is on WiFi signal–based human activity recognition, where Channel State Information provides a structured collection of wireless signals captured during various human activities [14]. CSI is obtained by leveraging the physical layer information from Wi-Fi signals [8], reflecting how wireless channels are influenced by environmental dynamics such as human motion [3], body orientation, and surrounding obstacles [2]. In this context, Wi-Fi devices operate as transmitters and receivers, while the wireless channel itself becomes a medium for sensing human activities [15]. CSI measurements, represented as complex values containing both amplitude and phase information across multiple subcarriers, provide fine-grained insight into signal variations caused by different activities [8]. However, due to the limited availability of publicly accessible CSI-based HAR datasets, this study employs the UCI HAR dataset as a benchmark for model training and evaluation. Although the UCI HAR dataset is based on wearable sensor data rather than WiFi CSI, it offers a standardized and well-annotated resource, making it an effective proxy for validating proposed approaches while keeping the broader goal of WiFi CSI–based HAR in perspective.

3.1. UCI HAR Dataset

The dataset includes recordings of common activities such as walking, sitting, standing, falling, and other movements, each of which produces distinct variations in the recorded signal patterns [14]. These variations are reflected in the sensor measurements over different axes, making it possible to differentiate between activities through signal analysis [5]. Each data trace is captured over time, resulting in a time-series dataset with multiple dimensions corresponding to sensor axes and placements [8]. Such multi-dimensionality enhances the discriminative power of the dataset, as certain activities influence specific signals more strongly than others [3]. An important characteristic of this dataset is its focus on real-world conditions, where the measurements are subject to variability, interference, and noise. These conditions make the data both challenging and realistic for machine learning applications in HAR [6]. To enable systematic experimentation, the dataset is pre-structured with labeled activity classes, allowing researchers to train and evaluate classification models effectively. The repository also provides preprocessing utilities to extract features from the raw signals, normalize the data, and prepare it for machine learning pipelines [14].

3.2. Reinterpreting UCI HAR as CSI Data

CSI data offers a rich, high-dimensional representation akin to accelerometer data in UCI HAR, but without requiring wearable sensors [1]. The overall pipeline mirrors the UCI HAR structure: activities are labeled, CSI data epochs are captured, preprocessed, and then fed into a classification algorithm specifically, an LSTM-based deep learning model [6]. This similarity suggests that UCI HAR classification tasks can, in principle, be reinterpreted within a CSI framework, using Wi-Fi signal variations as the functional equivalent of acceleration signals [9]. The repository includes tools for CSI data extraction (sendData/recvCSI), visualization, labeling, and model training, making it possible to replicate or extend CSI experiments similarly to how UCI HAR datasets are used for benchmarking [14]. Additionally, CSI-based frameworks such as Wi-ESP demonstrate the potential of standardized tools for device-free Wi-Fi sensing and reproducible experimentation [16].

4. METHODOLOGY

The proposed methodology leverages motion sensor data to perform HAR, drawing inspiration from the framework described below. The workflow begins with sensor data collection, where accelerometers and gyroscopes embedded in smartphones capture three-axis acceleration and angular velocity signals during different activities. Each sample represents motion dynamics across multiple axes, forming a high-dimensional time-series representation of human movement [14]. Since raw sensor measurements are often noisy and irregular, preprocessing is applied to ensure robustness. This involves steps such as denoising, signal calibration, and normalization of values to minimize device-specific distortions [5].

Following preprocessing, feature extraction is carried out to capture discriminative characteristics of human activities. Both time-domain features (e.g., mean, variance, energy) and frequency-domain features (e.g., spectral entropy, dominant frequencies) can be derived from the sensor signals. In some cases, advanced transformations such as Short-Time Fourier Transform (STFT) or Wavelet Decomposition are applied to preserve temporal and spectral information. The resulting feature vectors provide a compact and informative representation of activity patterns [8].

These feature vectors are then used for model training, where supervised learning algorithms such as Support Vector Machines (SVM), Random Forests, or deep neural networks are applied to learn discriminative mappings between extracted features and activity labels. The dataset is typically split into training and testing sets, and model parameters are optimized using validation techniques. Once trained, the model can classify unseen sensor signals into activity classes with high accuracy, completing the HAR pipeline [3].

4.1. Data Preprocessing

The data processing pipeline involves extracting raw signal measurements, removing noise, computing statistical and frequency-domain features, segmenting the data streams into overlapping windows, applying appropriate transformations, and normalizing the features to prepare them for classification [14]. To ensure consistent feature distributions, we perform the following preprocessing steps: Normalization: Z-score standardization is applied to each feature. Label Encoding: Activity labels are numerically encoded. Sequence Padding: For sequence-based models such as LSTM, sequences are padded to ensure uniform length. Train-Test Split: 80% of the data is used for training and 20% for testing [5].

