



OPEN

Dual Attention-Based recurrent neural network and Two-Tier optimization algorithm for human activity recognition in individuals with disabilities

Hend Khalid Alqahtani¹✉, Gouse Pasha Mohammed² & Radwa Marzouk^{3,4}

Human activity recognition (HAR) has been one of the active research areas for the past two years for its vast applications in several fields like remote monitoring, gaming, health, security and surveillance, and human-computer interaction. Activity recognition can identify/detect current actions based on data from dissimilar sensors. Much work has been completed on HAR, and scholars have leveraged dissimilar methods, like wearable, object-tagged, and device-free, to detect human activities. The emergence of deep learning (DL) and machine learning (ML) methods has proven efficient for HAR. This research proposes a Dual Attention-Based Two-Tier Metaheuristic Optimization Algorithm for Human Activity Recognition with Disabilities (DATTMOA-HARD) model. The main intention of the DATTMOA-HARD model relies on improving HAR to assist disabled individuals. In the initial stage, the Z-score normalization converts input data into a beneficial format. Furthermore, the binary firefly algorithm (BFA) model is employed for feature selection. Moreover, the proposed DATTMOA-HARD model implements the dual attention bidirectional gated recurrent unit (DABiG) technique for the classification process. Finally, the Tasmanian devil optimizer (TDO)-based hyperparameter selection is accomplished to enhance the detection results of the DABiG model. The experimental evaluation of the DATTMOA-HARD approach is examined under the HAR dataset. The comparison analysis of the DATTMOA-HARD approach portrayed a superior accuracy value of 98.66% over existing methods.

Keywords Two-Tier metaheuristic optimization algorithm, Human activity recognition, Disabilities, Feature selection, data normalization

HAR is vital in an individual's everyday routine because it can learn cutting-edge intelligence about human interference from raw sensor data¹. The HAR technology became prevalent in investigations targeted abroad and at home². Individuals can automatically categorize the kind of human motion and attain the data from everyday actions that present a source for another intellectual application³. Investigators have inquired about the role of diverse kinds of sensing technology in activity monitoring to enhance precision. HAR has become a complex field of research owing to the handiness of sensors in wearable gadgets such as smartwatches, smartphones, and more⁴. These consume less power and are cost-effective, comprising live cascading of time-series information. The HAR models might be classified into dual classes based on data resources: sensor- and visual-based⁵. Owing to the fast improvement of sensors and prevalent computing technologies, sensor-based HAR became continuously more prevalent, and it is extensively utilized with confidentiality being safely protected. Sensor-based HAR is generally used in smart gadgets since smartphones and their privacy are highly secured with the progression of prevalent sensor and computer automation⁶. As a wearable gadget, progressive smartphones have become extremely simple. Equipped with implanted sensors like gyroscopes, smartphones, accelerometers, ambient sensors, and Bluetooth, investigators can also analyze everyday lives⁷.

¹Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia. ²Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia.

³Department of Mathematics, Faculty of Science, Cairo University, Giza 12613, Egypt. ⁴King Salman Center for Disability Research, Riyadh 11614, Saudi Arabia. ✉email: hkAlqahtani@pnu.edu.sa

Over recent years, investigators have organized many studies to discover multiple sensing technologies and techniques projected for recognizing and modelling human activities⁸. Sensor-based HAR on a gadget might be deliberated as an ML technique intended to continually track the user's actions, although it is related to a person's body⁹. Conventional models have made significant strides in the performance of advanced DL models, comprising Support Vector Machine (SVM), Decision Tree (DT), Artificial Neural Networks (ANN), and Naive Bayes (NB). However, these conventional ML models might finally aim at heuristic, handcrafted feature extractors usually subject to professional background¹⁰. Nevertheless, the efficacy of traditional ML approaches in categorizing precision and other measurements is limited. HAR significantly enhances the quality of life, particularly for individuals with disabilities, by enabling better monitoring and assistance in daily tasks. With the widespread utilization of wearable devices equipped with various sensors, there is an opportunity to combine rich data for accurate and real-time activity analysis. There still exists a challenge due to noisy data, variability in human behaviour, and the complexity of interpreting sensor signals effectively. Reliable and adaptive systems may be achieved by addressing these challenges, which assists personalized healthcare and assistive technologies. Therefore, developing robust models that can effectually process and interpret sensor data is crucial for advancing HAR applications.

This research proposes a Dual Attention-Based Two-Tier Metaheuristic Optimization Algorithm for Human Activity Recognition with Disabilities (DATTMOA-HARD) model. The main intention of the DATTMOA-HARD model relies on improving HAR to assist disabled individuals. In the initial stage, the Z-score normalization converts input data into a beneficial format. Furthermore, the binary firefly algorithm (BFA) model is used for feature selection. Moreover, the proposed DATTMOA-HARD model implements the dual attention bidirectional gated recurrent unit (DABiG) technique for the classification process. Finally, the Tasmanian devil optimizer (TDO)-based hyperparameter selection is accomplished to enhance the detection results of the DABiG model. The experimental evaluation of the DATTMOA-HARD approach is examined under the HAR dataset. The key contribution of the DATTMOA-HARD approach is listed below.

- The DATTMOA-HARD model applies Z-score normalization to standardize input features, converting them to have zero mean and unit variance. This step enhances model convergence, stability, and performance and mitigates the influence of scale differences across features.
- The DATTMOA-HARD method employs the BFA technique to select the most relevant features, effectively mitigating dimensionality. This approach enhances computational efficiency and accelerates training time. Concentrating on crucial data attributes also assists in improving overall accuracy.
- The DATTMOA-HARD technique utilizes a DABiG methodology that integrates spatial and temporal attention mechanisms to capture intrinsic dependencies. This improves sequential data's accuracy and contextual comprehension and enhances the model's capability to concentrate on crucial features over time for better classification performance.
- The DATTMOA-HARD methodology implements the TDO technique to fine-tune hyperparameters efficiently, ensuring optimal classification performance. This optimization improves the model's adaptability to diverse datasets, mitigates training time, and improves overall accuracy.
- Integrating a dual attention mechanism within a Bi-GRU architecture, integrated with the BFA for feature selection and optimized by the TDO, presents a novel HAR framework. This framework uniquely balances accuracy, efficiency, and adaptability by capturing spatial-temporal dependencies and selecting relevant features. Its optimization approach ensures superior performance across various datasets. Overall, the model outperformed sequential models in both precision and computational cost.

