IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Traitement du Signal
Vol. 43, No. 3, June, 2026, pp. 1371-1381
Journal homepage: http://iieta.org/journals/ts

Hybridization of Convolutional Neural Network-Based Human Suspicious Activity Detection and Classification Using Radar Micro-Doppler Images for Defense Applications

Shanmugasami Abikayil Aarthi[1*], James Arputha Vijaya Selvi[2]

[1] Department of CSE, Kings College of Engineering, Thanjavur 613303, India
[2] Department of Electronics and Communication Engineering, Kings College of Engineering, Thanjavur 613303, India

Corresponding Author Email: aarthi.cse@kingsengg.edu.in

https://doi.org/10.18280/ts.430322

Received: 18 February 2026
Revised: 11 May 2026
Accepted: 18 May 2026
Available online: 30 June 2026

Keywords:
suspicious human activity, radar micro-Doppler image, EfficientNet-B6, deep learning, bidirectional gated recurrent unit

ABSTRACT

The recognition of suspicious human activities is a crucial component of national defense strategies. The human activity recognition (HAR) using vision-based models has several drawbacks. In videos, recognizing small human actions is a complex and labor-intensive task. As a result, HAR using radar is a developing research area owing to its resilience. Radar-based HAR generally depends on the micro-Doppler (m-D) features of target echoes. From the time-Doppler graph, the m-D features can emphasize the self-vibration and rotation of human limbs and the torso. Due to the distinctive correspondence between human behavior and m-D attributes, supervised learning models are regularly employed for radar-based HAR. In recent years, deep learning has been gradually evolving, and its outstanding classification performance has also gathered significant interest. Radar-based HAR research has become more sophisticated with the integration of deep learning systems. This research presents a Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) approach. This research aims to effectively identify and classify human activities using radar m-D images for defense applications. For accomplishing that, the proposed HNN-HSADC framework primarily pre-processes the input images using noise removal, contrast enhancement, image resizing, and normalization to refine image quality for subsequent analysis. For feature extraction, the EfficientNet-B6 is utilized to capture informative features from pre-processed radar images. Finally, the convolutional neural network with bidirectional gated recurrent unit is adopted for accurate classification of human activities. Extensive simulation studies are conducted to evaluate the enhanced performance of the HNN-HSADC model. The comparative result analysis reported the betterment of the HNN-HSADC method on recent approaches in terms of different evaluation measures.

1. INTRODUCTION

Human activity recognition (HAR) is crucial in many domains like counterterrorism, security monitoring, biomedical patient health monitoring, border surveillance, and initial recognition of public violent attacks/protests. The need for intelligent methods capable of detecting and classifying suspicious human behavior has become essential for triggering counter-action mechanisms, for controlling the state, and/or for situation analysis/post-scenario [1]. These days, various HAR methods are accessible depending on surveillance video, smart vision sensing, infrared, thermal, closed-circuit television (CCTV), Radio Frequency (RF) sensing (radars), and acoustic sensors [2]. HAR employing radar is a developing research area presently because of its flexibility for through-wall imaging, on-ground classification, air and/or foliage target images, ground penetrative recognition, low light, long-range operations, and severe weather conditions, among other things. Radar-based HAR typically relies on analyzing the micro-Doppler (m-D) effect present in the echoes reflected from a target. These distinct m-D signatures, often visualized in a time-Doppler spectrogram, effectively emphasize the subtle movements resulting from the self-vibration and rotation of a person's torso and limbs [3]. From a security perspective, frequently encountered suspicious human activities include individuals fighting or engaging in one-on-one assaults (such as punching or boxing); unauthorized entry for the purpose of pre-attack surveillance (like military marching); training exercises (such as army jogging); firing a weapon or fleeing with a rifle (e.g., jumping while holding a firearm); launching projectiles to cause harm or an explosion (stone-pelting or grenade-throwing); and covert movement for an attack or escape (e.g., army crawling) [4]. Timely and accurate identification/categorization of these specific types of alarming human behaviors, ideally at the earliest possible stage, is critical, and can be effectively achieved by analyzing their distinctive m-D signatures.

In recent times, scholars have employed Machine Learning (ML) models that require distinct time-frequency (T-F) map features' identification/retrieval methods, namely multilayer perceptron, principal component analysis (PCA), support vector machine (SVM), and linear predictive coding (LPC) for

HAR. Therefore, the effectiveness of feature extraction employing certain conventional models for identification is limited by previous data and the intricacy of classification issues [5]. In addition, the identification performance of the type of method degrades if the m-D signature resemblance is higher for a larger number of classes and/or across classes [6]. Recently, Deep Learning (DL) models' excellent classification outcomes have also gathered substantial interest. Radar-driven HAR studies have obtained more intelligence because of the integration of DL models [7]. DL acts as an automated feature extractor, with enhanced self-regulation, self-control, and self-actuation because of its decision-making skills as well as inherent computation, which motivated us to utilise the DL model for individual suspicious action identification [8]. Transformers, recurrent neural networks (RNNs), convolutional neural networks (CNNs), and hybrid networks are the four main types of DL models. This approach uses supervised learning for automating feature representation, therefore addressing the restrictions of traditional techniques to extract features. Employing recursive neural networks and time-sequence techniques may obtain temporal correlation features among information series [9]. Multiple studies have proved that enhancing long short-term memory (LSTM) as well as Bidirectional long short-term memory (BiLSTM) structures in a system may efficiently improve HAR's performance metrics.

1.1 Paper contribution

- **Proposed Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) method:** Aims to accurately detect and classify human activities using radar m-D images for defense applications.
- **Preprocessing pipeline:** Applied noise removal, contrast enhancement, image resizing, and normalisation to refine image quality.
- **Feature extraction**: Utilized EfficientNet-B6 to extract significant features from pre-processed radar images.
- **Classification:** Designed a CNN-Bidirectional Gated Recurrent Unit (BiGRU) to effectively classify human activities.
- **Comprehensive evaluation:** Conducted extensive experiments and a comparative study to validate the improvements of the HNN-HSADC method under different evaluation measures.

2. EXISTING RESEARCH ON SUSPICIOUS HUMAN ACTIVITY RECOGNITION FROM RADAR IMAGES

A summary of existing related works regarding suspicious human activity detection and classification using radar images is presented in this section. Pareek et al. [10] provided an innovative structure to enhance radar-driven individual suspicious activities identification by employing TL models. This study article starts with a detailed examination of field adaptation, feature extraction, and, finally, fine-grained TL approaches. This approach efficiently classifies suspicious activities through varied environmental conditions, indicating its generalisation as well as robustness capabilities. Moreover, the research classifies the implications and advantages of integrating TL into real-life observation and security methods. Waghumbare et al. [11] solved this task by presenting an augmentation approach personalized for radar m-D signatures. Models like frequency disturbance, time shift, and frequency shift are used for generating different information, thus enriching the training data and improving the technique's generalization. This limitation restricts the depth and performance of DCNN techniques. Furthermore, most DCNN techniques are computationally expensive and inherently difficult, which exacerbates the problem of overfitting in certain situations.

Table 1. Literature review of human suspicious activity detection using radar image

Authors	Models/ Dataset	Evaluations	Limitations
Pareek et al. [10]	TL models/ Radar image dataset	Accuracy of 99.77% and 99.8%	Limited by its evaluation on a single dataset, which might limit the generality of the proposed framework across diverse environments.
Waghumbare et al. [11]	DCNN/Huge dataset	Achieves a better outcome	Augmented data may not entirely represent the intricacy and variability of real-world radar m-D signatures.
Wang et al. [12]	Transformer/University of Glasgow (UoG) dataset	Accuracy of 96.61%	The main limitation is that the method still depends on large pre-trained models, which require high computational resources.
Nguyen et al. [13]	CNN techniques/Dual dataset	Accuracy of 95% and 99%	The model's performance may drop in real-world environments with complex noise or multiple moving subjects.
Guendel et al. [14]	CNN approaches	Reaches enhanced performance	The method requires multiple radar sensors and complex signal processing, making it hard to use in real-time or large-scale setups.
Zhou et al. [15]	Bi-GRU/NA	Accuracy of 90%	The major drawback is that the model's accuracy may decrease when dealing with overlapping or highly similar continuous activities.
Nguyen et al. [16]	DCNN/ImageNet and Noise datasets	Accuracy of 99%	The outcome can drop if the radar signals contain complex or unpredictable types of noise not handled by the denoising algorithm.