4.2. Feature Representation

Temporal dynamics are represented in the time-frequency domain using transformations such as STFT and complemented with statistical descriptors (mean, variance, skewness, and higher-order moments) to capture distributional characteristics of the signals. The combination of temporal, spectral, and statistical features produces a discriminative representation that preserves activity-specific signal modulations for effective classification. The original data provide a set of 561 features, including time-domain descriptors (mean, standard deviation, energy) and frequency-domain descriptors (FFT coefficients, spectral entropy). To incorporate temporal dependencies, feature vectors are reshaped into a sequence of shape (time, steps, features), particularly when using sequence-based deep learning models such as Bi-LSTM [11].

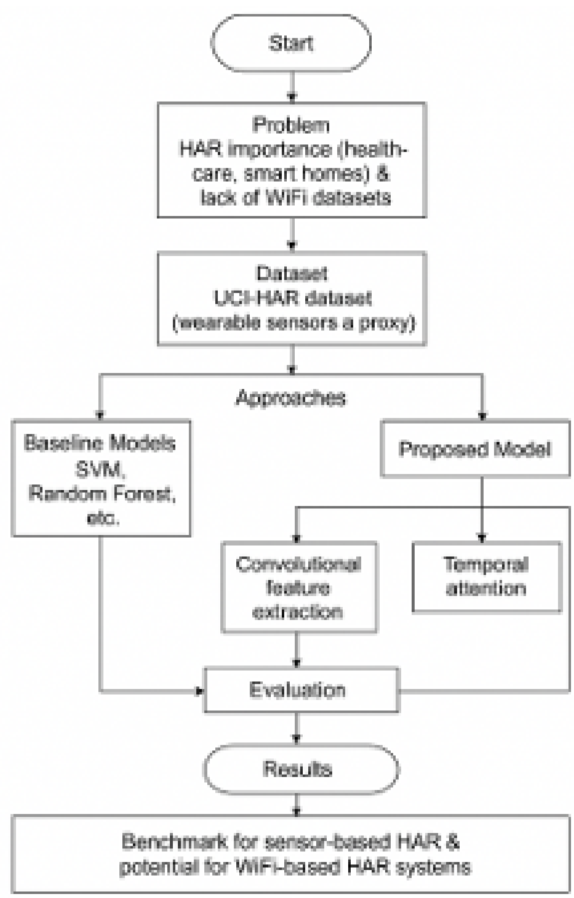

Fig. 3: Process Flow Diagram of Channel State Information-Based Human Sensing

4.3. Classification Models

Classification models are supervised learning algorithms that categorize input data into predefined classes based on patterns learned from labeled datasets [7]. They can handle tasks like binary classification (e.g., spam vs. not spam), multi-class classification (e.g., predicting fruit type), etc. Common models that we are using in our training dataset are Support Vector Machine, Random Forest, K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP) Neural Network, Bi-LSTM. **1) Support Vector Machine (SVM):** SVM is a powerful supervised learning approach that constructs an optimal separating hyperplane in a high-dimensional space, with kernel functions enabling efficient handling of non-linear class boundaries and high-dimensional feature mappings [12].

2) Random Forest: Random Forest operates as an ensemble of decision trees, employing bootstrap aggregation and random feature selection to enhance generalization, mitigate overfitting and achieve strong predictive performance on diverse datasets [13].

3) K-Nearest Neighbors (KNN): The model is simple and interpretable but sensitive to noisy features. KNN a non-parametric and instance-based learning algorithm, assigns class labels by evaluating the majority voting scheme within a specified neighborhood, thereby leveraging local proximity structures in feature space [17].

4) Multilayer Perceptron (MLP) Neural Network: It uses a MLP neural network that comprises several fully connected zones of sequentially shrinking sizes. The architecture consists of three hidden layers comprising 256, 128 and 64 neurons respectively, after which these layers are accompanied by ReLU activation functions. With its Adam optimizer and categorical cross-entropy loss, the training is done on the network. To increase generalization and avoid overfitting, Batch normalization and dropout regularization is used [4].