Related works

Khan et al.¹¹ proposed an ensemble model of LSTM and 1D-CNN for postural transition recognition caused by wearable sensors and wireless computing. Attracting pervasive learning will eventually induce the formation of adjustable gadgets allowed by multiple data analysis and relational learning models. This technique might be integrated to enable seamless learning and obtain co-relations with adaptive learning models. Duhayim¹² introduces an Improved Pelican Optimizer with Deep Transfer Learning enabled HAR (IPODTLHAR) method for impaired individuals. The IPODTL-HAR model's main objective was to identify impaired people's human activities. The projected method monitors data preprocessing to improve the data quality. Moreover, the EfficientNet approach originates valuable vector features, and the hyper-parameters are modified using Nadam optimization. Eventually, the IPO will employ DBN methods to classify and recognize human actions. It aids in effectually tuning the hyper-parameters connected to the DBN and EfficientNet techniques correspondingly. Serpush et al.¹³ introduced a specialized solution by developing an FSVM and DeepCNN to remove substantial aspects and label activities utilizing an FMF. By employing FSVM with FMF, the problem of transferring samples to the proper class is resolved, inducing more precise categorization and an adaptable technique. Dastbaravardeh et al.¹⁴ developed an enhanced approach to identify human actions in lower-size and lower-resolution videos by utilizing an autoencoder (AE) and CNN with a channel attention mechanism (CAM). By improving blocks with more demonstrative aspects, convolution layers remove discriminating aspects from several systems. In addition, random sampling of frames is utilized to enhance precision while leveraging less information. The objective is to raise implementations and overcome challenges like computational complexity, overfitting, and ambiguity by employing AE and CNN-CAM. Cob-Parro et al.¹⁵ projected a model for individual and action identification in the wild. HAR is accomplished by utilizing RNN, particularly an LSTM. The LSTM input is an ad-hoc, lightweight vector feature from the bounding box. Alotaibi et al.¹⁶ develop a novel arithmetic optimization algorithm with a long short-term memory autoencoder (AOA-LSTMAE) method for the HAR model in an IoT setting. During this projected model, the main goal is to identify multiple kinds of human actions in the IoT setting. The model primarily originates the P-ResNet method for extracting features to attain this. Furthermore,

the projected model leverages the LSTMAE classification technique to identify diverse actions. To enhance the effectiveness of the LSTMAE technique, AOA is utilized as a hyper-parameter optimization method.

In¹⁷, the relationship among properties and human activities of Wi-Fi CSI signals is examined on several receiving antennas and the signal assets that differ unusually in response to human actions. The dissimilarity between several antennas illustrates several sensitivities to human actions, instantly influencing detection performance. Consequently, an adaptive antenna extraction model that mechanically extracts the non-sensitive antenna is presented to identify human actions with greater efficacy. Sinha and Kumar¹⁸ developed an efficient ensemble-based framework for HAR using advanced preprocessing, feature extraction, and optimized classification techniques. Zhou et al.¹⁹ presented efficient, lightweight DL techniques for HAR on resource-constrained edge devices using TinyML. Sinha, Kumar, and Ghosh²⁰ proposed a four-stage framework for accurate HAR using advanced preprocessing, feature extraction, and a fusion of Bi-GRU and bidirectional long short-term memory (Bi-LSTM) classifiers, improved by the dempster-shafer theory (DST). Rafee, Zishan, and Noor²¹ aimed to enable real-time HAR and fall detection (FD) by utilizing an optimized DT and random forest (RF) models on low-power microcontrollers such as Arduino UNO and ESP32 for efficient on-device processing. Chen et al.²² proposed the multiple spectrogram fusion network (MSF-Net), a robust DL method using multiple spectrogram fusion and Transformer architecture, to enhance WiFi-based HAR despite environmental interference. Zhou et al.²³ proposed a covariance-based graph convolutional network (CovGCN) with a topology refinement module (CovTRM) to improve hand gesture recognition utilizing surface electromyography by dynamically learning muscle network topologies. Sinha and Kumar²⁴ improved HAR from UAV videos using a novel diminutive multi-dimensional locality coding-based convolutional neural network (DMLC-CNN) methodology with advanced segmentation, feature extraction, and encoding techniques. Ghous, Najam, and Jalal²⁵ aimed to enhance emotion detection in individuals with cognitive disabilities using EEG data and a multi-class SVM integrated with an advanced feature selection method. Bollampally et al.²⁶ improved HAR accuracy and efficiency using a BiLSTM model on the PAMAP2 dataset. Zhou et al.²⁷ proposed the cross-scale shuffle attention you only look once (CSSA-YOLO) network to improve classroom behaviour recognition by addressing occlusion and complex backgrounds through advanced feature optimization and attention mechanisms.

Despite improvements in HAR, the utilized studies face difficulty in several scenarios. Several methods encounter challenges with real-time deployment on resource-constrained devices, restricting their applicability in edge environments. Models depending on visual sensors are affected mainly by occlusion, background complexity, and lighting discrepancies, which mitigate detection accuracy. Techniques utilizing sEMG or EEG data encounter threats in capturing reliable topologies and subtle signal discrepancies. Moreover, AOA and IPO optimization, n techniques lack g, generalization across diverse HAR scenarios. Few approaches rely on heavy models such as CNN and LSTM, which are ineffective for low-power devices. A gap exists in developing lightweight, adaptive, and robust models that can effectively handle noisy data, sensor variability, and dynamic activity contexts while ensuring scalability across diverse HAR settings.

Materials and methods

This manuscript proposes the DATTMOA-HARD technique. The main intention of the DATTMOA-HARD technique is to improve HAR to assist disabled individuals. To accomplish that, the DATTMOA-HARD model has data normalization, dimensionality reduction, HAR classification, and parameter selection defined in Fig. 1.

