*Journal of Electrical and Electronic Engineering*
2025, Vol. 13, No. 4, pp. 143-153
https://doi.org/10.11648/j.jeee.20251304.11

Research/Technical Note

# Feasibility Study on Radar-based Human Activity Recognition

**Taiwo Samuel Aina***

School of Computer Science and Technology, University of Bedfordshire, Luton, United Kingdom

## Abstract

The growing interest in employing radar for human activity recognition is driven by the exponential rise in the incidence and risk of falls associated with aging, compounded by diminished leg strength, prolonged medication side effects, visual impairments, and other variables that contribute to decreasing strength. In comparison to contact devices and other non-contact devices, radar exhibits considerable advantages in terms of non-contact capability, accuracy, resilience, detection range, and privacy security. Radar-based Human Activity Recognition (HAR) works by using a Doppler frequency shift to figure out what people are doing. This shift creates unique Doppler signatures. The Doppler frequency shift is when electromagnetic waves change their frequency and wavelength depending on how fast the observer is moving compared to the source. This paper presents Radar based human activity recognition based on a convolutional neural network. Specifically, this paper utilized public datasets available by University of Glasgow, United Kingdom. The radar utilizes Novelda's X4 system-on-chip (SoC), with an integrated receiver and transmitter antenna, providing very precise distance and motion measurements. The target was located 0.45 meters from the radar at the time of data collection. The investigation makes use of PyTorch to implement classification through CNN architectures. The CNN model demonstrates effective ability to detect human activities within radar-based RF images. Although the model proves resilient it requires a larger collection of labelled data to reach higher performance standards.

## Keywords

Human Activity Recognition, Radar, CNN Architectures, Deep Learning Model

## 1. Introduction

By 2050, there will be a predicted 21.64% rise in the global population of elderly individuals over 65 [1]. With aging came an exponential increase in the effect and risk of falls because of decreased leg strength, long-term drug side effects, visual impairments, and other factors that reduced strength. Fall rates differ between countries. For instance, according to a Southeast Asian study, Japan has 20% of its older population fall per year, compared to 6 to 31% in China. According to research conducted in the American region, the annual percentage of senior individuals who fall varies from 21.6% in Barbados to 34% in Chile. Still, a lot of old individuals fall at home. According to estimates from 2002, 28.6% (26-31%) of Italians 65 years of age and older fall within a year. Of these, 43% have several falls. Home is where 60% of falls happen [2]. When it comes to older adults living independently in their own homes, about half of the falls happen in the home and its immediate surroundings. These falls typically happen in areas that are used frequently, like the kitchen, bathroom, living room, and

*Corresponding author: taiwo.aina@study.beds.ac.uk (Taiwo Samuel Aina)

**Received:** 19 February 2025; **Accepted:** 3 March 2025; **Published:** 4 July 2025

bedroom. The remaining falls happen in public places or in other people's homes [3]. The world has become quite concerned with fall detection and prediction in recent years [4]. Physiological reasons include age, a history of falls (for example, plantar phobia), mobility issues, sleep difficulties, and neurological illnesses are among the many factors that contribute to falls. Environmental factors also play a role. Dim light, smooth surfaces, and other environmental conditions are examples [5]. Though it doesn't stop falls from happening, fall prediction does necessitate taking into account all affecting circumstances. Health care providers are the only ones who should utilize fall risk assessment as a reference because fall prediction has a high likelihood of false alarms [6]. As such, the primary means of addressing the incidence of fall incidents is fall detection. Fall events and activities of daily living are the primary targets of fall detection systems now in use, which share a similar structure [7]. Sensing technologies play a crucial role in detecting and recognizing signals in healthcare. Over the past decade, a range of sensor technologies became available on the market. Unwanted outcomes and detrimental downstream impacts throughout the machine learning pipeline have been documented in recent research as a result of data problems [8, 9]. Given that falls are the primary cause of fatal injuries in the elderly, such as fractures, early detection of falls is crucial in preventing loneliness, fractures, loss of consciousness, and other related consequences. Therefore, the risk of falls in today's ageing society is a critical concern. Consequently, the number of systems designed to detect falls has significantly increased in recent years. Furthermore, medical studies of the harm caused by falls have shown that this is greatly influenced by the speed of response and rescue. These falls constitute a minimum of 50% of the reasons for hospitalization among the elderly, and around 40% of their non-natural causes of death. At a time when the identification and prevention of falls are vital for the welfare of older and susceptible persons, it is of paramount significance to get a high level of precision in detecting such occurrences [10].

The growing interest in employing radar for human activity recognition is driven by the exponential rise in the incidence and risk of falls associated with aging, compounded by diminished leg strength, prolonged medication side effects, visual impairments, and other variables that contribute to decreasing strength. Human activity recognition (HAR) with Doppler radar is essential for applications in smart homes, assisted living, and medical diagnostics [17]. In comparison to contact devices and other non-contact devices, radar exhibits considerable advantages in terms of non-contact capability, accuracy, resilience, detection range, and privacy security [11], greater resilience to variations in ambient conditions, including illumination, weather, temperature, and humidity [17]. Moreover, the power consumption and expense of radar-based systems are significantly cheaper than those of camera-based solutions. Radar-based Human Activity Recognition (HAR) works by using a Doppler frequency shift to figure out what people are doing. This shift creates unique Doppler signatures. The Doppler frequency shift is when electromagnetic waves change their frequency and wavelength depending on how fast the observer is moving compared to the source [12]. Typically, in radar-based human activity recognition (HAR), the raw radar signal is initially converted to the time-frequency domain using the short-time Fourier transform, resulting in a 2-D matrix known as the radar spectrogram [17]. Radar-based human activity recognition primarily encompasses continuous wave radar, ultra-wideband radar, and frequency-modulated continuous wave radar. FMCW radar is very good at finding people because it can measure quickly, has a low peak-to-average power ratio, and can measure both speed and distance at the same time [18].