5) Bi-LSTM: Bidirectional Long Short-Term Memory networks, as an advanced recurrent architecture, incorporate dual directional processing of sequential data, thereby modeling long-range dependencies and contextual cues more effectively than conventional unidirectional RNNs [10]. The table below represents a comparative analysis of commonly used machine learning and deep learning models for human activity recognition using Wi-Fi CSI signals, highlighting their key characteristics, strengths, limitations etc.: -

4.4. Evaluation Metrics

Evaluation metrics are measures used to assess the performance of a model by comparing its predictions with the actual outcomes. For classification tasks, common metrics include accuracy (the proportion of correct predictions) [18], precision (the proportion of correctly predicted positive instances out of all predicted positives) [1], recall (the proportion of correctly predicted positives out of all actual positives) [19], and the F1-score (harmonic mean of precision and recall, useful for imbalanced datasets) [9].

Table 1
Structural & performance-oriented comparison of ML Models

Model	Model Type	Training Time	Accuracy Rance	CSI Suitability
SVM	Classical ML	Medium	75-85%	Moderate
Random Forest	Ensembled (Tree-Based)	Fast	80-90%	Good

Evaluation using these metrics has been widely applied in wireless sensing and HAR frameworks [20].
As shown in Table 1.

Table 2: Structural & performance-oriented comparison of ML Models

Model	Model Type	Training Time	Accuracy Range	CSI Suitability
SVM	Classical ML	Medium	75-85%	Moderate
Random Forest	Ensemble (Tree-Based)	Fast	80-90%	Good
KNN	Instance based	None (Lazy)	70-85%	Moderate
MLP	Feed forward Neural Network	Medium to High	80-90%	Very Good
Enhanced Bi-LSTM	Deep Recurrent Network	High	85-95%	Excellent

5. RESULTS AND DISCUSSION

After applying the dataset on all the described algorithms we got the results in form of accuracy, precision, f1-score and confusion matrix. Below we have the description about the working and parameters used in implementation for each algorithm:

5.1. Results

1) SVM This classifier achieved an overall accuracy of 95.22%, indicating strong performance in distinguishing between the six human activity classes. The UCI-HAR dataset was employed to evaluate the effectiveness of a SVM classifier in multi-class activity recognition [21]. The raw data was preprocessed by imputing missing values with feature-wise means and applying standard normalization to ensure all features contributed equally during training. The classifier was implemented using the Radial Basis Function (RBF)

kernel, expressed as:

$$K(x_i, x_j) = exp(-\gamma ||x_i - x_j||^2) \tag{1}$$

which enables the model to map input features into a higher-dimensional space and capture non-linear decision boundaries [22]. The penalty parameter was set to C = 1.0 balancing the trade-off between maximizing the margin and minimizing misclassification errors, while the kernel coefficient was set to γ = scale adapting automatically to the feature space dimensionality. The optimization problem solved by the SVM can be formulated as:

$$min \frac{1}{2} ||\omega||^2 + c \sum_{i=0}^{N} \xi_i \tag{2}$$

subject to the soft margin constraints:

$$y_i(\omega^T \phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0 \tag{3}$$

Where $\phi(x_i)$ is the non-linear mapping induced by the kernel and ξ_i are slack variables allowing controlled misclassification. The resulting decision function is given by:

$$f(x) = sign(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b) \tag{4}$$

where α_i are support vector weights and b is the bias term. Model evaluation was conducted using accuracy, precision, F1-score, and confusion matrix analysis, ensuring a comprehensive performance assessment. These results confirm that, with carefully tuned parameters, kernelized SVMs provide robust non-linear decision boundaries well-suited for human activity recognition from CSI data.