Data normalization: Z-score

Initially, the Z-score normalization is employed to convert input data into a beneficial format²⁸. This model is chosen for its ability to effectively standardize features by centring them around a mean of zero and scaling to unit variance, which assists in handling data with varying scales and distributions. The technique is more robust to extreme values, preserving the overall data structure. This methodology also enhances the convergence speed of various ML techniques, particularly those sensitive to feature scales, such as neural networks and gradient-based models. Moreover, Z-score normalization maintains the relative relationships between data points, which is significant for time-series or sequential data. Its simplicity and efficiency in mitigating bias due to scale differences make it an ideal choice over other normalization methods for this model.

Z-score normalization is a statistical technique for standardizing the features in HAR tasks, mainly when functioning with sensor data. As an outcome, the malformed data has a standard deviation of 1 and a mean of 0, making it simpler to equate diverse features with fluctuating measures. It generally aids in decreasing the impact of outliers and certifies that every feature donates similarly to the ML technique. It is particularly beneficial in HAR when sensor readings differ significantly across diverse devices or atmospheres. This normalization method boosts the convergence and performance of several ML models applied to HAR classification tasks.

Dimensionality reduction: BFA

Afterwards, the feature selection is performed by BFA²⁹. This method is chosen for its effectiveness in solving combinatorial optimization problems such as feature selection. Unlike conventional methods such as principal component analysis (PCA), which transform features into a new space and may lose interpretability, BFA directly chooses the most relevant original features, preserving their meaning. The model also effectively explores the solution space, avoiding local minima better than simpler heuristic methods. Furthermore, this technique effectively balances exploration and exploitation through firefly brightness and attractiveness, resulting in a robust selection of feature subsets that improve model accuracy while mitigating computational complexity. Its binary encoding suits feature selection tasks, making it an appropriate choice over other metaheuristic algorithms for this application. Figure 2 specifies the BFA method.

Like numerous biologically inspired models, the firefly algorithm (FA) uses observable behaviour from the real world to update heuristic methods to solve composite optimizer problems, ranging from continuous to

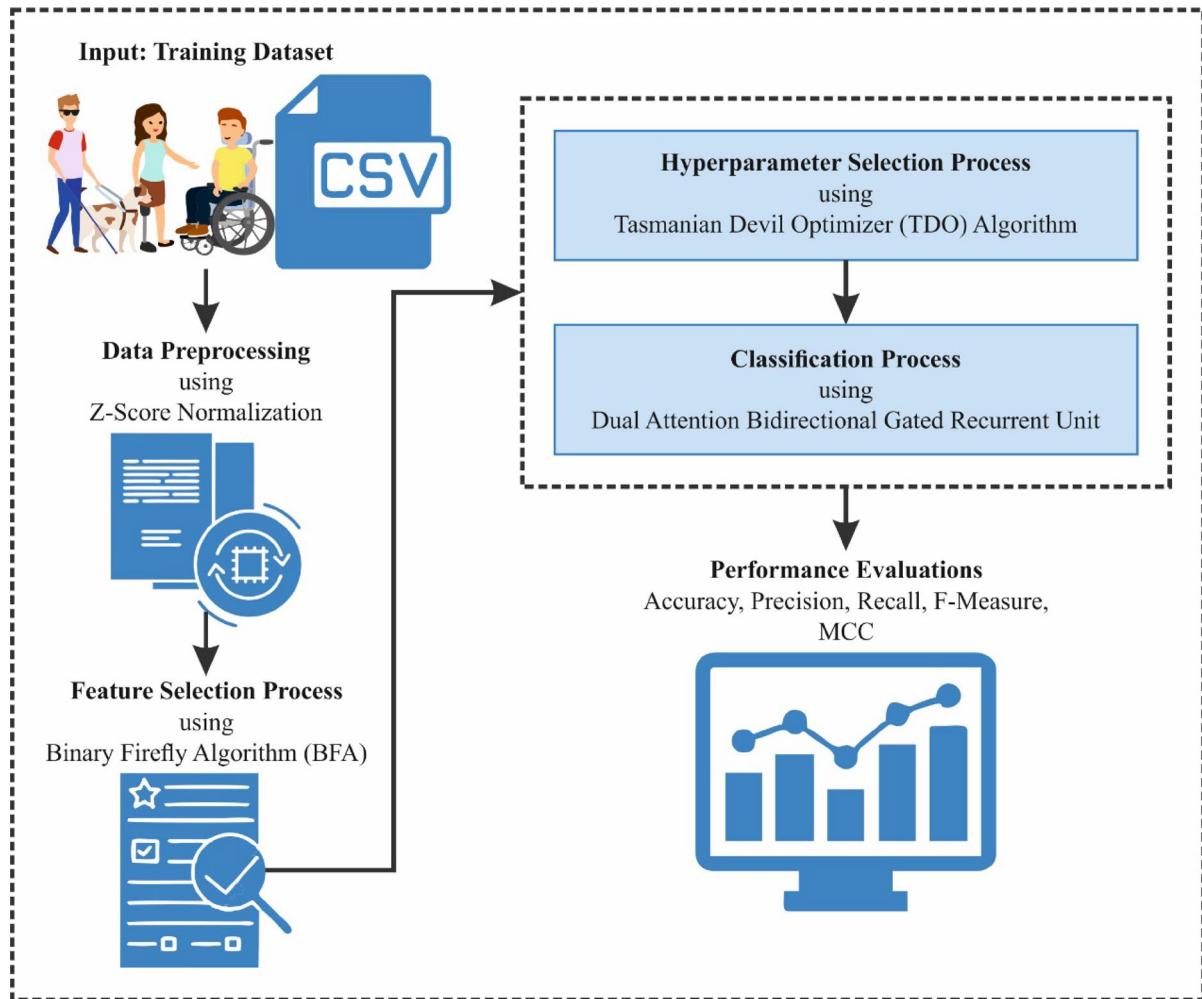


Fig. 1. Overall process of DATTMOA-HARD model.