Recent approaches in human motion classification have utilized deep neural networks (DNNs). Deep neural networks are limited, though, by the fact that they can't use large data sets for proper training and performance validation. Also, DNNs are only used for time-frequency (TF) representations, and their performance hasn't been fully tested on a difficult database like ours, which has a lot of participants, locations, and viewing angles [16]. This paper presents Radar based human activity recognition based on a convolutional neural network.

## 2. Radar Principle of Operation

Throughout the measurement process, an FMCW radar continuously emits a signal with a linearly modulated frequency [19]. The fundamental representation of a radar signal reflecting towards a target is given by [13]:

$$s_t(t) = \sin(2\pi f_c t + \pi a t^2) \tag{1}$$

In equation 1, the slope is defined as α and mathematically represented as:

$$\alpha = \frac{\Omega}{T_c}, \tag{2}$$

$T_c$ = chirp duration. $\Omega$ = chirp bandwidth

$$\text{The range resolution } R_{res},\ R_{res} = \frac{c}{2\Omega} \tag{3}$$

The range resolution $R_{res}$ is impacted by the chirp bandwidth, $\Omega$, and the speed of light, $c$

The received signal $s_r(t)$, corresponding to an attenuated and delayed $T_d$ copy of the transmitted signal $s_t$ (t), reflects from a target positioned at distance $d$ and is given by the following:

$$s_r(t) = \xi s_t(t - T_d) \tag{4}$$

attenuation coefficient = $\xi$, thus $T_d$ is given by:

$$T_d = \frac{2R}{c} \tag{5}$$

$R$ = target range.

Intermediate frequency (IF) signal is written as:

$$y(t) = -\frac{\xi}{2}\cos 2\pi\alpha T_d t + 2\pi f_c T_d - \pi\alpha T_d^2 \tag{6}$$

The distance calculation, maximum distance, velocity calculation and maximum velocity is given by [14]:

$$f_b = \frac{2d}{c} \cdot S \tag{7}$$

$S$ = the slope of the chirp and $d$ = distance between the radar and the objects, $d$ can be estimated as shown below:

$$d = \frac{f_b c}{2S} \tag{8}$$

The maximum distance is given as: [29].

$$d_{max} = \sqrt[4]{\frac{\sigma P_{tx} G_{tx} G_{rx} \lambda^2}{(4\pi)^3 L P_{rx}}} \tag{9}$$

In equation 9, $d_{max}$ = maximum detectable range

$P_{tx}$ = transmission power

$G_{tx}$ and $G_{rx}$ = the antenna gains of the transmitter and receiver

$\lambda$ = wavelength

$L$ = loss of the entire system

$P_{rx}$ = maximum power required for the reciever to receive the signal.

Assuming an object is moving at velocity $v$, then the displacement of the object between any two successive chirps would be $T_c \cdot v$.

The second chirp signal would have travelled an additional distance of $\Delta d = 2 \cdot T_c \cdot v$

The phase of a sinusoid before and after travelling a distance $\Delta d$ can be written as

$$2\pi f \left(\tau + \frac{\Delta d}{c}\right)$$

The phase shift would be:

$$\Delta\phi = 2\pi f \left(\tau + \frac{\Delta d}{c}\right) - 2\pi f \tau = 2\pi f_0 \frac{\Delta d}{c} = 2\pi \frac{\Delta d}{c} \tag{10}$$

Therefore, the phase shift between two chirps due to the velocity is:

$$\Delta\phi = 2\pi \frac{2 \cdot T_c \cdot v}{\lambda} \tag{11}$$

By transposition:

$$v = \frac{\lambda \Delta\phi}{4\pi T_c} \tag{12}$$

Since the phase of a signal always has a range of [$-\pi$, $\pi$],

Unambiguous measurement of the phase requires $|\Delta\phi| \leq 180°$ or $|\Delta\phi| \leq \pi$.

Considering the extreme situation where $\Delta\phi = \pi$, t

Then:

$$v_{max} = \frac{\lambda\pi}{4\pi T_c} = \frac{\lambda}{4T_c} \tag{13}$$

The maximum range, $R_{max}$,

$$R_{max} = \frac{c}{2\alpha T_s} \tag{14}$$

## 3. Related Work

The study in [12] suggests a radar-based system for detecting human activity that combines range-time-Doppler maps with range-distributed convolutional neural networks. To see how well the proposed model worked, tests were done using the "radar signatures of human activities" dataset from the University of Glasgow. The proposed model did better at identifying things and making recognition errors less frequent compared to CNNs with the same number of parameters. The authors in reference [16] presented Radar Data Cube Processing for Human Activity Recognition utilizing Mult subspace Learning. The proposed RDC-based method starts with a step called eCLEAN, which is meant to get rid of any unwanted data distortions and noise artifacts. They showed that the multidimensional PCA method is better than those that use predefined features, 1-D and 2-D PCA features, and a 12-layer CNN by using MDC. They collected two sets of data from two different indoor scenarios. Five distinct motions were examined in the experiments: falling (191 samples), sitting (213 samples), bending (203 samples), kneeling (108 samples), and walking (112 samples). They executed movements in five distinct orientations: 0°, 30°, 45°, 60°, and 90°. For a 5-class activity recognition task, the suggested method worked 97.2% of the time using 1D-PCA, preset features, 2D-PCA, and CNN on the spectrograms. For 1D-PCA, pre-defined features, 2D-PCA, CNN, and MPCA, the average classification accuracy is 65.32%, 73.65%, 83.10%, 84.54%, and 91.4%, in that order. In study [17], a simple data enhancement method for micro-Doppler radar-based human activity identification (HAR) was described. Their method for adding to the training dataset works well because it keeps the kinematic information in the spectrograms. The proposed technique can be seamlessly included into the training process of the DNNs without necessitating any pretraining phase. By mitigating the overfitting issue resulting from inadequate data, radar-based human activity recognition can employ larger networks with deeper architectures. Their investigations revealed that the augmentation strategy enhances generalization to previously unencountered data, which is crucial for practical applications, and also exhibits robustness across many contexts. The research in [18] mostly examines the implementation of continuous HAR technology. Their research is mostly about finding the best order and window for the fractional order domain of radar data. Utilizing the multi-input multi-task (MIMT) recognition network, the characteristics of each domain are concurrently evaluated, and various input representations are merged to achieve continuous activity categorization results with high accuracy. They compare the segmentation efficacy of the activity detector using VW-STA/LTA with two other segmentation methods to examine its impact. This article suggests the CMDN, which effectively gets around the problems with current methods that rely on single-domain data and fixes the issue of how to accurately separate and find ongoing human activities. The authors in reference [21] developed a real-time radar-based gesture recognition system implemented on an edge-computing platform. Their proposed multi-feature encoder might effectively store the gesture profile as a feature cube, which can then be fed into a shallow CNN for gesture classification. They used a variety of inputs to get continuous activity classification results with a high level of 98.47% precision and an average level of 93.11% accuracy, using gestures from people who were told what to do and people who weren't told what to do. Their suggested radar-based gesture recognition system could be used for many things, like mobile and wearable tech, because it works well at detecting gestures and doesn't require a lot of processing power. Future projects will require the creation of diverse gesture datasets tailored to specific applications. Also, the proposed system might not be able to accurately classify things when the radar is not fixed in relation to the user. Table 1 below depict summarized literature review on related work.

***Table 1.*** *Radar-Based HAR.*

| Research Description | Methodology | Efficiency | Research Gaps/ Recommendation | Reference |
|---|---|---|---|---|
| Radar-Based Human Activity Recognition Combining Range-Time-Doppler Maps and Range-Distributed-Convolutional Neural Networks | Tests were done using the "radar signatures of human activities" dataset from the University of Glasgow | The proposed model did better at identifying things | Recognition errors less frequent compared to CNNs with the same number of parameters. | [12] |
| Radar Data Cube Processing for Human Activity Recognition Using Multi subspace Learning | Mult subspace Learning, PCA | Multidimensional PCA method is better than those that use predefined features | For 1D-PCA, pre-defined features, 2D-PCA, CNN, and MPCA, the average classification accuracy is 65.32%, 73.65%, 83.10%, 84.54%, and 91.4%, in that order. | [16] |
| RadarSpecAugment: A Simple Data Augmentation Method for Radar-Based Human Activity Recognition | Augmentation method | Exhibits robustness across many contexts. | Their investigations revealed that the augmentation strategy enhances generalization to previously unencountered data, which is crucial for practical applications | [17] |
| CMDN: Continuous Human Activity Recognition Based on Multi-domain Radar Data Fusion | short-time fractional Fourier transform (STFrFT) to map radar data into the fractional domain | Utilizing the multi-input multi-task (MIMT) recognition network, the characteristics of each domain are concurrently evaluated, and various input representations are merged to achieve continuous activity categorization results with high accuracy. | This article suggests the CMDN, which effectively gets around the problems with current methods | [18] |
| HAROOD: Classifying Human Activity and Detecting Out-of-Distribution using short-range FMCW | Suggested a two-stage network for classification that combines triplet loss, contrastive loss, and intermediate reconstruction loss | For out-of-distribution detection, an average AUROC of 95.04% and an average classification | Future research could examine how this approach performs in more complicated settings and how it can | [19] |

| Research Description | Methodology | Efficiency | Research Gaps/ Recommendation | Reference |
|---|---|---|---|---|
| radar | in the first stage, and cross-entropy loss in the second. | accuracy of 96.51% were obtained. | be used to a wider variety of tasks. | |
| A metric learning method for activity recognition based on MIMO radar | Used a metric learning technique for categorization after capturing micro-doppler and angular velocity signatures using a MIMO radar. | 88.9% classification accuracy was attained for eight activities; 86.42% accuracy was attained for 10 activities using few-shot learning. | In order to improve model resilience, future research might concentrate on expanding the dataset's size and diversity and investigating how this method might be applied in practical situations. | [20] |
| Real-time gesture recognition and detection using radar integrated into an edge computing platform | Created a framework for real-time data processing with a 60 GHz FMCW radar system, extracting detailed hand profiles and utilizing a shallow CNN to recognize gestures. | Displayed the capacity to accurately classify 12 gestures in real time while maintaining a high F1-score. | Future research could examine how this system can be integrated into multiple applications and evaluate how well it performs in various real-world situations. | [21] |
| Human Activity Recognition (HAR) using millimeter wave radar for a medical monitoring robot | For real-time monitoring, a lightweight deep neural network system with a light-PointNet backbone and a bidirectional lightweight LSTM model was proposed for a moveable robot-mounted mmWave radar system. | Beat prior research on both continuous and discrete HAR tasks by a significant margin. | Future studies could examine how this approach is implemented in different healthcare environments and evaluate how well it functions in practical situations. | [22] |
| Recognizing radar-based activities with CNN-LSTM network architecture | Used convolutional layers to train features and LSTM layers to improve temporal information in a CNN-LSTM architecture for radar micro-doppler signature picture classification. | Accuracy for training and testing data was 96.8% and 93.5%, respectively. | Future research could examine how well this architecture performs with various radar systems and investigate its application to a broader range of activities. | [23] |
| Radar HAR with a Deep Learning Network based on attention | Suggested a deep learning network for radar HAR that is based on attention, with an emphasis on improving feature extraction utilizing attention mechanisms. | Showed higher recognition accuracy as compared to conventional methods. | Future studies could examine how attention mechanisms can be applied to different kinds of radar data and how well they work in various HAR tasks. | [24] |
| HAR based on DenseNet and frequency-modulated continuous waves | Gathered FMCW radar point clouds, then using a DenseNet neural network to identify human activity from these readings. | For five tasks, 100% recognition accuracy was attained. | In order to evaluate the system's resilience and generalizability, future research might concentrate on testing it in increasingly complicated settings and with a wider range of tasks. | [25] |
| mmWave radar-based attention-based vision transformer for classifying human activities | Used a slice segment technique and time-frequency feature representation of the micro-Doppler map to create a modified vision transformer network for radar-based HAR. | Demonstrated improved performance over conventional methods. | Future studies should examine how vision transformers can be applied to other kinds of radar data and how well they work in various HAR tasks. | [26] |
| Continuous classification of human activity using Bi-LSTM networks from FMCW radar | Classified continuous human activities from FMCW radar data using Bi-LSTM networks, paying particular attention to | High categorization accuracy was attained for a number of activities. | Bi-LSTM networks may be used to additional kinds of radar data in future research, and their performance in | [27] |