Note: TL = Transfer Learning; DCNN = Deep Convolutional Neural Network; CNN = convolutional neural network; Bi-GRU = Bidirectional Gated Recurrent Unit; m-D = micro-Doppler.

Wang et al. [12] suggested an innovative method that uses Low-Rank Adaptation (LoRA) fine-grained in the weight space for enabling knowledge transfer from pre-trained ViT methods. Moreover, for fine-grained characteristic extraction, to enhance feature extraction in the combined serial-parallel adapters in characteristic space. The novel fine-grained approach (tuned to radar-driven time-Doppler signatures) increases the precision of HAR by a significant amount. Efficient monitoring is needed to prevent serious damages and fatalities resulting from falls; fall detection. Nguyen et al. [13] presented the use of a frequency-modulated continuous wave (CW) radar sensor for detecting movement. A stacked-residual convolutional neural network (SRCNN) is proposed to classify the daily individual actions that rely on the m-D characteristics of recovered radar signals. This approach is based on a dual-layer SR framework for re-using the previous features, which improves the identification accuracy. In the study of Guendel et al. [14], the issues of multi-path in radar sensor networks for HAR have been studied. The multi-path is known as a source of additional clutter, and the multi-path is being analysed for the possibility of being exploited during the generation of virtual radar nodes. A novel processing step is presented to realize this concept, and the data from multi-path signals are obtained to enhance the HAR.

Zhou et al. [15] proposed a method to classify continuous individual actions, which are standing, jumping, squatting, walking, and running. The technique relies on the properties of the m-D formed from continuous-wave radar waves. Initially, the procedure involves the radar signals from individual actions for producing m-D spectrograms. Then, a Bi-GRU system is used for training and testing these characteristics extracted. To eliminate the white Gaussian noise in the unprocessed radar data before detecting individual behaviours, Nguyen et al. [16] introduced a new method based on a denoising algorithm using a DCNN. In particular, the denoising model is used in the pre-processing phase. Then, a Cross-Residual convolutional neural network (CRCNN) with flexible CR links is suggested for identification. Table 1 portrays a comparison analysis of previous works, comprising their models, datasets, evaluations, and limitations.

3. PROPOSED SYSTEM

This study presents an HNN-HSADC technique to accurately identify and classify human activities using radar m-D images for defense applications. Figure 1 illustrates the workflow of the HNN-HSADC model. As depicted in Figure 1, the proposed method proceeds through a sequence of phases.

•Apply noise removal, contrast enhancement, image resizing, and normalization to prepare input images for analysis.

•Feed pre-processed images into EfficientNet-B6 to obtain deep feature vectors.

•The extracted features are passed to a CNN-Bi-GRU hybrid classifier to learn spatial patterns and temporal dependencies to perform human activity classification.

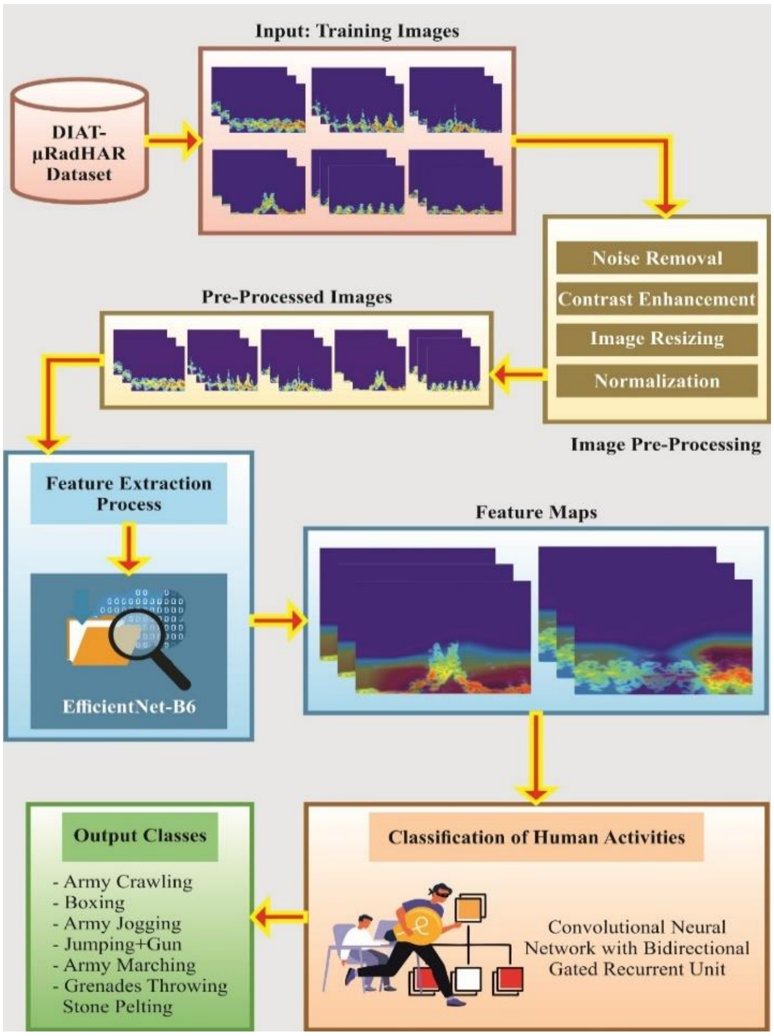

Figure 1. Overall workflow of the proposed Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) system

3.1 Dataset description

The DIAT-μRadHAR dataset [17, 18] contains m-D spectrograms captured employing a custom X-band CW 10 GHz radar system, targeting detection of suspicious human behaviour like jogging, marching, army crawling, boxing, jumping while holding a gun, and grenade throwing/stone-pelting. It was gathered in an open-field setting with numerous human subjects of differing heights, weights, and genders performing the listed actions at distances from 10 m to 0.5 km and at aspect angles of 0°, ±15°, ±30°, and ±45° relative to the radar. Intended to enable deep-learning-driven detection of human suspicious activities in radar information. This dataset consists of 2560 samples under six actions as displayed below in Table 2. Figure 2 shows the spectrogram images.

Table 2. Details of the dataset

Activities	No. of Instances
Army Crawling	400
Boxing	420
Army Jogging	410
Jumping + Gun	400
Army Marching	490
Grenades_Throwing Stone_Pelting	440
Total Instances	**2560**

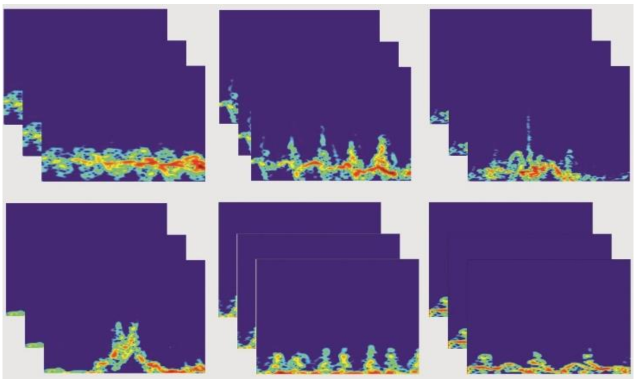

Figure 2. Spectrogram images

To validate the robustness and consistency of the proposed framework, the dataset was evaluated under two different train-test split configurations, namely 80:20 and 70:30. The training data was used to learn the model, and the testing data was used to test the model on unseen data. All the samples underwent the same preprocessing operations: noise removal, contrast enhancement, resizing, and normalization, prior to training.