separate fields. In feature selection (FS), all fireflies characterize particular combinations of chosen features from the dataset, usually signified as a binary vector. At the same time, individual bits specify the exclusion or inclusion of specific characteristics. As the iterative method advances, fireflies pass through the feature subsets space, consistently dropping towards structures with larger appraising scores. Concurrently, they present stochastic disturbances to guarantee a complete exploration of the searching area. This interaction of probability-based exploration and predetermined attraction allows subtle search approaches, permitting *FA* to detect an optimum or close-optimum feature subdivision, which improves prognostic precision and possibly dimensionality reduction. The major mathematical equation for the solution vector to the optimizer issue is represented as:

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{i,j}^2} (x_j^t - x_i^t) + \alpha \epsilon_i^t \quad (1)$$

Whereas α refers to the scaling feature that controls the arbitrary walk's step size, ϵ_i^t signifies a randomly generated vector number that is extracted from a Gaussian distribution at all iterations, γ stands for scale-dependent parameter controls the firefly's visibility, β_0 denotes attractiveness constant after the distance amongst two fireflies is 0. It has to be noted that the distance $r_{i,j}$ amongst i th and j th fireflies is described as their distance from Cartesian to eliminate some vagueness within greater sizes. Moreover, the *FA* is intended to optimize issues constantly. The problem of FS is stated as a boolean value array, so the FA's binary version is needed. This was achieved by combining a threshold-based (threshold = 0.5) binarization procedure while the fireflies' constant positional values are round and related beside the threshold to make binary decisions for FS. The fireflies' movement near optimistic individuals is attuned by the attractive feature (β) that reduces with distance and is additionally affected by randomization terms. The formulation is given below:

$$\beta = \beta_0 - \beta_{\min} * e^{-\gamma r^2} + \beta_{\min} \quad (2)$$



Fig. 2. Overall process of DATTMOA-HARD model.

Whereas β_{\min} refers to minimal attractiveness, this binary adaptation allows the model to carry out FS tasks by explaining all bits as the absence or presence of features instead of a location in a continuous area.

Therefore, the non-linear attraction mechanism of the FA permits shorter-distance attractions. In these sub-swarms, the arrival of the global solution is conceptually guaranteed. Moreover, provided the ability of the FA to work with numerous swarms, it is perfectly appropriate to deal with non-linear, multi-modal optimization challenges.

The fitness function (FF) reveals the accuracy of classification and the quantity of voted features. It maximizes the classification accuracy and decreases the dimension of the chosen features. Therefore, the FF mentioned below is applied for evaluating individual solutions, as set in Eq. (3).

$$\text{Fitness} = \alpha * \text{ErrorRate} + (1 - \alpha) * \frac{\#SF}{\#All_F} \quad (3)$$

ErrorRate is the classification rate of error. It is computed as the ratio of improper classification to many classifications set between 0 and 1. $\#SF$ signifies the number of chosen features; $\#All_F$ refers to the complete number of features. α is applied for controlling the import of classification excellence and sub-set length.

HAR classification: DABiG

Besides, the proposed DATTMOA-HARD model designs a DABiG technique for the classification process³⁰. This model is chosen for its efficiency in capturing spatial and temporal dependencies in sequential data. Unlike standard GRUs or LSTMs, the dual attention mechanism allows the model to concentrate selectively on crucial features across time and space, improving contextual understanding. This enhances recognition accuracy, especially in complex and noisy environments typical of HAR tasks. The bidirectional structure processes data in both forward and backward directions, capturing past and future context, which is significant for precise activity recognition. Compared to other models like CNNs or single-attention RNNs, DABiG presents a better computational efficiency and performance balance, making it appropriate for real-time HAR applications. Figure 3 portrays the infrastructure of DABiG.

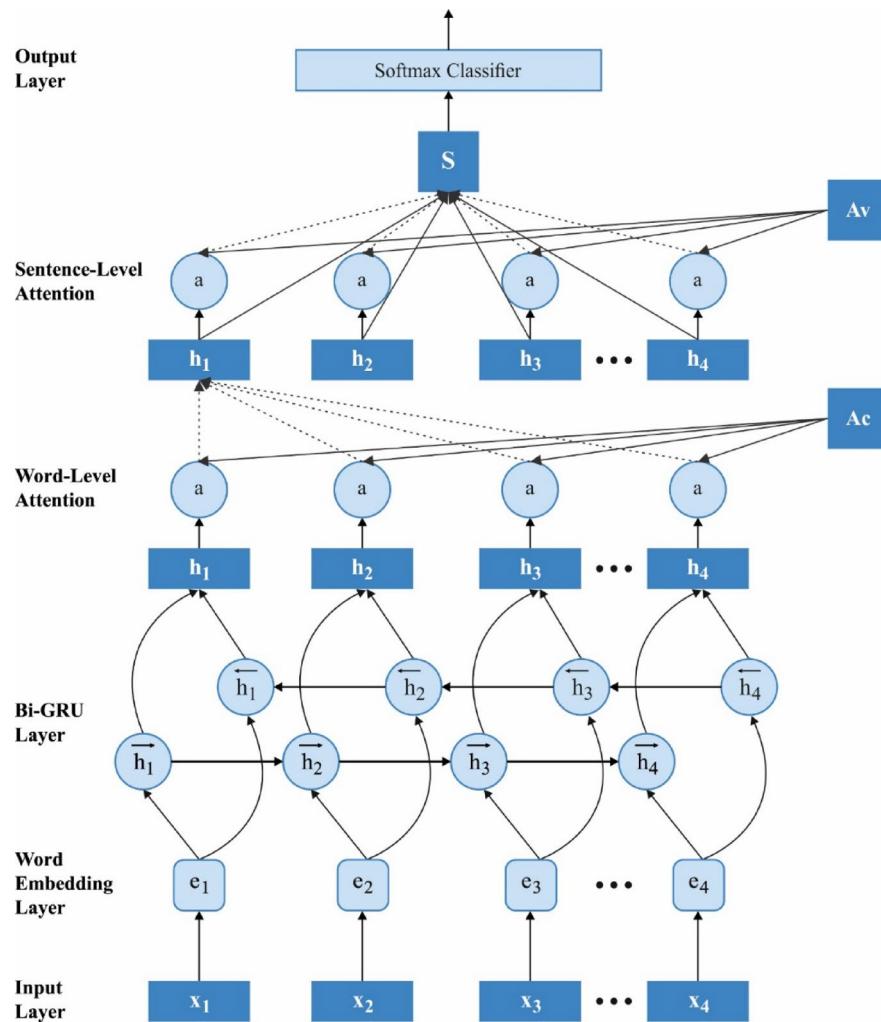


Fig. 3. Architecture of DABiG.