| Research Description | Methodology | Efficiency | Research Gaps/ Recommendation | Reference |
|---|---|---|---|---|
| | temporal dependencies in the data. | | various HAR tasks may be examined. | |
| Classification of human activity using Deep Convolutional Neural Networks based on micro-Doppler signatures | Used deep convolutional neural networks to categorize human activity using radar-captured micro-Doppler signals. | Showed remarkable classification accuracy across a range of tasks. | Future studies could concentrate on enhancing the model's resistance to data fluctuations and investigating how it might be used in practical situations. | [28] |

## 4. Methodology

This section thoroughly investigated human activity recognition using radar-based radio frequency (RF) signals. The aim of the system is to classify various human activities based on the RF signal pattern, which enables real-time and batch classification.

### 4.1. Dataset Description

This investigation utilized public datasets available by University of Glasgow, United Kingdom. The dataset link: https://researchdata.gla.ac.uk/view/author/39488.html. The Radar Dataset used in this paper was divided into six distinct folders: A, E, I, O, U, and Empty. The dataset comprises several activities collected by the UWB Radar sensor (Xethru X4M03). The radar utilizes Novelda's X4 system-on-chip (SoC), with an integrated receiver and transmitter antenna, providing very precise distance and motion measurements. The target was located 0.45 meters from the radar at the time of data collection. Every task required 6 seconds for completion. We linked Radar to a PC with Intel(R) Core (TM) i7-7700 3.60 GHz processors and 16 GB RAM using the modular connection XEP. We employed the experimental configuration to obtain the data [15].

*Table 2. Details of the Data Set (Folders, Files, Description, and Number of Samples) [15].*

| Number of Subjects | Folder Name | Technology | Description | Number of Samples per Class |
|---|---|---|---|---|
| 1 | Subject1(M)_With Mask | Radar | Male subject pronounced the vowel and Empty Data using Mask | 300 |
| | Subject1(M)_Without Mask | | Male subject pronounced the vowel and Empty Data using Without Mask | 300 |
| 2 | Subject2(F1) _With Mask | Radar | Female subject pronounced the vowel and Empty Data using Mask | 300 |
| | Subject2(F1) _Without Mask | | Female subject pronounced the vowel and Empty Data using Without Mask | 300 |
| 3 | Subject2(F2) _With Mask | Radar | Female subject pronounced the vowel and Empty Data using Mask | 300 |
| | Subject2(F2) _Without Mask | | Female subject pronounced the vowel and Empty Data using Without Mask | 300 |

The features are derived from images along with the main categories which include:

1) Pixel-level Feature (Raw Image Data)
2) Convolutional Features (Learned Features)
3) This investigation adopts the letters derived from metadata labelling as its target feature known as Class Labels (Target Feature).

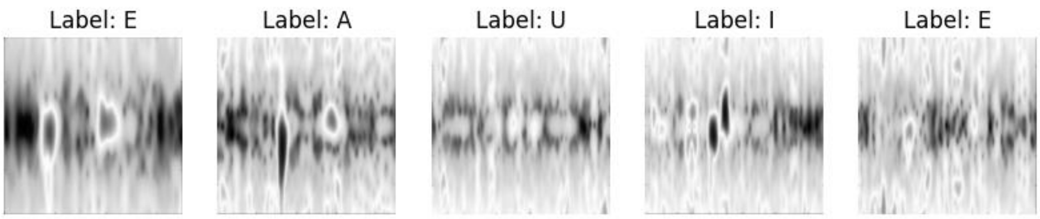


***Figure 1.*** *Target feature known as Class Labels.*

### 4.2. Data Acquisition

The first stage of radar based human activity identification through deep learning involves data acquisition. Various environmental and operational frequency events and antenna configuration patterns affect the quality profile of data acquisition. The Authors in [16] selected commercially available FMCW radars because of their compact physical design and Doppler distance information together with time and range capabilities. They also used a multilinear subspace learning method to analyses radar data cubes that combine slow-time, fast-time and Doppler frequency elements to improve HAR effectiveness.