3.2 Image preprocessing

Initially, noise removal, contrast enhancement, image resizing, and normalisation are used as preprocessing models to improve the image quality for further analysis. Before training with Doppler power spectra imaging, images must be pre-processed. In the database, the images are three-channel (RGB) colour images of varying dimensions, which might be affected by dissimilar noise signals [19]. To make an imaging into an even dimensions, eliminate the noise, and improve the quality, this image is handled utilizing a pre-processing structure. CLAHE is employed on an input image for enhancing edge definition and local contrast. Then, the histogram equalization output is exposed to removal of noise utilizing Wiener filter (WF). The WF is nothing but a linear adaptive filter, which refines noise while maintaining edges and other higher-frequency image portions. Next, the resizing process was performed on the denoised imaging for measuring down the image size to 120 × 120 pixels with 3 channels. Then, the output images are normalized within the range of [0, 1] to make the calculation easier. Afterward, the pre-processed images were employed to train the classifier methods.

3.3 Feature extraction via EfficientNet-B6

For feature extraction, the EfficientNet-B6 is deployed to capture beneficial features from pre-processed radar images. EfficientNet-B6 is part of the EfficientNet series, proposed to improve the efficacy of CNNs over a balanced scaling model [20]. Traditional scaling methods generally increase one dimension of the network (depth, width, or resolution) independently, which often leads to sub-optimal performance and increased computational cost. EfficientNet overcomes this restriction with compound scaling, which measures every three dimensions evenly, depending upon a compound coefficient ϕ.

Compound Scaling: The main advance in EfficientNet-B6 is the compound scaling model. The network's depth d, width w, and resolution r are determined using a predetermined set of scaling parameters: α, β, and γ.

$$d = \alpha^{\phi} \tag{1}$$

$$w = \beta^{\phi} \tag{2}$$

$$r = \gamma^{\phi} \tag{3}$$

For EfficientNet-B6, the compound coefficient ϕ is fixed to 6, which defines the extent to which every dimension is scaled. The specific values of α, β, and γ originate from a grid search intended to enhance the trade-off between accuracy and computational efficacy.

Architectural Building Blocks: EfficientNet-B6 is constructed utilizing Mobile Inverted Bottleneck Convolution (MBConv) blocks with optimization of Squeeze-and-Excitation (SE). Every MBConv block includes:

Expansion Phase: Enlarges an input network by a factor, permitting richer feature extraction:

$$Conv_{expand}(x) = ReLU6\big(Conv2D_{1\times 1}(x)\big) \tag{4}$$

Depthwise Convolution Phase: Employs depth-wise separable convolutions, which perform spatial convolution independently across every channel, decreasing the cost of computation:

$$\mathrm{Convdw}\ (\mathrm{x}) = \mathrm{ReLU6}\big(\mathrm{DepthwiseConv2D}(\mathrm{x})\big) \tag{5}$$

SE Phase: Recalibrates channel-wise feature response by demonstrating interdependencies among channels:

$$\mathrm{SE}(\mathrm{x}) = \sigma(\mathrm{W}_2 \mathrm{ReLU}(\mathrm{W}_1 \mathrm{GlobalAvgPool}\ (\mathrm{x}))) \tag{6}$$

Projection Phase: Decreases the channel count back to preferred output size, enabling effective data flow over the network:

$$\mathrm{Conv}_{\mathrm{project}}(\mathrm{x}) = \mathrm{Conv2D}_{1\times 1}(\mathrm{x}) \tag{7}$$

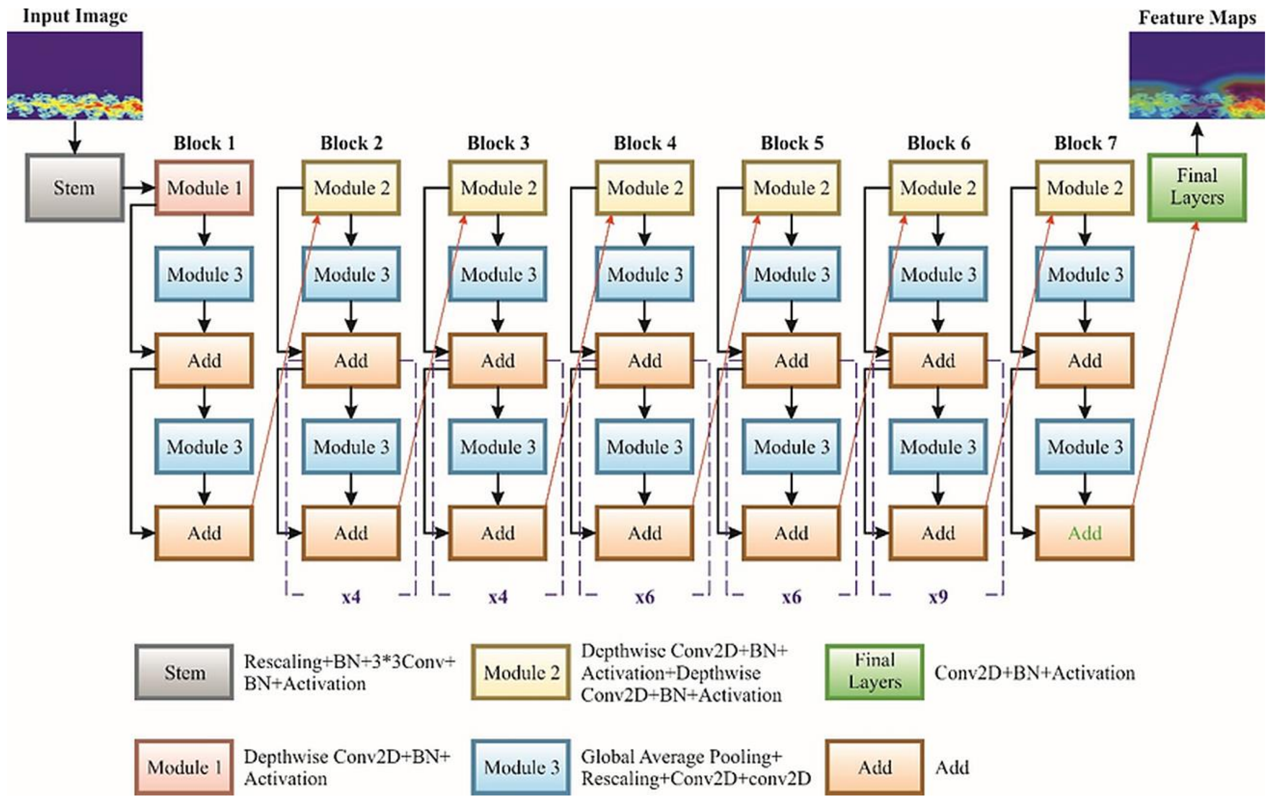

Figure 3. Architecture of EfficientNet-B6

EfficientNet-B6 Architecture: EfficientNet-B6 is organized with repeated MBConv blocks, where every block is arranged according to the compound scaling tactic. The network starts with a standard convolutional layer, followed by a sequence of MBConv blocks divided into phases. Every stage functions at a different resolution and channel depth, allowing the network to capture hierarchical features. The structure of EfficientNet-B6 is summarized below:

Stem: Initial 3 × 3 convolution with 32 filters.

Stage: Repeated MBConv blocks with varying expansion factors, kernel dimensions, and output channels.

Top: Final 1 × 1 convolution to yield the desired number of output classes, tracked by global average pooling layers and fully connected (FC) layers for classification.