The Bi-GRU is a recurrent neural network (RNN) that processes sequential data bi-directionally. It contains dual layers of GRU; one processes the data in the forward direction and the other in the backward direction. GRUs are recognized for their capability to seize longer-term dependencies in sequential data while being computationally efficient. All GRU cells provide a hidden state, which encodes information from preceding time steps and utilizes it to upgrade its present state. Moreover, spatial attention concentrates on information from dissimilar spatial positions (for example, sensors on the abdomen, back, and chest), whereas data is gathered. The spatial attention mechanism calculates attention weights for all spatial positions, representing which sensors are most relevant for the classification task. These weights are frequently learned in training. The spatial attention weights combine the spatial information into a single, weighted representation. For every time step, the temporal attention mechanism calculates an attention or weight score, representing the significance of that time step's data for the others in the sequence. The temporal attention weights generate a weighted sequence representation, highlighting the most relevant time steps for classification.

Primarily, the candidate state \tilde{p}_T is assessed by considering the result of the reset gate u_T . The candidate estimation state \tilde{p}_T is written as:

$$\tilde{p}_T = \tanh(Q_p [u_T \times p_{T-1}, D_T]) \quad (4)$$

Now, u_T indicates the reset gate and is specified as:

$$u_T = \text{Sig}(Q_u [p_{T-1}, D_{T\tau}]) \quad (5)$$

By utilizing the reset gate, unsuitable details are eliminated, and then, the updated gate l_T controls the process as follows:

$$l_T = \text{Sig}(Q_l [p_{T-1}, D]) \quad (6)$$

At last, the hidden state outcome p_T is stated as:

$$p_T = (1 - l_T) \times p_{T-1} + l_T \times \tilde{p}_T \quad (7)$$

whereas the weight is described as Q , D_T stands for the input, and Sig means the sigmoid function. GRU networks are recognized for their capability to seize long-range dependencies in sequential data. Nevertheless, they can occasionally contend with particular patterns, which need an understanding of either past or future contexts. Bi-GRU alleviates this problem by using information from either direction, making it best appropriate to capture longer-term dependences. The result of the forward and backward hidden state result is delineated as follows:

$$\vec{p}_T = GRU(D_T, \vec{p}_{T-1}) \quad (8)$$

$$\overleftarrow{p}_T = GRU(D_T, \overleftarrow{p}_{T-1}) \quad (9)$$

$$p_T = A_T \vec{p}_T + B_T \overleftarrow{p}_T + f_T \quad (10)$$

Meanwhile, weights regarding the backward and forward hidden layer (HL) are correspondingly designated as A_T and B_T . The state of the HL in either the forward or backward handling is written as \vec{p}_T and \overleftarrow{p}_T correspondingly. At the time T , the state of the HL is designated as f_T . Formerly, the weighting mechanism was developed utilizing the double attention mechanism called spatial and temporal attention to enhance the classification precision.

Temporal Attention Module: A temporal attention mechanism allocates weights to all time steps inside the sequence. These weights characterize the relative significance of all sequence parts in making the last classification decision. After computing the attention weights, the method calculates the input sequence's weighting amount, multiplying all time steps by their equivalent attention weight.

$$G'_H = \text{softmax}(W_G X_H + Z_H) \quad (11)$$

$$Y_H'' = G'_H \Theta X_H = (G_{1,H} X_{1,H}, G_{2,H} X_{2,H}, \dots \dots .G_{T,H} X_{T,H}) \quad (12)$$

whereas, $G'_H = (G_{1,H}, G_{2,H}, \dots .G_{T,H})$ denotes the temporal weights of the attention layers, Z_H and W_G signify the bias and weight Y_H'' refers to temporal attention output and the dot product is specified as Θ .

Spatial Attention Module: The mechanisms allocate significance weights to dissimilar spatial features inside the input data. These weights characterize the relative importance of all features for the classification task. Features with greater weight are measured more significantly when making the last classification decision. After calculating the attention weights, the model typically carries out a weighted sum or a weighted collection of the input data according to these weights.

$$R'_H = \text{softmax}(W_R X_H + Z_H) \quad (13)$$

$$Y_H'' = R'_H \Theta X_H = (R_{1,H} X_{1,H}, R_{2,H} X_{2,H}, \dots \dots .R_{T,H} X_{T,H}) \quad (14)$$

Here, $R'_H = (R_{1,H}, R_{2,H}, \dots .R_{T,H})$ represents the weights of the spatial attention layer, Z_H and W_R signify the bias and weight and Y_H'' denotes spatial attention output.

Softmax Classification: At last, the softmax classification is developed to classify.

Parameter selection: TDO

Finally, the TDO-based hyperparameter range method is achieved to enhance the detection outcomes of the DABiG model^[31]. This method is chosen for its effectual global search capability and convergence behaviour inspired by the natural foraging strategy of Tasmanian devils. Compared to conventional optimization methods such as grid or random search, this model effectually balances exploration and exploitation, mitigating the risk of getting trapped in local optima. Its adaptive nature allows for dynamic adjustment of search parameters, resulting in faster and more precise tuning of hyperparameters. The model is effectual as it needs fewer iterations to find optimal solutions, thus saving computational resources. This makes it specifically advantageous for complex models like DABiG, where fine-tuning numerous parameters can be computationally expensive and time-consuming. Figure 4 indicates the flow of the TDO technique.