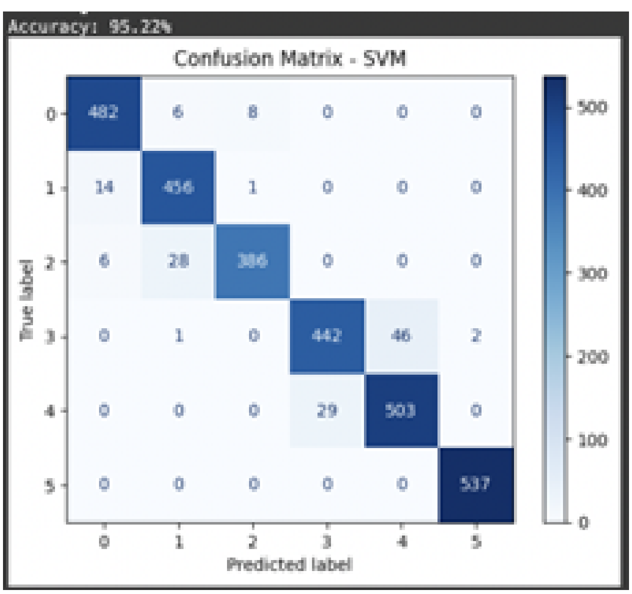

Fig. 4: Accuracy and confusion matrix for support vector machine

2) Random Forest This classifier achieved an accuracy of 92.5%, demonstrating reliable performance across all activity classes. In this approach data preprocessing was performed by imputing missing values with feature-wise means. The Random Forest was implemented with 100 decision trees (n_estimators = 100) and a fixed random state of 42 for reproducibility. A Random Forest constructs an ensemble of decision trees, each trained on a bootstrap sample of the training data with random feature selection at every split, thereby reducing correlation among trees and enhancing generalization. For an input sample x, the final prediction is determined by majority voting across all trees:

$$\hat{y} = mode\{h_t(x), t = 1, 2....T\} \tag{5}$$

where $h_t(x)$ denotes the output of the t^{th} decision tree and T is the total number of trees. Each decision tree within the ensemble determines its splits based on an impurity measure such as the Gini index, given by:

$$G = 1 - \sum_{k=1}^{K} p_k^2, \tag{6}$$

where p_k is the proportion of samples belonging to class k at a node. Alternatively, splits can be guided by Information Gain (IG), which quantifies the reduction in entropy after a split:

$$I.G.(D, A) = H(D) - \sum_{v \in Values(A)} \frac{|D_v|}{|D|} H(D_v), \tag{7}$$

where H(D) is the entropy of the dataset D, and D_v is the subset of D when attribute A takes value v. The overall predictive power of the Random Forest can also be formalized as the expected reduction in classification error:

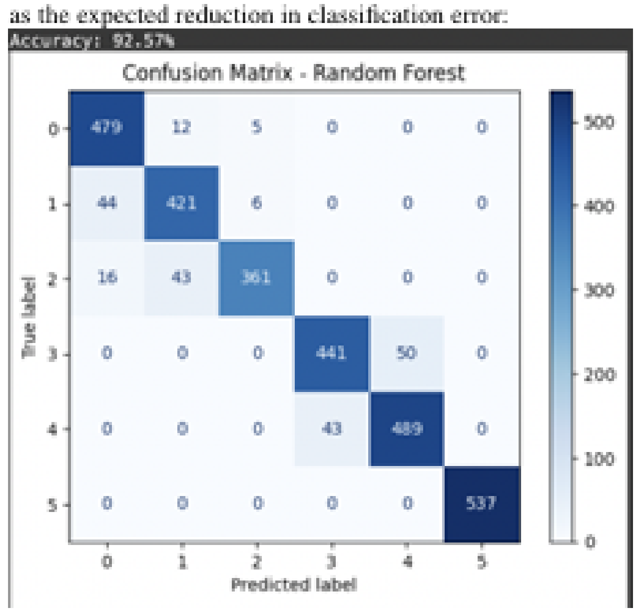

Fig. 5: Accuracy and confusion matrix for Random Forest

$$E = \frac{1}{T} \sum_{t=1}^{T} I(h_t(x) \neq y) \tag{8}$$

Where I is the indicator function. The trained Random Forest was evaluated using accuracy, precision, F1-score, and confusion matrix analysis, providing a well-rounded assessment of predictive performance. The results highlight that ensemble-based models like Random Forest achieve high robustness by combining the strengths of multiple weak learners, making them highly effective for complex tasks such as human activity recognition.

3) K-nearest neighbours This classifier achieved an overall accuracy of 88.02%, indicating reasonable performance but slightly lower than other models. The K-Nearest Neighbors (KNN) algorithm was employed for activity recognition using the UCI HAR dataset, with preprocessing steps including missing value imputation and feature scaling. The classifier was implemented with k = 5 nearest neighbors, classifying each test instance based on the majority class among its five closest training samples. The similarity between two data points x_i and x_j is computed using the Euclidean distance:

$$d(x_i, x_j) = \sqrt{\sum_{m=1}^{n} (x_i m - x_j m)^2} \tag{9}$$

Where n denotes the number of features, x_{im} is the m^{th} feature of instance and x_{jm} is the corresponding feature of instance x_j. The classification decision is then made by majority voting over the nearest neighbors:

$$\hat{y}(x_i) = mode\{y_j | x_j \in N_k(x_i)\}, \tag{10}$$

Where $N_k(x_i)$ represents the set of nearest neighbors of x_i. In weighted variants of KNN, closer neighbors are given higher importance using weights defined as:

$$w_{ij} = \frac{1}{d(x_i, x_j) + e} \tag{11}$$

Where e is a small constant to avoid division by zero. For this implementation, the unweighted version was used, ensuring all selected neighbors contribute equally. Model evaluation was carried out using accuracy, precision, F1-score, and confusion matrix, demonstrating that KNN effectively captures local neighborhood structure in data and provides competitive results for activity classification tasks.