EfficientNet-B6 is trained utilizing standard models, comprising dropout, data augmentation, and batch normalization. The Adam optimiser is generally employed to update the model parameters, with a learning rate schedule to enable convergence. EfficientNet-B6 leverages its balanced compound scaling to deliver a highly effective and computationally efficient method. The usage of MBConv blocks with SE components improves its feature extraction abilities, making it a powerful model for image classification and other computer vision tasks. By enlarging input resolution and scaling width and depth regularly, EfficientNet-B6 attains a good trade-off among efficiency and accuracy, making it appropriate for higher-resolution, intricate image analysis tasks. Figure 3 demonstrates the architecture of EfficientNet-B6.

3.4 Human activity classification model

Finally, the CNN-BiGRU is used for accurate human activity classification. CNN is composed of pooling and convolutional layers, in which the convolutional layers represent the non-linear local features of the data and the pooling layers represent the reduction of attributes in order to increase the generality of the data [21]. This framework enhances the capability of the model to acquire significant data and generalise effectively. Bi-GRU is an improved version of GRU that is able to process sequential data in both forward and backward directions. They show significant advancement in RNN, making them increasingly relevant in different areas. In the CNN-BiGRU technique, the input data is supervised by the CNN layer that removes the relevant spatial features. These attributes are combined and simplified to achieve a low-dimensional model. This compressed model is fed to the Bi-GRU layer, which has the ability to learn sequential dependency from the forward and backward directions and enhances the retrieved attributes. The features have been pre-processed to be normalized and converted into embeddings of numerical features. These embeddings leverage the input to the CNN-BiGRU framework, which enhances the outcome by leveraging the elements of CNN to recognize complex spatial patterns and the elements of Bi-GRU to enhance the detection of the model. In the dual classifier, CNN-BiGRU is used to handle the input.

Implementing the proposed HNN-HSADC framework was done in Python using the TensorFlow and Keras libraries. To extract the features, the EfficientNet-B6 model was used, and the CNN-BiGRU model was trained using the Adam optimizer with a learning rate of 0.001. The batch size was set to 32, and the model was trained for 200 epochs. The convolution layers used kernel sizes of 3 × 3 and ReLU as the activation function. To effectively capture temporal dependencies from the extracted feature representations, the Bi-GRU layer used 128 hidden units. To avoid overfitting and to enhance the generalization ability, the dropout regularization technique was also employed.

Input Layer: Assume the input data depicted as a matrix $X \in \mathbb{R}^{n*d}$, here n denotes the no. of instances, and d signifies the no. of attributes.

Convolutional Layers: Dual 1D convolution layers were employed sequentially to eliminate spatial attributes from the input.

$$H^{(1)} = \mathrm{ReLU}(X * W_1 + b_1) \tag{8}$$

Now b_1 implies the bias, $*$ signifies the convolutional operation, and $W_1 \in \mathbb{R}^{k_1 * d}$ refers to the kernel of size k_1. The Output: $H^{(1)} \in \mathbb{R}^{n-k_1+1*f_1}$, where f_1 signifies the no. of filters. For the 2nd convolution layer, the Input: $H^{(1)}$ and convolution.

$$H^{(2)} = \mathrm{ReLU}\left(H^{(1)} * W_2 + b_2\right) \tag{9}$$

Here, $W_2 \in \mathbb{R}^{k_2 * f_2}$ and b_2 imply kernel and bias, correspondingly. The Output: $H^{(2)} \in \mathbb{R}^{n-k_1-k_2+2*f_2}$, now f_2 refers to the no. of filters.

Every output of the convolutional layer undergoes BN to stabilize training.

$$H_{BN}^{(l)} = \gamma \frac{H^{(l)} - \mu}{\sigma} + \beta \tag{10}$$

Now γ and β denote learnable scaling parameters, μ and σ refer to the mean and SD of the batch.

Bi-GRU Layer: The outcome from convolution layers, $H_{BN}^{(2)}$ is passed to Bi-GRU. In the forward pass of time-step t.

$$\overrightarrow{h_t} = \mathrm{GRU}_{\mathrm{forward}}(h_{t-1}, x_t) \tag{11}$$

Now $\mathrm{GRU}_{\mathrm{forward}}$ leverages a gating mechanism to update the hidden state.

$$\overleftarrow{h_t} = \mathrm{GRU}_{\mathrm{backward}}(h_{t+1}, x_t) \tag{12}$$

The Bi-directional Output integrated hidden state has been specified.

$$h_t = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right] \tag{13}$$

Here [;] signifies concatenation. The outcome is $H_{BiGRU} \in \mathbb{R}^{n*2h}$, now h implies the no. of hidden components.

FC Layer: The outcome from the Bi-GRU layers has been flattened.

$$H_{flat} = \mathrm{Flatten}\,(H_{BiGRU}) \tag{14}$$

Subsequently, the FC layer maps the attributes to a high-dimensional space.

$$Z = \mathrm{ReLU}(H_{flat} \cdot W_{fc} + b_{fc}) \tag{15}$$

Here, $W_{fc} \in \mathbb{R}^{d*64}$ and b_{fc} imply the weights and biases. Dropout is employed to avoid overfitting.

$$Z_{drop} = \mathrm{Dropout}(Z) \tag{16}$$

Afterwards, in the last FC layer, a single unit is output.

$$y = \mathrm{sigmoid}\left(Z_{drop} \cdot W_{out} + b_{out}\right) \tag{17}$$

Here $W_{out} \in \mathbb{R}^{64*1}$.

Output Layer: The outcome of sigmoid activation is a probability $y \in [0,1]$ utilizing threshold τ.

$$class = \begin{cases} & y \geq \tau \\ & y < \tau \end{cases} \tag{18}$$

In a multi-class process, a Softmax activation can substitute the sigmoid function.

$$y = \mathrm{Softmax}\left(Z_{drop} \cdot W_{out} + b_{out}\right) \tag{19}$$

Higher outcomes of the classifier are accomplished by this framework's effective integration of CNN and Bi-GRU, and FC layers have resilient categorization.

4. EXPERIMENTAL EVALUATION

In this section, the performance analysis of the HNN-HSADC system is examined under the DIAT-μRadHAR dataset.

Figure 4 shows the confusion matrices created by the HNN-HSADC model below 80:20 of training and testing phases. The result implies that the HNN-HSADC approach has effective classification and detection of every class accurately.

Confusion Matrix
Training Phase (80%)

Actual / Predicted	Army Crawling	Boxing	Army Jogging	Jumping+ GIM	Army Marching	Grenades Throwing Stone Pelting
Army Crawling	324	1	1	0	1	1
Boxing	1	311	2	6	3	2
Army Jogging	3	0	319	3	0	0
Jumping+ GIM	3	0	5	300	0	2
Army Marching	1	1	0	1	389	3
Grenades Throwing Stone Pelting	0	0	6	3	0	356

(a)

Confusion Matrix
Testing Phase (20%)

Actual / Predicted	Army Crawling	Boxing	Army Jogging	Jumping+ GIM	Army Marching	Grenades Throwing Stone Pelting
Army Crawling	69	1	1	0	0	1
Boxing	1	90	0	2	0	2
Army Jogging	0	1	82	2	0	0
Jumping+ GIM	0	2	0	88	0	0
Army Marching	1	2	3	0	89	0
Grenades Throwing Stone Pelting	0	0	1	0	0	74

(b)

Figure 4. Confusion matrices of (a) 80% training phase and (b) 20% testing phase

Table 3. Activities detection of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model under 80% and 20%