### 4.3. Data Acquisition and Segmentation

Similar to the mathematical normalization of a dataset, data preprocessing can have a substantial effect on the accuracy of feature extraction and picture analysis results. The initial data collected by radar systems normally carries both unwanted noise levels and superfluous data elements, hence require process of fine-tuning involves to modify the parameters of the previously trained model to better suit the features of the target dataset. The primary functions of preprocessing steps involve signal quality improvement while extracting relevant features from data. The three main preprocessing techniques include clutter removal and normalization procedures and transferring the data into time-frequency forms that appear as spectrograms. Radar Spec Augment represents an impressive technique which executes data augmentation procedures directly on radar spectrograms by implementing time shifts as well as frequency masking to enhance neural network capabilities during restricted training processes [17]. During segmentation, continuous radar data are broken into distinct segments which corresponds to individual activities. In [18], Short-Time Fractional Fourier Transform (STFrFT) techniques were utilised to convert radar data into fractional fields for extracting precise motion features for continuous HAR.

## 5. Implementation Result

### 5.1. Confusion Matrix

In Figure 2, we plotted confusion matrix of proposed system for the subject labelled as A, E, I, O, U, and Empty. A confusion matrix displays classification performance through its depiction of actual labels on the y-axis and predicted labels on the x-axis. The confusion matrix contains cells which present the number of predicted samples that fall under specific labels. Higher values in the data correspond with dark color blocks while lower values appear as light colors. Model performance analysis starts with the confusion matrix because it reveals vital information regarding which classes have better or worse prediction results. Correct classifications appear in the diagonal cells when predicted labels match the true labels. True label "A" along with other labels result in misclassifications appearing in cells located outside the main diagonal of the confusion matrix. The model shows complete ability to identify "Em" class data points while facing challenges when classifying "A" and "E" cases which produces non-zero readings outside its diagonal values. The graphical display provides insights into specific performance areas of individual classes to help developers improve modeling effectiveness (for example through weight modifications or increased training examples for specific classes). The matrix helps analysts understand both false categorization percentages and measures for model quality enhancement.

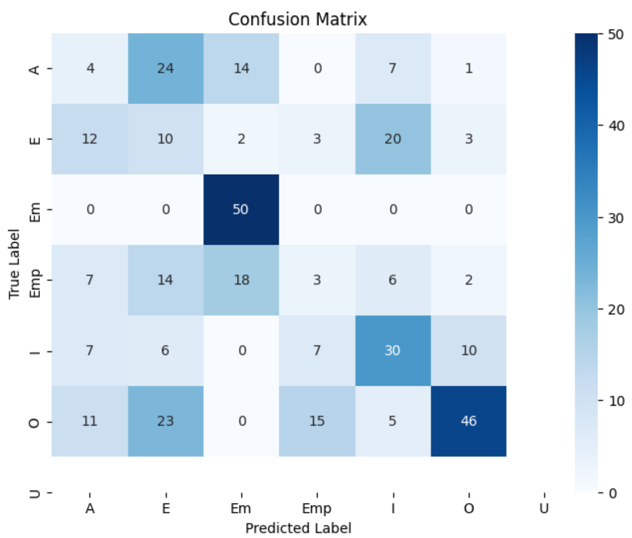


***Figure 2.*** *Confusion matrix of proposed system.*

## 5.2. True Against Predicted Label Plot

In Figure 3, we plotted True against Predicted Label. The given scatter plot illustrates both actual blue dots representing true labels along with predictive red crosses for different samples. The y-axis shows the label values from 0 to 6 while the x-axis displays the sample indexes. This display shows what degree the model matches its predicted tags with true label values. The correct predictions are shown through the overlapping of matching blue dots with red crosses in the plot. Non-overlapping points indicate misclassifications. A mismatch will appear in this plot when the model predicts value 2 for a sample with actual label 3. The model struggles most to predict correctly based on how the red crosses deviate from the actual labels.