4) Multilayer Perceptron Model This model achieved an accuracy of 95.11%, showing strong capability in capturing non-linear relationships between features. For this experiment, a Multilayer Perceptron classifier was implemented to perform human activity recognition on the UCI HAR dataset. The input features were standardized to zero mean and unit variance before training. The neural network consisted of a single hidden layer with 100 neurons, trained using the Adam optimizer with a maximum of 300 iterations and a regularization parameter of $\alpha = 10^{-4}$ .Each neuron computes a weighted sum of the inputs followed by a non-linear activation function. The forward propagation through a hidden layer can be expressed as Equation (12).

$$h^{(l)} = f(W^{(l)} x^{(l-1) + b^{(l)}} \tag{12}$$

Where $W^{(l)}$ and $b^{(l)}$ represent the weight matrix and bias vector of layer l, $x^{(l-1)}$ is the input from the previous layer, and $f()$ is the activation function (ReLU in this case).

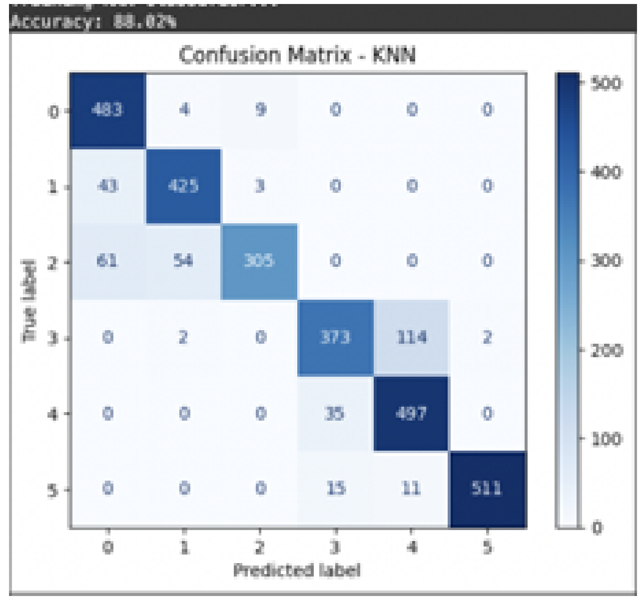

Fig. 6: Accuracy and confusion matrix for K-nearest neighbours

The final prediction is obtained using the softmax output layer:

$$\hat{y}_i = \frac{exp(z_i)}{\sum_{j=1}^{K} exp(z_j)} \tag{13}$$

Where z_i denotes the logit for class i and K is the total number of activity classes. The network parameters are optimized by minimizing the cross-entropy loss:

$$L = -\sum_{i=1}^{N} \sum_{k=1}^{K} y_{ik} \log \hat{y}_{ik} \tag{14}$$

With y_{ik} representing the true label in one-hot form. Model performance was assessed using accuracy, precision, F1-score, and confusion matrix visualization, confirming the capability of neural networks to learn complex non-linear mappings between sensor signals and human activities.

5) Enhanced Bidirectional Long Short-Term Memory: We proposed an Enhanced Bi-LSTM model for human activity recognition using WiFi CSI data. Raw CSI sequences are normalized and augmented via jittering, scaling, time warping, and mixup to improve generalization. Local temporal features are first extracted using residual CNN blocks, followed by bidirectional LSTM layers that capture long-range dependencies. A temporal attention mechanism highlights informative time steps, producing a context-aware representation, which is then classified through dense layers with dropout and a softmax output. The model is trained with categorical cross-entropy and label smoothing, employing either a cosine decay restart schedule or ReduceLROnPlateau for adaptive learning rate control, along with early stopping to prevent overfitting. This architecture efficiently models complex spatiotemporal patterns in CSI data, achieving high accuracy in human activity recognition. In this implementation, a hybrid deep learning architecture was designed

by combining convolutional layers, residual connections, bidirectional LSTMs, and temporal attention. This model aims to capture both local feature patterns from signals and long-term dependencies across time, while attention ensures that the most informative segments are emphasized during classification. Residual Learning To ensure stable gradient flow and avoid vanishing gradients, residual connections were used:

$$H(x) = F(x, W) + x \tag{15}$$

Here F(x,W) represents the output of stacked convolution and normalization layers, while the input x is directly added as a shortcut. This allows the network to preserve original information while learning transformations.