The TDO model considers other TD positions in the population promising corpse spots inside the search space. One such phenomenon, in which the i th TD selects the k th population member as the objective carrion, is demonstrated by Eq. (15) as an arbitrary selection. With the possible exemption of the number i , k is randomly selected from 1 to N to stop a TD from selecting itself as the goal.

$$C_i = x_{\kappa_t} \quad i = 1, 2, \dots, N \\ k \in \{1, 2, \dots, N | k \neq i\} \quad (15)$$

C_i specifies the carrion selected by the i th TD. After choosing the objective carrion, the TD has a novel location in the searching region. Based on this approach, the TD will adhere to the goal carrion during the simulated motion if its objective function is greater; if not, it will move out. Equation (16) characterizes this movement model. In this approach, the TD will preserve its present posture till it discovers a novel place, which gives an

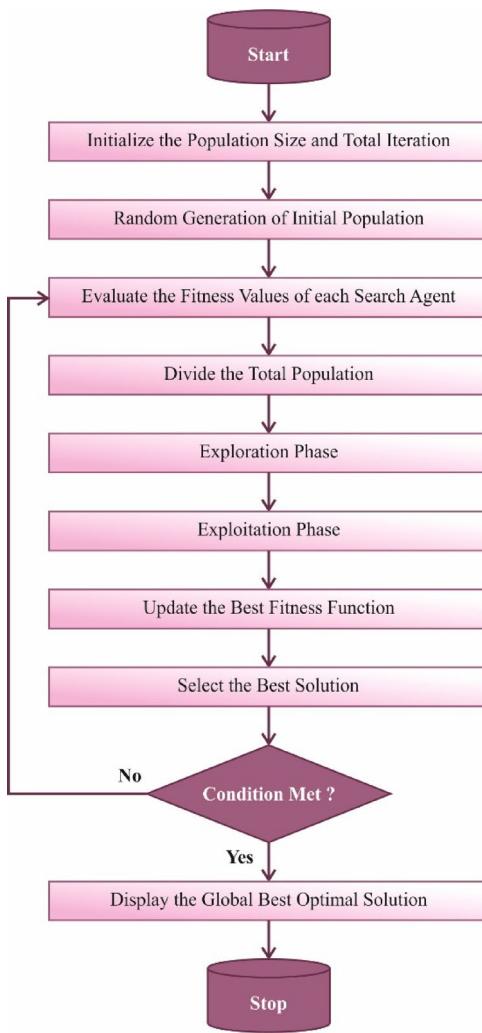


Fig. 4. Working flow of the TDO method.

improved objective function value; at this time, it will accept the novel position. Equation (17) describes how the upgrade is produced:

$$x_{i,j}^{new,S1} = \begin{cases} x_{i,j} + r.(c_{i,j} - l.x_{i,j})_t & F_{C_i} < F_i^* \\ x_{i,j} + r.(x_{i,j} - c_{i,j})_t & \text{otherwise} \end{cases} \quad (16)$$

$$X_i = \begin{cases} X_{i,j}^{new,S1} & F_{i,j}^{new,S1} < F_i \\ X_i, & \text{else} \end{cases} \quad (17)$$

r signifies the randomly generated numerical value inside the range [0,1]. F_{C_i} represents the function or objective value associated with the selected carrion i substitutes for a random integer, which may have two dissimilar values. The objective function value is characterized by $F_{i,j}^{new,S1}$, and the upgraded position of the i th TD following the application of the initial model is characterized by $X_{i,j}^{new,S1}$. After using Tactic 1, the j th variable value at the new location of the i th TD is characterized by $X_{i,j}^{new,S1}$.

Tactic 1: using prey as a nutrition source

The dual major phases of TD's second feeding approach are searching and consumption of prey. It finds the target and starts an attack primarily (initial phase). The prey is then followed and condensed immobile before consumption (second phase). Interestingly, this second approach's initial stage, which entails choosing a target carrion, is modelled comparably to the initial modelling approach. Equation (17) over Eq. (20) pretends the initial portion of the next model, in which the TD selects its and attacks. Utilizing this next model, another population member is considered food supplies and prey. The location of the TD is upgraded. Equation (17) demonstrates how prey is selected:

$$P_i = X_k, \quad i = 1, 2, \dots, N, \\ k \in \{1, 2, \dots, N | k \neq i\} \quad (18)$$

The number selected randomly inside the interval of $(1-N)$ is represented by k . P_i characterizes the prey selected by the i th TD. Next is victim identification, where the TD estimates its novel location. This position defines whether the present prey contains a lower or greater objective function value. The TD moves towards its prey when the conditions are correct; if not, it retreats. The TD can improve its search tactics by utilizing this approach. This procedure is validated in Eq. (19). The TD in the second model, Eq. (20), will only acknowledge the recently projected location if it increases the target function's value.

$$x_{i,j}^{new,S2} = \begin{cases} \begin{cases} x_{i,j} + r \bullet (p_{i,j} - I.x_{i,j}), & F_{pi} < F_i^* \\ x_{i,j} + r \bullet (x_{i,j} - p_{x_{i,j}}), & otherwise \end{cases} \end{cases} \quad (19)$$

$$x_{i,j}^{new,S2} = \begin{cases} \begin{cases} x_{i,j} + r \bullet (p_{i,j} - I.x_{i,j}), & F_{pi} < F_i^* \\ x_{i,j} + r \bullet (x_{i,j} - p_{x_{i,j}}), & otherwise \end{cases} \end{cases} \quad (20)$$

$x_{i,j}^{new,S2}$, the TD's i th present status is determined utilizing the second model. After using the second model, $x_{i,j}^{new,S2}$ characterizes the j th parameter value for the i th TD. After using the second model, the objective function value related to the i th TD is characterized as $F_i^{new,S2}$. F_{pi} presents the objective function value equivalent to the TD's chosen meal. The main change between the 1st and 2nd tactics is that the latter contains a prey-chasing stage, equivalent to a limited search within the search area. This exercise emphasizes the TD's capability to benefit from opportunities and move towards improved candidate solutions. The TD observes prey close to the attacking place to imitate this search. The numerical description of Eqs. (21) to (23).

$$R = 0.01 \left(1 - \frac{t}{T}\right) \quad (21)$$

$$X_{i,j}^{new} = x_{i,j} + (2r - 1) \bullet R.x_{i,j} \quad (22)$$

$$X_i = \begin{cases} X_i^{new}, F_i^{new} < F_i^* \\ X_i, otherwise \end{cases} \quad (23)$$

The iteration counter is characterized by t . X_i represents the maximal iterations, and the updated or novel location of the i th TD inside the area of its preceding location is signified by X_i^{new} . The objective function value related to the i th TD at its upgraded or novel position is considered in F_i^{new} . The sum of the j th parameter for the i th TD at its novel position is $X_{i,j}^{new}$. The neighbourhood's perimeter is characterized by the sign R that has the assault point of TD in its centre. After all iterations, the TDO model computes novel locations and objective function values for all members of the TD. The model upgrades the number of TDs up to the end condition before moving on to the following iteration.