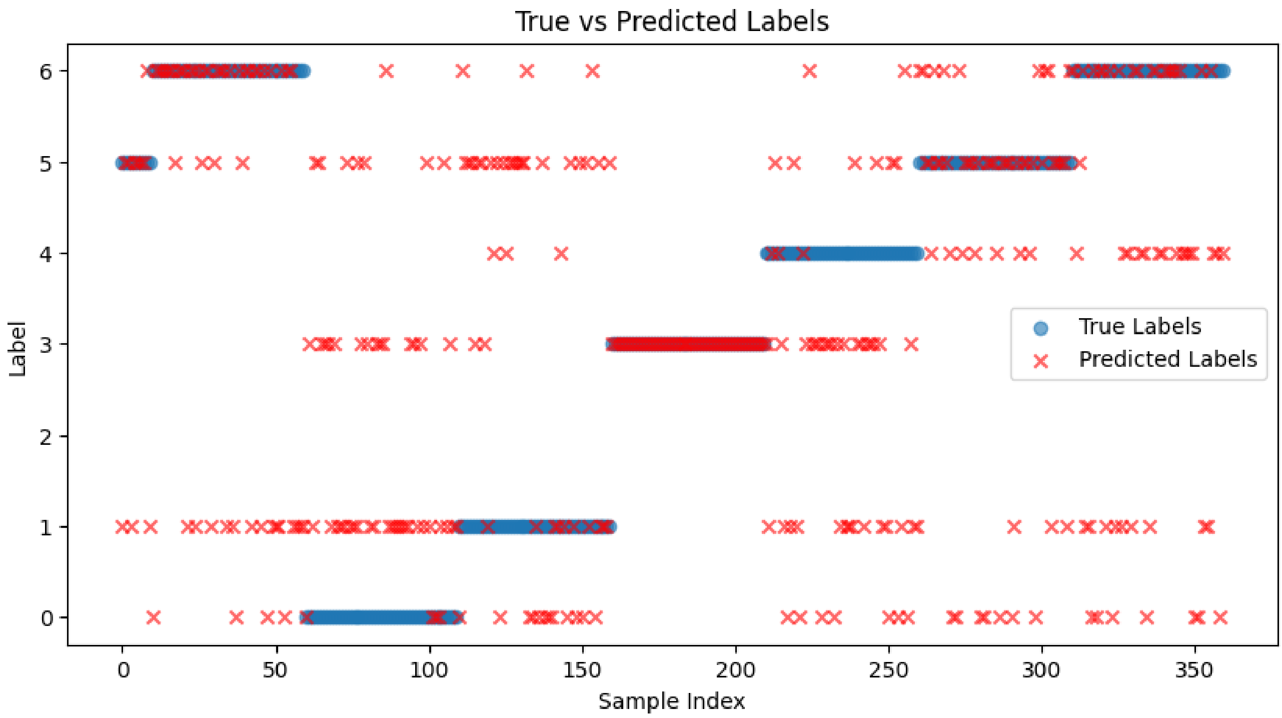


***Figure 3.*** *True against Predicted Label plot.*

## 5.3. Training Loss and Accuracy Plot

In Figure 4, we plotted Training Loss and Accuracy. Two lines represent the data points in the graph where loss is shown in blue and accuracy appears in red while both axes display values against epochs. The horizontal axis displays epochs whereas the vertical axis represents values for both loss and accuracy. The model performs error reduction successfully during training which loss measurements demonstrate. The blue line demonstrates that the loss remains at a low steady state throughout all epochs thus indicating the model has reached a point of stability in loss minimization. The accuracy measurement exhibits growth through each epoch cycle that is visible on the red line. Accuracy increases steadily throughout the epochs until it reaches an optimal point where it shows minor improvements. This pattern shows that the model converges to an optimal accuracy level after which more training probably would not enhance its performance significantly.

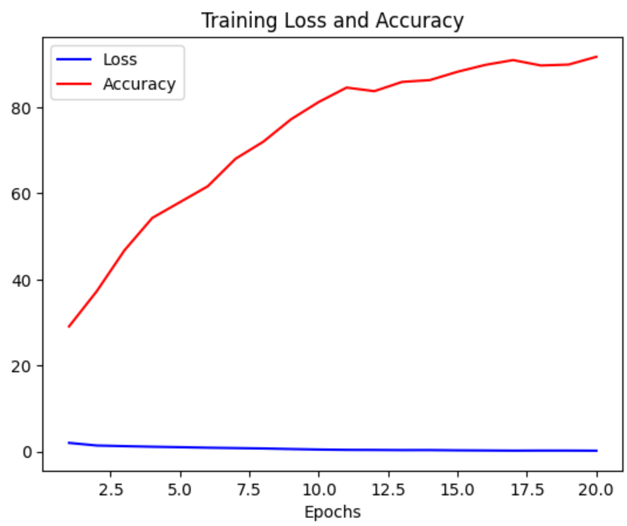


***Figure 4.*** *Training loss and accuracy.*

### 5.4. Velocity against Time Plot

In Figure 5, we plotted Velocity against Time Plot. The presented graph displays velocity values (red line) that change according to each epoch (period of time). The velocity data points stand on the y-axis alongside epoch numbers located across the x-axis. The speed of model learning or change across epochs represents velocity for this purpose. Changes in learning dynamics become visible through variations in the velocity measurement. Model learning speed accelerates during intervals of significant velocity changes between epochs but velocity reduction suggests learning stagnation occurs. The large differences between velocities show that learning patterns are unstable during this process. The model requires better learning rate adjustments possibly due to data-related fluctuations in training efficiency throughout different epochs. Academic research shows that a flat velocity graph indicates the best training conditions because it represents system stability and learning consistency during the process.

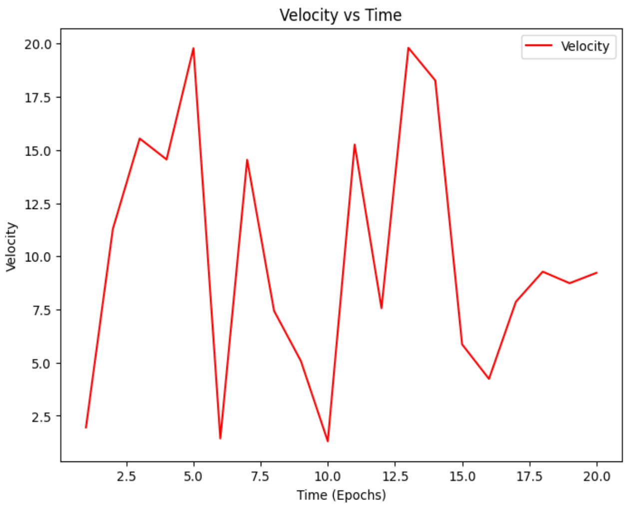


***Figure 5**. Velocity against Time Plot.*

## 6. Performance Evaluation

There are four main ways to measure how well the proposed radar domain and CNN models work: accuracy, precision, recall, and F1 score in this paper. The metrics are derived by juxtaposing projected outcomes with actual findings, yielding four potential returns: true positive (TP), true negative (TN), false positive (FP), and false negative (FN) [13].

### 6.1. Accuracy

Accuracy denotes the comprehensive efficacy of the model. It is defined as the ratio of accurate forecasts, encompassing both true positives and true negatives, to the total predictions produced. [13].

$$\text{Accuracy (\%)} = \frac{TP+TN}{TP+FP+TN+FN}100$$

### 6.2. Precision

Precision assesses the model's reliability in categorizing an incident as positive. It emphasizes the ratio of true positives (TP) to the total number of positive predictions, encompassing both true positives and false positives (FP). Precision is crucial for reducing Type-1 errors (false positives) [13].