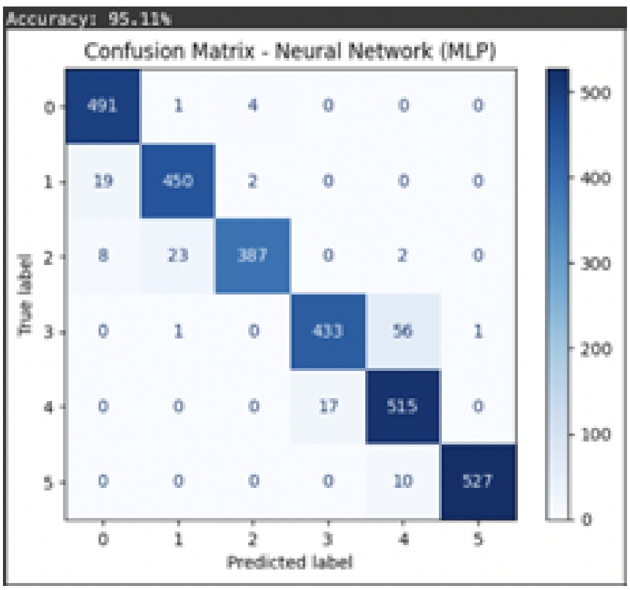

Fig. 7: Accuracy and confusion matrix for neural network (multiplayer perceptron model)

Temporal Attention Mechanism Instead of treating all timesteps equally, the model applies attention to weight important temporal features:

$$a_t = \frac{exp(e_t)}{\sum_{k=1}^{T} exp(e_k)}, c = \sum_{t=1}^{T} a_t h_t \tag{16}$$

Where h_t is the hidden state at timestep t, a_t is the learned importance score, and c is the final context vector representing the activity. This ensures that key signal segments contribute more to the decision.

Cosine Annealing Learning Rate To optimize training stability, a cosine decay with restarts strategy is used for the learning rate:

$$\eta_t = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min})[1 + cos(\frac{T_{cur}}{T_i}\pi)] \tag{17}$$

Where η_t is the learning rate at iteration t, T_{cur} is the current training step, and t_i is the restart period. This cyclic behavior prevents premature convergence. Evaluation Metrics Standard classification metrics-accuracy, precision, recall, and F1-score were used, as described earlier in the methodology section. These metrics evaluate the model's robustness in differentiating between human activities.

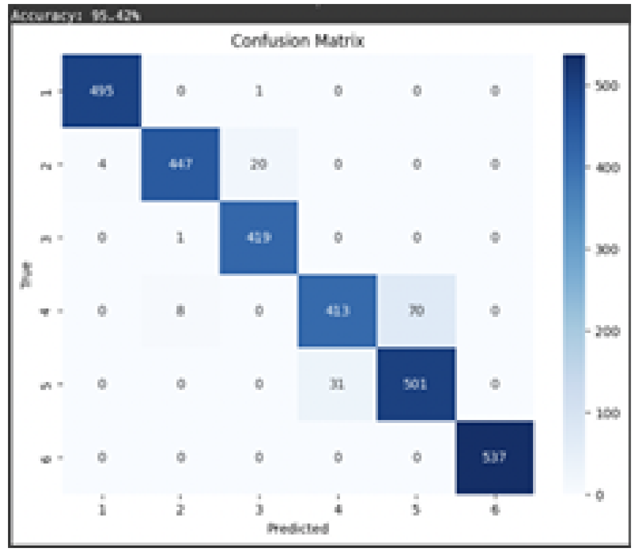

Fig. 8: Accuracy and confusion matrix Bidirectional LSTM

5.2. Performance Comparison

This section represents the evaluation results computed over different machine learning models. The results are summarized as Table 3.

Table 3: Evaluation results for selected ML models

Model	Accuracy	Precision	Recall	F1-Score
SVM	95.2%	92.1%	91.8%	91.9%
Random Forest	92.5%	91.5%	91.2%	91.3%
KNN	88.02%	88.4%	88.1%	88.2%
MLP	95.11%	93.3%	92.9%	93.0%
Enhanced Bi-LSTM	95.42%	95.5%	95.4%	95.4%

In the proposed framework, the sequence modeling stage is performed using Bi-LSTM layers, which effectively capture temporal dependencies in CSI signals by processing information from both past and future contexts. This bidirectional modeling proves crucial for human activity recognition, as activity patterns often depend on signal variations across entire time windows. By integrating dropout mechanisms to reduce overfitting and coupling the Bi-LSTM outputs with a temporal attention mechanism, the model selectively emphasizes the most informative segments of the sequence. This combination enables robust feature representation, leading to improved classification performance, with the proposed architecture achieving an accuracy of approximately 95-96% on CSI-based human activity recognition tasks.