Class Labels	$Accur_y$	$Preci_n$	$Recal_l$	$F1_{Score}$	AUC_{Score}
Training Phase (80%)					
Army Crawling	99.41	97.59	98.78	98.18	99.16
Boxing	99.22	99.36	95.69	97.49	97.79
Army Jogging	99.02	95.80	98.15	96.96	98.67
Jumping + Gun	98.88	95.85	96.77	96.31	98.01
Army Marching	99.51	98.98	98.48	98.73	99.12
Grenades Throwing Stone Pelting	99.17	97.80	97.53	97.67	98.53
Average	**99.20**	**97.56**	**97.57**	**97.56**	**98.55**
Testing Phase (20%)					
Army Crawling	99.02	97.18	95.83	96.50	97.69
Boxing	97.85	93.75	94.74	94.24	96.65
Army Jogging	98.44	94.25	96.47	95.35	97.65
Jumping + Gun	98.83	95.65	97.78	96.70	98.41
Army Marching	98.83	100.00	93.68	96.74	96.84
Grenades Throwing Stone Pelting	99.22	96.10	98.67	97.37	98.99
Average	**98.70**	**96.16**	**96.19**	**96.15**	**97.71**

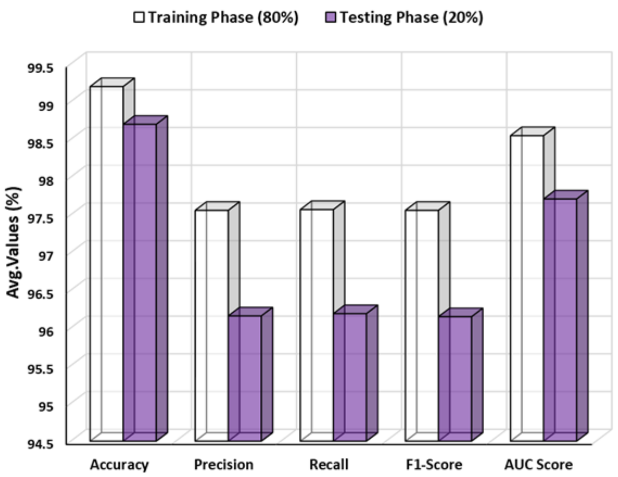

Figure 5. Average values of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model under 80% and 20%

Table 3 and Figure 5 depict the activity detection of the HNN-HSADC model under 80% and 20%. On 80% training, the proposed HNN-HSADC method obtains an average $accur_y$, $preci_n$, $recal_l$, $F1_{score}$, and AUC_{score} of 99.20%, 97.56%, 97.57%, 97.56%, and 98.55%, correspondingly. Also, under 20% testing, the proposed HNN-HSADC model attains an average $accur_y$, $preci_n$, $recal_l$, $F1_{score}$, and AUC_{score} of 98.70%, 96.16%, 96.19%, 96.15%, and 97.71%, respectively.

The training (TRAN) and validation (VALD) accuracy of the model HNN-HSADC obtained over 200 epochs are shown in Figure 6. Both curves are converging slowly and steadily, meaning that the model is learning effectively. The VALD accuracy is always slightly better than the TRAN accuracy, which suggests little overfitting, meaning good generalisation to the test set. The slightly different accuracy results are reasonable considering the difficulty of the task, but the general upward trend indicates increased accuracy and stability of the model.

In this, the HNN-HSADC model is trained with 80:20 over 200 epochs, with loss shown in Figure 7, which exemplifies the TRAIN and VALID loss. Both curves show a constant downward trend, indicating that the model is effectively reducing error during learning. Loss on VALD is slightly higher than the training loss in the majority of epochs, suggesting good generalization and no overfitting phenomenon. Though it experiences some fluctuations, which are monitored, it is gradually becoming steady and consistent.

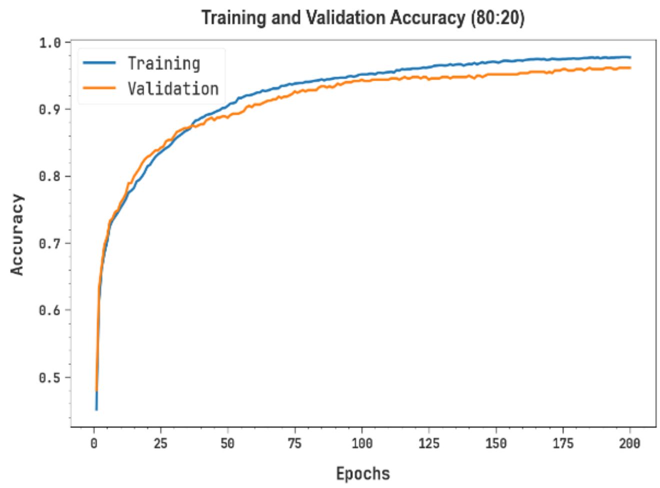

Figure 6. Accuracy curve of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) algorithm under 80:20

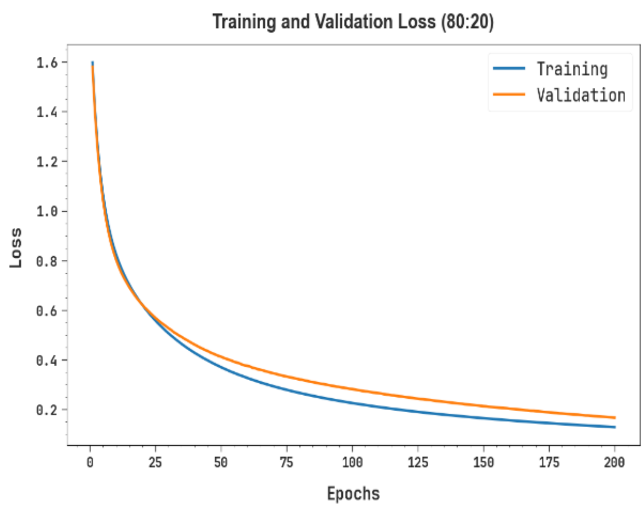

Figure 7. Loss curve of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) algorithm under 80:20

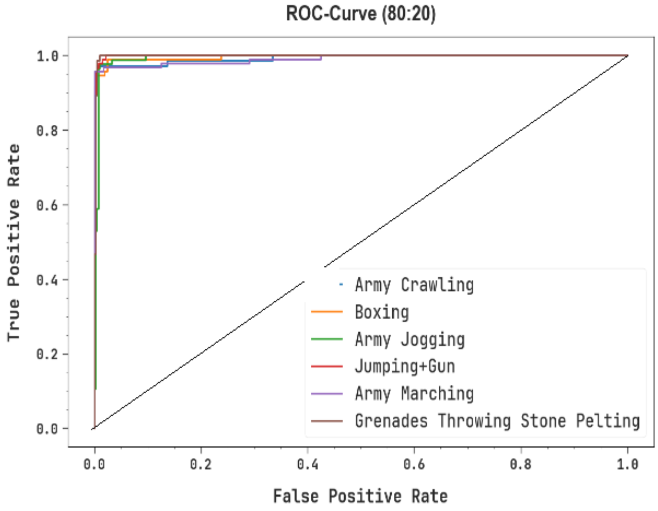

Figure 8. Receiver Operating Characteristic (ROC) curve of the Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) algorithm under 80:20

The Receiver Operating Characteristic (ROC) graph of the HNN-HSADC technique under 80:20 is discussed in Figure 8. The result indicates that the HNN-HSADC model achieves better ROC results for all classes, indicating a significant power to distinguish between the classes. This consistent improvement of the ROC values with multiple classes suggests the excellent performance of the HNN-HSADC model in the classification process, with greater efficacy in predicting classes.

Figure 9 shows the confusion matrices created by the HNN-HSADC model below 70:30 of training and testing phases. The result indicates that the HNN-HSADC technique has effective classification and detection of each class accurately.