Fitness choice is a substantial feature that manipulates the outcome of the TDO model. The selection of a hyperparameter procedure contains an encoded solution technique to assess the efficacy of the candidate solution. At this point, accuracy is considered by the TDO method as the leading standard to project FF. Its mathematical formulation is expressed below:

$$Fitness = \max (P) \quad (24)$$

$$P = \frac{TP}{TP + FP} \quad (25)$$

Here, TP signifies the positive value of true, and FP denotes the positive value of false.

Experimental validation

This study examines the performance assessment of the DATTMOA-HARD technique under the HAR dataset³². The dataset holds 30,000 samples, with six classes defined in Table 1. Five features are in total, four of which were selected.

Activity	No. of Instances
Walking	5000
Jogging	5000
Upstairs	5000
Downstairs	5000
Sitting	5000
Standing	5000
Total Instances	30,000

Table 1. Details of the dataset.

Figure 5 displays the classifier performances of the DATTMOA-HARD method on the test dataset. Figure 5a and b exemplifies the confusion matrix through precise classification and identification of all classes on a 70%:30% TRASE/TESSE. Figure 5c signifies the PR examination, which notified higher performance over all classes. Eventually, Fig. 5d demonstrates the ROC study, which signifies proficient solutions with great ROC values for different class labels.

The classification results of the DATTMOA-HARD technique are demonstrated in Table 2; Fig. 6. The experimental values infer the effectual capability of the DATTMOA-HARD technique on the recognition process. With 70%TRASE, the DATTMOA-HARD technique obtains an average $accu_y$ of 98.53%, $prec_n$ of 95.61%, $reca_l$ of 95.61%, $F_{measure}$ of 95.60%, and MCC of 94.73%. Moreover, using 30%TESSE, the DATTMOA-HARD method attains an average $accu_y$ of 98.66%, $prec_n$ of 95.98%, $reca_l$ of 95.97%, $F_{measure}$ of 95.97%, and MCC of 95.17%.

Figure 7 depicts the TRA $accu_y$ (TRAAY) and validation $accu_y$ (VLAAY) performances of the DATTMOA-HARD approach. The values of $accu_y$ are computed through a period of 0–25 epochs. The figure underscored that the values of TRAAY and VLAAY show an increasing trend, indicating the capacity of the DATTMOA-HARD technique with maximum performance through multiple repetitions. Moreover, the TRAAY and VLAAY values remain close across the epochs, notified lesser overfitting, and present superior performance of the DATTMOA-HARD technique, guaranteeing reliable calculation on unseen samples.

Figure 8 shows the TRA loss (TRALO) and VLA loss (VLALO) graph of the DATTMOA-HARD method. The loss values are computed across a period of 0–25 epochs. The values of TRALO and VLALO represent a declining tendency, notifying the proficiency of the DATTMOA-HARD approach in harmonizing a tradeoff between generalization and data fitting. The successive reduction in loss values assures the improved performance of the DATTMOA-HARD approach and tunes the prediction solutions after a while.

Table 3; Fig. 9 portray the overall comparison analysis of the DATTMOA-HARD technique is portrayed^{33–35}. The results showed that the DATTMOA-HARD technique achieved a greater outcome. According to $accu_y$, the DATTMOA-HARD model has gained a maximal $accu_y$ of 98.66%. At the same time, the JDS-TL, Multi-