5.3. Confusion Matrix Analysis

The confusion matrix for the Enhanced Bi-LSTM model reveals that while the model performs reasonably well on dynamic activities such as Walking, Walking Upstairs, and Walking Downstairs, it struggles to distinguish between static activities like Sitting and Standing. These two activities generate very similar signal patterns in the CSI data due to minimal body movement, leading to overlapping feature representations. As a result, the model is often misclassified between them. This indicates that the Bi-LSTM model, in its current configuration, lacks sufficient discriminative power for static postures and requires further optimization in terms of sequence modeling, feature enhancement, or training parameters to improve its class-wise accuracy.

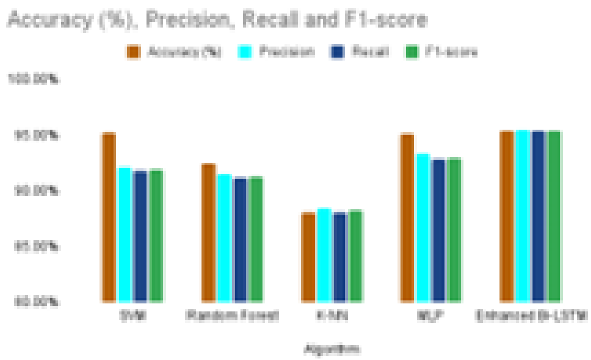

Fig. 9: Performance metrics of Algorithms on the UCI-HAR dataset

5.4. Practical Considerations

The Multilayer Perceptron neural network presents a practical balance between classification accuracy and computational efficiency. While it does not achieve the highest accuracy among all tested models (compared to Enhanced Bi-LSTM), it still performs significantly better than traditional classifiers in modeling non-linear relationships within the feature space. Due to its relatively simple architecture and feedforward nature-without the recurrent overhead of sequence modeling-MLP requires fewer computational resources, less memory, and shorter inference time. These attributes ensure that it fits the platform well where its resource requirements are low, i.e., either mobile devices, embedded systems, or edge computing units where real-time computing is required. In contrast, the classical machine learning models-particularly the Random Forest classifier-are not only computationally efficient but also highly interpretable. Random Forests are ensembles of decision trees which stratify features recursively, and the significance of the features may be measured straightforwardly. It facilitates an explanation that helps the practitioners to know what predetermines the decision of classification, what features are contributing the most to the predictions, and correcting/debugging of the prospective misclassification. Moreover, Random Forests are generally easy to tune in terms of hyperparameters, and they are resistant to overfitting, which means that they are a safe default algorithm to apply to HAR tasks, where little training data is available or the data is noisy. All in all, these models offer versatile solutions based on the application limitations: MLP offers moderately difficult, deployable at the edges, but could use better accuracy, while Random Forest offers fast and interpretable applications, either in the prototyping phase or low-latency systems.

6. CONCLUSION AND FUTURE WORK

We applied different algorithms on the UCI-HAR dataset to find it out that how these algorithms are responding for classification of different activities represented in the dataset. The classic machine learning algorithms like support vector machine is giving significantly good result. In the same way standard deep learning model like multilayer perceptron model also giving better performance. But the deep learning model Enhanced Bi-LSTM has given the best results till now. Data normalization and augmentation, local temporal feature extraction with CNN blocks and then applying this data on dense layered architecture of Bi-LSTM has improved the efficiency and performance of Bi-LSTM algorithm. So basically, it's observed, if we add some more layers for preprocessing of data on a layered deep learning architecture like Bi-LSTM so it may enhance the performance of the algorithm which would outperform the performance of other models. In the future work the focus would be on preparing the dataset by capturing it from WiFi signals and creating a sensor-less and contact-less system to detect the human activities by Wi-Fi.

References

[1] William Taylor, Syed Aziz Shah, Kia Dashtipour, Adnan Zahid, Qammer H Abbasi, and Muhammad Ali Imran. An intelligent non-invasive real-time human activity recognition system for next-generation healthcare. *Sensors*, 20(9):2653, 2020.

[2] Chenshu Wu, Zheng Yang, Zimu Zhou, Xuefeng Liu, Yunhao Liu, and Jiannong Cao. Non-invasive detection of moving and stationary human with wifi. *IEEE Journal on Selected Areas in Communications*, 33(11):2329–2342, 2015.