Confusion Matrix – Training Phase (70%)

Actual / Predicted	Army Crawling	Boxing	Army Jogging	Jumping+ GIM	Army Marching	Grenades Throwing Stone Pelting
Army Crawling	267	2	3	2	3	3
Boxing	2	281	2	1	2	2
Army Jogging	1	1	251	6	11	3
Jumping+ GIM	2	6	3	267	4	4
Army Marching	2	0	6	4	335	4
Grenades Throwing Stone Pelting	5	2	1	1	0	302

(a)

Confusion Matrix – Testing Phase (30%)

Actual / Predicted	Army Crawling	Boxing	Army Jogging	Jumping+ GIM	Army Marching	Grenades Throwing Stone Pelting
Army Crawling	116	1	1	1	1	0
Boxing	0	127	1	0	0	1
Army Jogging	1	2	124	4	5	1
Jumping+ GIM	2	1	1	105	2	3
Army Marching	1	1	3	1	129	4
Grenades Throwing Stone Pelting	0	1	0	3	2	123

(b)

Figure 9. Confusion matrices of (a) 70% training phase and (b) 30% testing phase

Table 4 and Figure 10 depict the activity detection of the HNN-HSADC model under 70% and 30%. On 70% training, the proposed HNN-HSADC method reaches an average $accur_y$, $preci_n$, $recal_l$, $F1_{score}$, and AUC_{score} of 98.34%, 95.05%, 94.96%, 95.00%, and 96.98%, correspondingly. Likewise, under 30% testing, the proposed HNN-HSADC approach accomplishes an average $accur_y$, $preci_n$, $recal_l$, $F1_{score}$, and AUC_{score} of 98.09%, 94.27%, 94.31%, 94.28%, and 96.58%, respectively.

Figure 11 shows the accuracy of the HNN-HSADC model in 200 epochs for TRAN and VALD. Both curves are monotonously increasing and slowly merging, suggesting a good learning rate. The VALD accuracy constantly remains slightly higher than the TRAN accuracy, implying that the model is not overfitting and generalises well to unseen data. The fluctuations in accuracy would be expected due to the complexity of the task, but the overall increase in accuracy is indicative of greater accuracy and stability of the model.

Table 4. Activities detection of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model under 70% and 30%

Class Labels	$Accur_y$	$Preci_n$	$Recal_l$	$F1_{Score}$	AUC_{Score}
Training Phase (70%)					
Army Crawling	98.55	95.36	95.36	95.36	97.25
Boxing	98.83	96.23	96.56	96.40	97.92
Army Jogging	97.94	94.36	91.94	93.14	95.48
Jumping + Gun	98.16	95.02	93.36	94.18	96.21
Army Marching	97.99	94.37	95.44	94.90	97.03
Grenades Throwing Stone Pelting	98.60	94.97	97.11	96.03	98.01
Average	**98.34**	**95.05**	**94.96**	**95.00**	**96.98**
Testing Phase (30%)					
Army Crawling	98.96	96.67	96.67	96.67	98.02
Boxing	98.96	95.49	98.45	96.95	98.76
Army Jogging	97.53	95.38	90.51	92.88	94.78
Jumping + Gun	97.66	92.11	92.11	92.11	95.36
Army Marching	97.40	92.81	92.81	92.81	95.61
Grenades Throwing Stone Pelting	98.05	93.18	95.35	94.25	96.97
Average	**98.09**	**94.27**	**94.31**	**94.28**	**96.58**

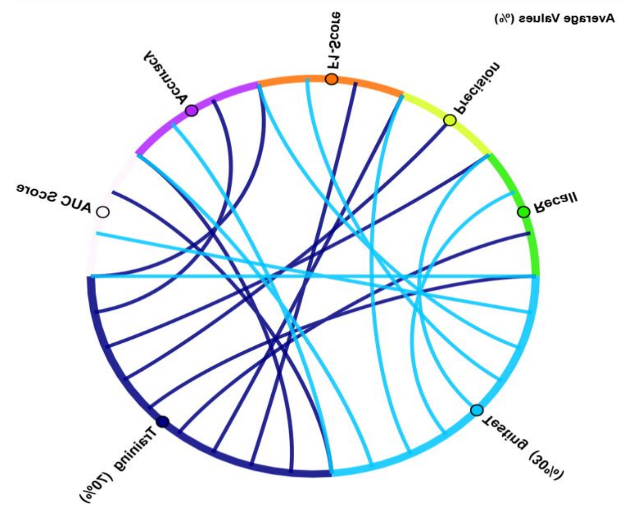

Figure 10. Average values of the Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model under 70% and 30%

Figure 12 shows the loss of the technique HNN-HSADC for the TRAN model and VALD for 70:30 for 200 epochs. Both curves show a steady decrease, suggesting that the model is able to reduce the error across learning. The VALD loss stays slightly lower than the training loss through most epochs, implying good generalization and no signs of overfitting. Some of these fluctuations are, however, monitored and stabilize and become more constant.

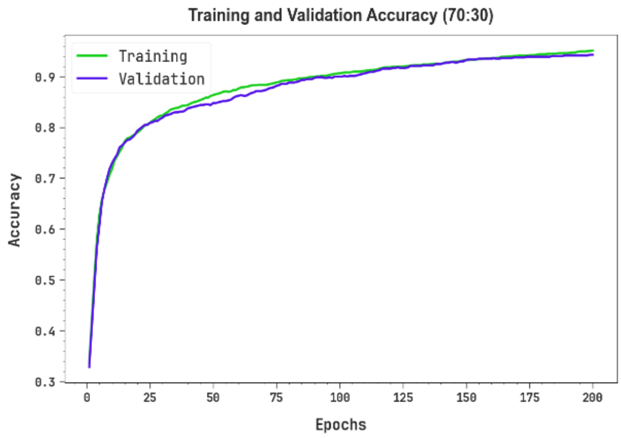

Figure 11. Accuracy curve of the Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) algorithm under 70:30

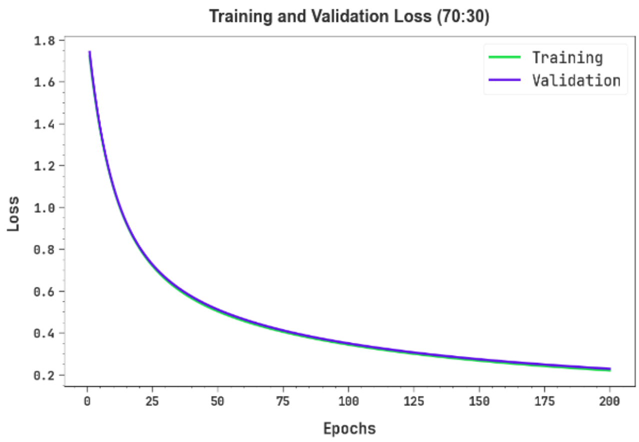

Figure 12. Loss curve of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) algorithm under 70:30

In Figure 13, the ROC curve of the HNN-HSADC model under 70:30 is examined. The outcomes indicate that the HNN-HSADC technique achieves improved ROC outcomes across all classes, indicating an important ability to discriminate between the classes. This persistent tendency of enhanced ROC values over several classes implies the efficacious operation of the HNN-HSADC model in predicting classes, emphasizing the stronger nature under the classification procedure.

Table 5 and Figure 14 signify the comparative analysis of the HNN-HSADC model with existing techniques under different measures [22]. This research emphasises that the proposed HNN-HSADC model has achieved higher performance. Based on $accur_y$, the presented HNN-HSADC method has attained superior $accur_y$ of 99.20% whereas the SVM7, ShuffleNet41, CrossViT44, VGG-19, DenseNet-201, BlazeFace, and 1.0 MobileNet-224 technologies have obtained lower $accur_y$ of 59.10%, 88.60%, 87.50%, 93.00%, 84.00%, 99.22%, and 94.67%, correspondingly. Also, based on $preci_n$, the proposed HNN-HSADC model has reached a superior $preci_n$ of 97.56%, while the SVM7, ShuffleNet41, CrossViT44, VGG-19, DenseNet-201, BlazeFace, and 1.0 MobileNet-224 approaches have inferior $preci_n$ of 94.62%, 70.20%, 71.65%, 76.09%, 96.18%, 95.67%, and 96.74%, correspondingly. In addition, based on $F1_{Score}$, the proposed HNN-HSADC model has obtained greater $F1_{Score}$ of 97.56%, while the SVM7, ShuffleNet41, CrossViT44, VGG-19, DenseNet-201, BlazeFace, and 1.0 MobileNet-224 approaches have got lesser $F1_{Score}$ of 90.58%, 92.38%, 81.84%, 80.91%, 79.72%, 93.31%, and 87.68%, correspondingly.