Precision is computed as TP divided by TP plus FP.

$$\text{Precision (\%)} = 100\frac{TP}{TP+FP}$$

### 6.3. Sensitivity

In machine learning, sensitivity or recall, quantifies the ability of a model to correctly identify positive events. It is the ratio of true positives accurately identified as positive [13].

$$\text{Sensitivity (\%)} = 100\frac{TP}{TP+FN}$$

### 6.4. F1 score

In order to offer a balanced assessment of a system's performance, the F1 score takes into account both precision and recall. The F1 score is a harmonic mean of recall and precision, offering a balanced assessment of both measurements [13].

Mathematically,

$$F1 = \frac{2+(\text{Precision}+\text{Recall})}{\text{Precision}+\text{Recall}}$$

F1 is the weighted average of Precision and Recall.

The model's performance is analyzed using accuracy scores, loss curves, and validation performance. Result includes:

Epoch 1, Loss: 2.0473, Accuracy: 29.10%
Epoch 2, Loss: 1.4396, Accuracy: 37.29%
Epoch 3, Loss: 1.2930, Accuracy: 46.81%
Epoch 4, Loss: 1.1545, Accuracy: 54.31%
Epoch 5, Loss: 1.0486, Accuracy: 57.99%
Epoch 6, Loss: 0.9264, Accuracy: 61.67%
Epoch 7, Loss: 0.8400, Accuracy: 68.06%
Epoch 8, Loss: 0.7548, Accuracy: 72.01%
Epoch 9, Loss: 0.6214, Accuracy: 77.22%
Epoch 10, Loss: 0.5053, Accuracy: 81.25%
Epoch 11, Loss: 0.4302, Accuracy: 84.58%
Epoch 12, Loss: 0.4121, Accuracy: 83.75%
Epoch 13, Loss: 0.3744, Accuracy: 85.90%
Epoch 14, Loss: 0.3806, Accuracy: 86.32%
Epoch 15, Loss: 0.3195, Accuracy: 88.26%
Epoch 16, Loss: 0.2843, Accuracy: 89.86%
Epoch 17, Loss: 0.2447, Accuracy: 90.97%
Epoch 18, Loss: 0.2649, Accuracy: 89.72%
Epoch 19, Loss: 0.2595, Accuracy: 89.93%
Epoch 20, Loss: 0.2314, Accuracy: 91.74%

## 7. Conclusion

Human activity recognition is an essential task with several applications in the entertainment, security, and medical sectors. In recent decades, the discipline of Human Activity Recognition has seen a revolution thanks to machine learning techniques, especially deep learning algorithms. This paper mainly investigated human activity recognition based on a convolutional neural network. The classification of the subsections reflects the researchers' primary focus on specific stages and highlights the primary contributions of this study. Analyzing individual sample predictions becomes easier through inspection of the true versus predicted plot. The training loss and accuracy plot demonstrates how the model learns during its entire process. A confusion matrix provides detailed performance assessment by examining model success or failure according to each classification category. The velocity-time plot enables observers to assess training dynamics while helping detect abnormal patterns during the process. The trained deep learning system achieves high performance for detecting human activities by processing radar-based RF image data. An improvement in model results would require larger datasets that have been appropriately annotated.

## Abbreviations

| | |
|---|---|
| HAR | Human Activity Recognition |
| CNN | Convolutional Neural Network |

## Author Contributions

Taiwo Samuel Aina is the sole author. The author read and approved the final manuscript.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

[1] D FAD (2018). Health Quality and Safety Commission, New Zealand. Falls in People Aged 50 and Over. https://www.hqsc.govt.nz/our-programmes/health-quality-evaluation/projects/atlas-of-healthcare-variation/falls/

[2] D Castro, F., Dentamaro, V., Gattulli, V., Impedovo, D. (2023). Fall Detection with LSTM and Attention Mechanism. Workshop on Advances of Mobile and Wearable Biometrics, Athens, Greece pp. 1-14.

[3] D Mehta, J., Czanner, G., Harding, S., Newsham, D., Robinson, J. (2022). Visual Risk Factors for Falls in Older Aadults: a case-control study BMC Geriatr 22(1): 1-9 https://doi.org/10.1186/S12877-022-02784-3/TABLES/3

[4] D Usmani, S.; Saboor, A.; Haris, M.; Khan, M. A.; Park, H. (2021). Latest Research Trends in Fall Detection and Prevention Using Machine Learning: A systematic review. Sensors 21: 5134.

[5] D Oumaima, G.; Hamd, A. A.; Youness, T.; Rachid, O. H. T.; Omar, B. (2021). Vision-based Fall Detection and Prevention for the Elderly People: A review and ongoing research. In Proceedings of the 2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS), Fez, Morocco, pp. 1-6.

[6] D Chaccour, K.; Darazi, R.; El Hassani, A. H.; Andres, E.(2016). From Fall Detection to Fall Prevention: A generic classification of fall-related systems. IEEE Sensor Journal 17: 812-822.

[7] D Xefteris, V. R.; Tsanousa, A.; Meditskos, G.; Vrochidis, S.; Kompatsiaris, I. (2021). Performance, Challenges, and Limitations in Multimodal Fall Detection Systems: A review. IEEE Sensor Journal 21: 18398-18409.

[8] D Paullada, A., Raji, I. D., Bender, E. M., Denton, E., Hanna, A. (2021). Data and its (dis) contents: A survey of dataset development and use in machine learning research. Patterns 2(11): 100336.

[9] D Renggli, C., Rimanic, L., Gürel, N. M., Karlaš, B., Wu, W., Zhang, C. (2021). A Data Quality-Driven View of MLOps. Data Engineering pp. 11.