[3] Jiao Liu, Guanlong Teng, and Feng Hong. Human activity sensing with wireless signals: A survey. *Sensors*, 20(4):1210, 2020.

[4] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. Human respiration detection with commodity wifi devices: Do user location and body orientation matter? In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, pages 25–36, 2016.

[5] Feng Zhang, Chenshu Wu, Beibei Wang, Hung-Quoc Lai, Yi Han, and KJ Ray Liu. Widetect: Robust motion detection with a statistical electromagnetic model. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(3):1–24, 2019.

[6] Iftikhar Ahmad, Arif Ullah, and Wooyeol Choi. Wifi-based human sensing with deep learning: Recent advances, challenges, and opportunities. *IEEE Open Journal of the Communications Society*, 5:3595–3623, 2024.

[7] Boyu Yang and Ting Jiang. Wifi based multi-task sensing via selective sharing module. In *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, pages 1–6. IEEE, 2023.

[8] Steven M Hernandez and Eyuphan Bulut. Wifi sensing on the edge: Signal processing techniques and challenges for real-world systems. *IEEE Communications Surveys & Tutorials*, 25(1):46–76, 2022.

[9] Hasmath Farhana Thariq Ahmed, Hafisoh Ahmad, et al. Device free human gesture recognition using wi-fi csi: A survey. *Engineering Applications of Artificial Intelligence*, 87:103281, 2020.

[10] Yukai Song, William Taylor, Yao Ge, Kia Dashtipour, Muhammad Ali Imran, and Qammer H Abbasi. Design and implementation of a contactless ai-enabled human motion detection system for next-generation healthcare. In *2021 IEEE International Conference on Smart Internet of Things (SmartIoT)*, pages 112–119. IEEE, 2021.

[11] Guozhen Zhu, Beibei Wang, Weihang Gao, Yuqian Hu, Chenshu Wu, and KJ Ray Liu. Wifi-based robust human and non-human motion recognition with deep learning. In *2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pages 769–774. IEEE, 2024.

[12] Xuanqi Meng, Jiarun Zhou, Xiulong Liu, Xinyu Tong, Wenyu Qu, and Jianrong Wang. Secur-fi: A secure wireless sensing system based on commercial wi-fi devices. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*, pages 1–10. IEEE, 2023.

[13] Fahd Saad Abuhoureyah, Yan Chiew Wong, and Ahmad Sadhiqin Bin Mohd Isira. Wifi-based human activity recognition through wall using deep learning. *Engineering Applications of Artificial Intelligence*, 127:107171, 2024.

[14] Andrii Zhuravchak. Human activity recognition based on wifi csi data. 2020.

[15] KJ Ray Liu and Beibei Wang. *Wireless AI: Wireless Sensing, Positioning, IoT, and Communications*. Cambridge University Press, 2019.

[16] Muhammad Atif, Shapna Muralidharan, Heedong Ko, and Byounghyun Yoo. Wi-esp—a tool for csi-based device-free wi-fi sensing (dfws). *Journal of Computational Design and Engineering*, 7(5):644–656, 2020.

[17] Yuqian Hu, Feng Zhang, Chenshu Wu, Beibei Wang, and KJ Ray Liu. A wifi-based passive fall detection system. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1723–1727. IEEE, 2020.

[18] Muhammad Bilal Khan, Zhiya Zhang, Lin Li, Wei Zhao, Mohammed Ali Mohammed Al Hababi, Xiaodong Yang, and Qammer H Abbasi. A systematic review of non-contact sensing for developing a platform to contain covid-19. *Micromachines*, 11(10):912, 2020.

[19] Jun Qi, Po Yang, Lee Newcombe, Xiyang Peng, Yun Yang, and Zhong Zhao. An overview of data fusion techniques for internet of things enabled physical activity recognition and measure. *Information Fusion*, 55:269–280, 2020.

[20] Steve Blandino, Tanguy Ropitault, Claudio RCM da Silva, Anirudha Sahoo, and Nada Golmie. Ieee 802.11 bf dmg sensing: Enabling high-resolution mmwave wi-fi sensing. *IEEE Open Journal of Vehicular Technology*, 4:342–355, 2023.

[21] Yunhao Bai and Xiaorui Wang. Carin: Wireless csi-based driver activity recognition under the interference of passengers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1):1–28, 2020.

[22] Daqing Zhang, Hao Wang, Yasha Wang, and Junyi Ma. Anti-fall: A non-intrusive and real-time fall detector leveraging csi from commodity wifi devices. In *International Conference on Smart Homes and Health Telematics*, pages 181–193. Springer, 2015.