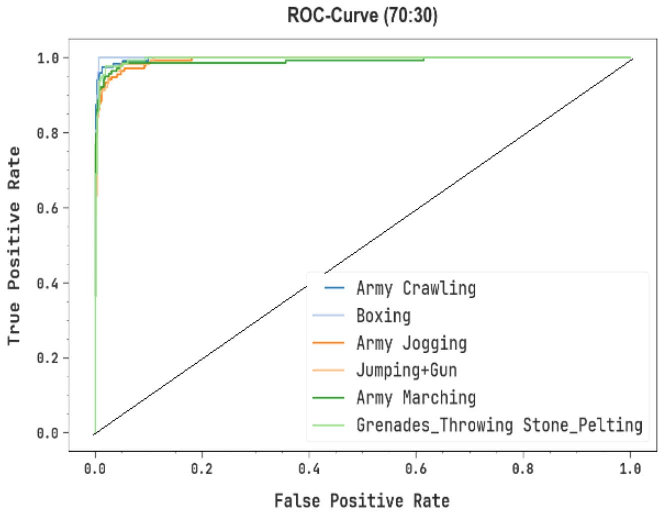

Figure 13. Receiver Operating Characteristic (ROC) curve of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) algorithm under 70:30

Table 5. Comparative analysis of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model with existing methodologies

Networks	$Accur_y$	$Preci_n$	$Recal_l$	$F1_{Score}$
SVM7	59.10	94.62	70.81	90.58
ShuffleNet41	88.60	70.20	71.62	92.38
CrossViT44	87.50	71.65	81.92	81.84
VGG-19	93.00	76.09	83.58	80.91
DenseNet-201	84.00	96.18	94.02	79.72
BlazeFace	99.22	95.67	69.65	93.31
1.0 MobileNet-224	94.67	96.74	95.69	87.68
HNN-HSADC	99.20	97.56	97.57	97.56

The strong performance of the HNN-HSADC framework is primarily due to the synergistic effect of EfficientNet-B6 and CNN-BiGRU. The compound scaling and MBConv layers are able to effectively capture the discriminative spatial representations, while the Bi-GRU captures the temporal dependence that may exist in the radar m-D signature.

Table 6 and Figure 15 represent the execution time (ET) of the HNN-HSADC technique. The existing models such as SVM7, CrossViT44, VGG-19, DenseNet-201, and 1.0 MobileNet-224 has obtained higher ET of 19.88sec, 16.55sec, 20.94sec, 20.24sec, and 14.55sec, correspondingly. Meanwhile, the ShuffleNet41 and ShuffleNet41 methods have got somewhat lower ET of 7.74sec and 6.23sec, respectively. Therefore, the proposed HNN-HSADC method has a lower ET of 3.67sec.

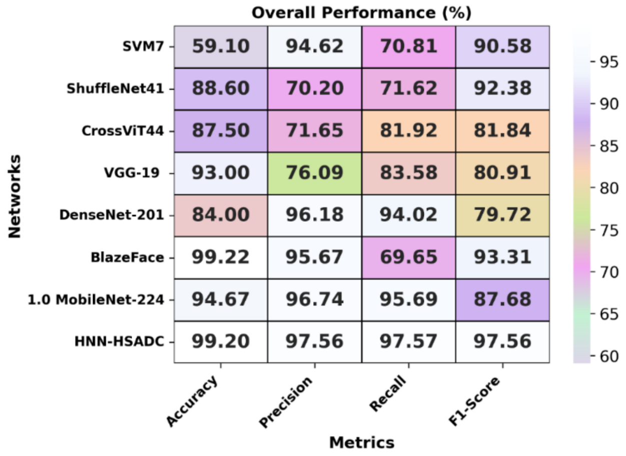

Figure 14. Comparative analysis of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model with existing methodologies

Table 6. Execution time of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model with existing systems

Networks	Execution Time (sec)
SVM7	19.88
ShuffleNet41	7.74
CrossViT44	16.55
VGG-19	20.94
DenseNet-201	20.24
BlazeFace	6.23
1.0 MobileNet-224	14.55
HNN-HSADC	3.67

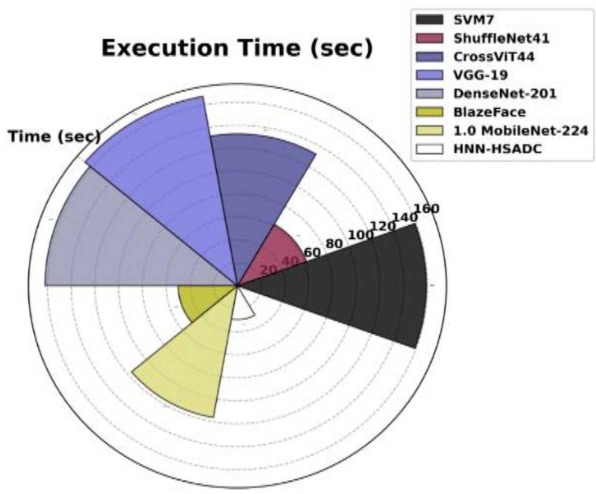

Figure 15. Execution time of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model with existing systems

The proposed HNN-HSADC performed well, regardless of the training/testing split ratio and the evaluation measure used. The stability shown in the confusion matrices, ROC curves, and comparative analyses confirms the robustness and reliability of the proposed method for suspicious HAR based on radar m-D images.

Table 7 offers the computational efficiency of the HNN-HSADC technique with several architectures based on FLOPs, inference time (IT), and Graphical Processing Unit (GPU). The HNN-HSADC method proves higher efficacy with lower FLOPs of 0.68 G, fastest IT of 0.076 ms, and the highest GPU of 9003 M. Conventional techniques like Hidden Markvov Model (HMM) and DeiT also display relatively less IT of 0.72 ms and 1.40 ms, respectively, but vary in computational load. In contrast, the existing approaches, such as LSTM, GRU, and LSTM-BiLSTM, require considerably more FLOPs and display longer ITs, signifying superior computational complexity. The Stack3-LSTM reveals very high FLOPs of 446.47 G and moderate inference time, implying a heavily parallelised architecture. Overall, the outcomes emphasise that HNN-HSADC obtains higher efficiency, making it more appropriate for human-suspicious action detection using radar m-D imaging.

Table 7. Computational efficiency outcome of Hybrid Neural Network-Based Human Suspicious Activity Detection and Classification (HNN-HSADC) model

Networks	FLOPs	Inference Time	GPU
HMM	274.6 G	0.72 ms	2049 M
LSTM	7.88 G	38.48 ms	2872 M
GRU	5.91 G	36.43 ms	3445 M
DeiT	1.08 G	1.40 ms	2577 M
LSTM-BiLSTM	10.6 G	32.30 ms	3534 M
Stack3-LSTM	446.47 G	5.17 ms	3642 M
Mobile-RadarNet	3.11 G	2.61 ms	2205 M
HNN-HSADC	0.68 G	0.076 ms	9003 M

Note: HMM = Hidden Markvov Model; LSTM = long short-term memory; GRU = Gated Recurrent Unit; BiLSTM = Bidirectional long short-term memory; GPU = Graphical Processing Unit.