[10] D Umer, M. et al. (2024) 'Enhancing fall prediction in the elderly people using LBP features and Transfer learning model', Image and Vision Computing, 145, p. 104992. https://doi.org/10.1016/j.imavis.2024.104992

[11] Y. Yao, W. Liu, G. Zhang, and W. Hu, "Radar-based human activity recognition using hyperdimensional computing," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 3, pp. 1605-1619, Mar. 2022. https://doi.org/10.1109/tmtt.2021.3134992

[12] W.-Y. Kim and D.-H. Seo, "Radar-based human activity recognition combining range-time-doppler maps and range-distributed-convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-11, 2022. https://doi.org/10.1109/tgrs.2022.3162833

[13] F. Ayaz *et al.*, *Radar Signal Processing and its impact on Deep Learning-driven human activity recognition*, Dec. 2024. https://doi.org/10.20944/preprints202412.0637.v1

[14] H. Cui, "Human activity recognition using millimetre-wave radars with machine learning," *University of Bristol*. http://research-information.bristol.ac.uk/

[15] Hameed, H., Usman, M., Tahir, A., Hussain, A., Abbas, H., Cui, TJ., Imran, M. and Abbasi, Q. (2022), Pushing the Limits of Remote RF Sensing: Reading Lips Under Face Mask, University of Glasgow http://dx.doi.org/10.5525/gla.researchdata.1282

[16] B. Erol and M. G. Amin, "Radar Data Cube processing for human activity recognition using Multisubspace learning," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 6, pp. 3617-3628, Dec. 2019. https://doi.org/10.1109/taes.2019.2910980

[17] D. She, X. Lou, and W. Ye, “Radarspecaugment: A simple data augmentation method for radar-based human activity recognition,” *IEEE Sensors Letters*, vol. 5, no. 4, pp. 1-4, Apr. 2021. https://doi.org/10.1109/lsens.2021.3061561

[18] X. Feng, P. Chen, Y. Weng, and H. Zheng, “CMDN: Continuous human activity recognition based on Multi-domain Radar Data Fusion,” *IEEE Sensors Journal*, pp. 1-1, 2025. https://doi.org/10.1109/jsen.2025.3530921

[19] K. S. Mustafa, M. S. Yavuz, and E. Steinbach, “HAROOD: Human Activity Classification and Out-of-Distribution Detection with Short-Range FMCW Radar,” *arXiv (Cornell University)*, Jan. 2023, https://doi.org/10.48550/arxiv.2312.08894

[20] F. Aziz, O. Metwally, P. Weller, U. Schneider, and M. F. Huber, “A MIMO Radar-Based Metric Learning Approach for Activity Recognition,” *arXiv (Cornell University)*, Jan. 2021, https://doi.org/10.48550/arxiv.2111.01939

[21] Y. Sun, T. Fei, X. Li, A. Warnecke, Ernst Warsitz, and N. Pohl, “Real-Time Radar-Based Gesture Detection and Recognition Built in an Edge-Computing Platform,” *IEEE Sensors Journal*, vol. 20, no. 18, pp. 10706-10716, Sep. 2020, https://doi.org/10.1109/jsen.2020.2994292

[22] Z. Gu, X. He, G. Fang, C. Xu, F. Xia, and W. Jia, “Millimeter Wave Radar-based Human Activity Recognition for Healthcare Monitoring Robot,” *arXiv.org*, 2024. https://arxiv.org/abs/2405.01882 (accessed Feb. 02, 2025).

[23] Dr. A. H. Victoria, S. V. Manikanthan, Dr. V. H R, M. A. Wildan, and K. H. Kishore, “Radar Based Activity Recognition using CNN-LSTM Network Architecture,” *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 14, no. 3, pp. 303-312, Jan. 2023, https://doi.org/10.17762/ijcnis.v14i3.5630

[24] S. Huan, L. Wu, M. Zhang, Z. Wang, and C. Yang, “Radar Human Activity Recognition with an Attention-Based Deep Learning Network,” *Sensors*, vol. 23, no. 6, p. 3185, Mar. 2023, https://doi.org/10.3390/s23063185

[25] W. Jiang *et al.*, “Human Activity Recognition Based on Frequency-Modulated Continuous Wave and DenseNet,” *Journal of Computer and Communications*, vol. 11, no. 07, pp. 15-28, Jan. 2023, https://doi.org/10.4236/jcc.2023.117002

[26] Z. Wang, S. Huan, L. Wu, Q. Wang, J. Liu, and Z. Hu, “Attention-based Vision Transformer for Human Activity Classification using mmWave Radar,” pp. 128-134, Nov. 2022, https://doi.org/10.1145/3577164.3577184

[27] Z. Li, X. Liu, T. Wang, W. He, and C. Xia, “GHSCN: A Graph Neural Network-Based API Popularity Prediction Method in Service Ecosystem,” *IEEE Access*, vol. 8, pp. 137032-137051, Jan. 2020, https://doi.org/10.1109/access.2020.3011439

[28] L. Barbato, A. Formisano, R. Martone, and F. Villone, “Error Field Impact on Plasma Boundary in ITER Scenarios,” *IEEE Transactions on Magnetics*, vol. 52, no. 3, pp. 1-4, Sep. 2015, https://doi.org/10.1109/tmag.2015.2480417

[29] D. Weyer *et al.*, “Design considerations for Integrated Radar Chirp synthesizers,” *IEEE Access*, vol. 7, pp. 13723-13736, 2019. https://doi.org/10.1109/access.2019.2893313

## Biography

**Taiwo Samuel Aina** is currently pursuing the Ph.D. degree in electronic engineering with University of Bedfordshire, Luton, United Kingdom. His current research interests include human activity recognition, radar signal processing, deep learning, and fusion of sensing technologies. He received MSc. degrees in electronic engineering with Project Management from University of Bedfordshire, Luton, United Kingdom, in 2023 and the M. Eng. degree from the Department of Electrical and Electronics Engineering, University of Benin, Nigeria, in 2016. Since 2022, He has been working with the Bedford College Group, United Kingdom as Electrical installation, Electronics and mathematics Lecturer.