5. CONCLUSION

This manuscript proposed an HNN-HSADC technique for accurately identifying and classifying human activities utilizing radar m-D images for defense applications. To achieve this, the presented HNN-HSADC system initially pre-processed the input images employing noise removal, contrast enhancement, image resizing, and normalization to enhance the image quality for subsequent analysis. For the feature representation process, the EfficientNet-B6 is leveraged to extract beneficial features from pre-processed radar images. Lastly, the CNN-BiGRU is applied for precise classification of human activities. Extensive simulation analyses were conducted to guarantee an improved performance of the HNN-HSADC approach. The comparative analysis demonstrated the superiority of the HNN-HSADC model over other state-of-the-art models concerning diverse evaluation metrics. The proposed framework of HNN-HSADC led to better classification results, but there are still some limitations. The DIAT-μRadHAR dataset is a small collection of action categories in specific environmental conditions, and this can limit the generalization of the system in very dynamic scenarios. Moreover, the training of EfficientNet-B6 requires more computational resources and memory during training. The framework can be refined in the future with larger multi-environment radar datasets, lightweight architectures for real-time applications, as well as cross-domain adaptation for various radar operating conditions.

DATA AVAILABILITY

The data that support the findings of this study are openly available in the Kaggle repository at https://ieee-dataport.org/documents/diat-mradhar-radar-micro-doppler-signature-dataset-human-suspicious-activity-recognition#files [20].

REFERENCES

[1] Tan, T.H., Tian, J.H., Sharma, A.K., Liu, S.H., Huang, Y.F. (2024). Human activity recognition based on deep learning and micro-Doppler radar data. Sensors, 24(8): 2530. https://doi.org/10.3390/s24082530

[2] Ahmed, S., Cho, S.H. (2023). Machine learning for healthcare radars: Recent progresses in human vital sign measurement and activity recognition. IEEE Communications Surveys & Tutorials, 26(1): 461-495. https://doi.org/10.1109/COMST.2023.3334269

[3] Chakraborty, M., Kumawat, H.C., Dhavale, S.V., Raj, A.A.B. (2022). DIAT-μ RadHAR (micro-Doppler signature dataset) & μ RadNet (a lightweight DCNN)-For human suspicious activity recognition. IEEE Sensors Journal, 22(7): 6851-6858. https://doi.org/10.1109/JSEN.2022.3151943

[4] Papadopoulos, K., Jelali, M. (2023). A comparative study on recent progress of machine learning-based human activity recognition with radar. Applied Sciences, 13(23): 12728. https://doi.org/10.3390/app132312728

[5] Li, X., He, Y., Fioranelli, F., Jing, X. (2021). Semi-supervised human activity recognition with radar micro-Doppler signatures. IEEE Transactions on Geoscience and Remote Sensing, 60: 1-12. https://doi.org/10.1109/TGRS.2021.3090106

[6] Waghumbare, A., Singh, U., Singhal, N. (2022). DCNN-based human activity recognition using micro-Doppler signatures. In 2022 IEEE Bombay Section Signature Conference (IBSSC), Mumbai, India, pp. 1-6. https://doi.org/10.1109/IBSSC56953.2022.10037310

[7] Huan, S., Wang, Z., Wang, X., Wu, L., Yang, X., Huang, H., Dai, G.E. (2023). A lightweight hybrid vision transformer network for radar-based human activity recognition. Scientific Reports, 13(1): 17996. https://doi.org/10.1038/s41598-023-45149-5

[8] Yousaf, J., Yakoub, S., Karkanawi, S., Hassan, T., Almajali, E., Zia, H., Ghazal, M. (2024). Through-the-wall human activity recognition using radar technologies: A review. IEEE Open Journal of Antennas and Propagation, 5(6), 1815-1837. https://doi.org/10.1109/OJAP.2024.3459045

[9] Triani, L.R., Adiono, T., Haar, S., Constandinou, T., Ahmadi, N. (2024). Radar-based human activity recognition using optimised CNN on edge devices. In 2024 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Penang, Malaysia, pp. 284-288. https://doi.org/10.1109/IECBES61011.2024.10990893

[10] Pareek, A., Kadian, P., Arora, S.M., Verma, J., Kumar, R., Khullar, V., Popli, R. (2024). Enhanced radar-based human suspicious activity classification using transfer learning methods. AIP Conference Proceedings, 3209(1): 050003. https://doi.org/10.1063/5.0228505

[11] Waghumbare, A.A., Singh, U., Raj, A.A.B., Singhal, N. (2025). A lightweight DCNN algorithm for radar-based suspicious human activities classification with data augmentation techniques. IEEE Aerospace and Electronic Systems Magazine, 40(2): 4-18. https://doi.org/10.1109/MAES.2025.3581164

[12] Wang, Y., Wang, Y., Xu, C., Yao, S., Wu, Q. (2025). SelaFD: Seamless adaptation of vision transformer fine-tuning for radar-based human activity recognition. In ICASSP 2025 IEEE International Conference on Acoustics, Speech and Signal Processing, Hyderabad, India, pp. 1-5. https://doi.org/10.1109/ICASSP49660.2025.10888271

[13] Nguyen, N., Doan, V.S., Pham, M., Le, V. (2024). SRCNN: Stacked-residual convolutional neural network for improving human activity classification based on micro-Doppler signatures of FMCW radar. Journal of Electromagnetic Engineering and Science, 24(4): 358-369. https://doi.org/10.26866/jees.2024.4.r.235

[14] Guendel, R.G., Kruse, N.C., Fioranelli, F., Yarovoy, A. (2024). Multipath exploitation for human activity recognition using a radar network. IEEE Transactions on Geoscience and Remote Sensing, 62: 1-13. https://doi.org/10.1109/TGRS.2024.3363631

[15] Zhou, J., Sun, C., Jang, K., Yang, S., Kim, Y. (2023). Human activity recognition based on continuous-wave radar and bidirectional gated recurrent unit. Electronics, 12(19): 4060. https://doi.org/10.3390/electronics12194060

[16] Nguyen, N., Pham, M., Doan, V.S., Le, V. (2024). Improving human activity classification based on micro-Doppler signatures of FMCW radar with the effect of noise. PLoS ONE, 19(8): e0308045. https://doi.org/10.1371/journal.pone.0308045

[17] Chakraborty, M., Kumawat, H.C., Dhavale, S.V. (2022). DIAT-RadHARNet: A lightweight DCNN for radar-based classification of human suspicious activities. IEEE Transactions on Instrumentation and Measurement, 71: 1-10. https://doi.org/10.1109/TIM.2022.3154832

[18] Thampy, P.B., Shailesh, S., MV, J., Kottayil, A. (2022). A convolutional neural network approach to Doppler spectra classification of 205 MHz radar. Theoretical and Applied Climatology, 149(3): 1769-1783. https://doi.org/10.1007/s00704-022-04126-0

[19] Tiwary, P.K., Johri, P., Katiyar, A., Chhipa, M.K. (2025). Deep learning-based MRI brain tumour segmentation with EfficientNet-enhanced UNet. IEEE Access. https://doi.org/10.1109/ACCESS.2025.3554405

[20] Fatima, T., Xia, K., Yang, W., Ul Ain, Q., Perera, P.L. (2025). Diabetes prediction using ADASYN-based data augmentation and CNN-BiGRU deep learning model. Computers, Materials & Continua, 84(1): 811. https://doi.org/10.32604/cmc.2025.063686

[21] Chakraborty, M., Kumawat, H.C., Dhavale, S.V. (2022). Application of DNN for radar micro-Doppler signature-based human suspicious activity recognition. Pattern Recognition Letters, 162: 1-6. https://doi.org/10.1016/j.patrec.2022.08.005

[22] Chakraborty, M., Kumawat, H.C., Dhavale, S.V., Raj, A.B. (2022). DIAT-μ RadHAR radar micro-Doppler signature dataset for human suspicious activity recognition. IEEE Dataport, 22(7): 6851-6858. https://doi.org/10.1109/JSEN.2022.3151943