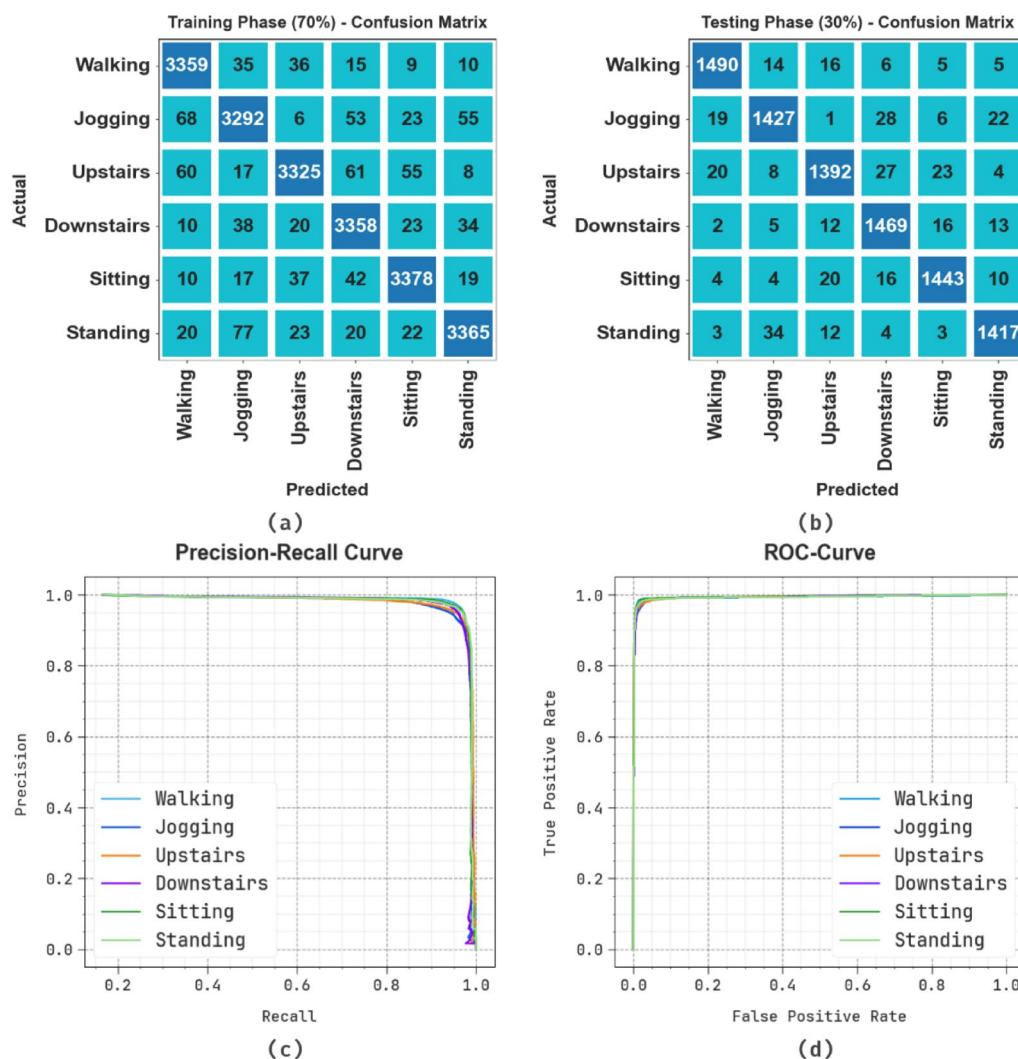
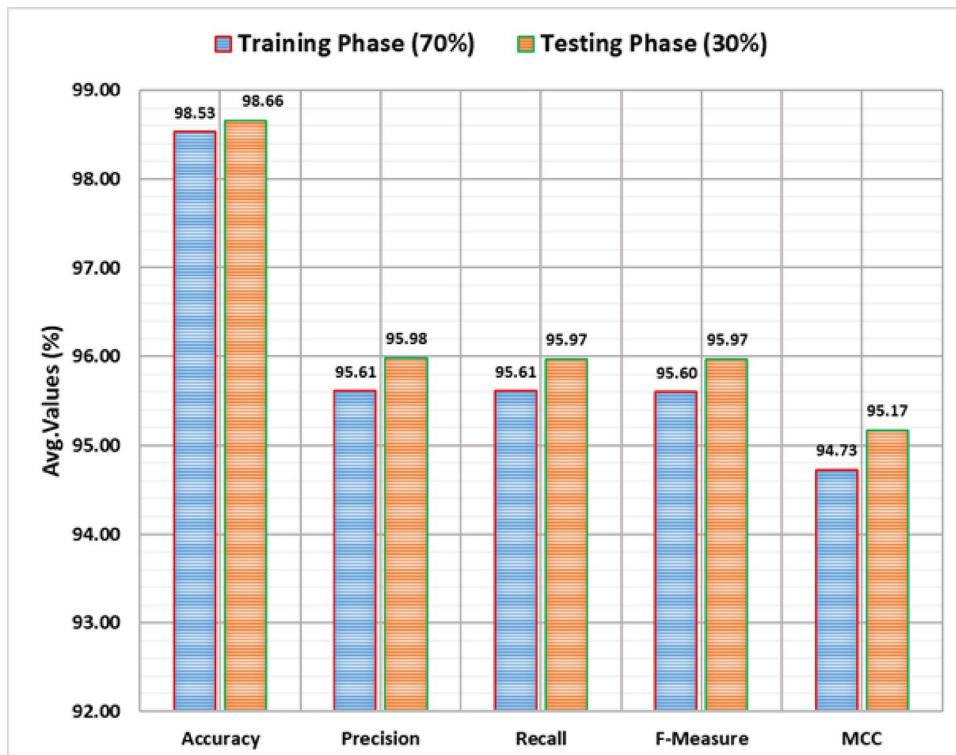


Fig. 5. (a-b) confusion matrix and **(c-d)** PR and ROC curves.

Class Labels	<i>Accu_y</i>	<i>Prec_n</i>	<i>Recal_t</i>	<i>F_{Measure}</i>	MCC
TRASE (70%)					
Walking	98.70	95.24	96.97	96.09	95.32
Jogging	98.15	94.71	94.14	94.42	93.31
Upstairs	98.46	96.46	94.30	95.37	94.45
Downstairs	98.50	94.62	96.41	95.51	94.61
Sitting	98.78	96.24	96.43	96.34	95.60
Standing	98.63	96.39	95.41	95.90	95.07
Average	98.53	95.61	95.61	95.60	94.73
TESSE (30%)					
Walking	98.96	96.88	97.01	96.94	96.31
Jogging	98.43	95.64	94.94	95.29	94.35
Upstairs	98.41	95.80	94.44	95.11	94.17
Downstairs	98.57	94.77	96.84	95.79	94.94
Sitting	98.81	96.46	96.39	96.42	95.71
Standing	98.78	96.33	96.20	96.26	95.53
Average	98.66	95.98	95.97	95.97	95.17

Table 2. Classifier outcome of DATTMOA-HARD technique with 70%TRASE and 30%TESSE.**Fig. 6.** Average of DATTMOA-HARD model with 70%TRASE and 30%TESSE.

channel CNN, WISNet, Bi-LSTM, Ensemble-GRU, Vid2Doppler, and MBGAN methodologies have achieved lesser *accu_y* of 87.89%, 95.51%, 95.82%, 94.10%, 96.96%, 81.64%, and 89.90%, respectively. In addition, according to *prec_n*, the DATTMOA-HARD model has attained a maximum *prec_n* of 95.98%. At the same time, the JDS-TL, Multi-channel CNN, WISNet, Bi-LSTM, Ensemble-GRU, Vid2Doppler, and MBGAN methodologies have accomplished minimal *prec_n* of 91.31%, 85.49%, 88.29%, 93.14%, 85.46%, 90.46%, and 94.02%, correspondingly.

Furthermore, according to *recal_t*, the DATTMOA-HARD method has reached an improved *recal_t* of 95.97%. At the same time, the JDS-TL, Multi-channel CNN, WISNet, Bi-LSTM, Ensemble-GRU, Vid2Doppler, and MBGAN techniques have attained lower *recal_t* of 82.13%, 94.44%, 87.05%, 93.72%, 85.80%, 87.12%, and 93.17%, correspondingly. Finally, according to *F_{measure}*, the DATTMOA-HARD technique has accomplished

Training and Validation Accuracy

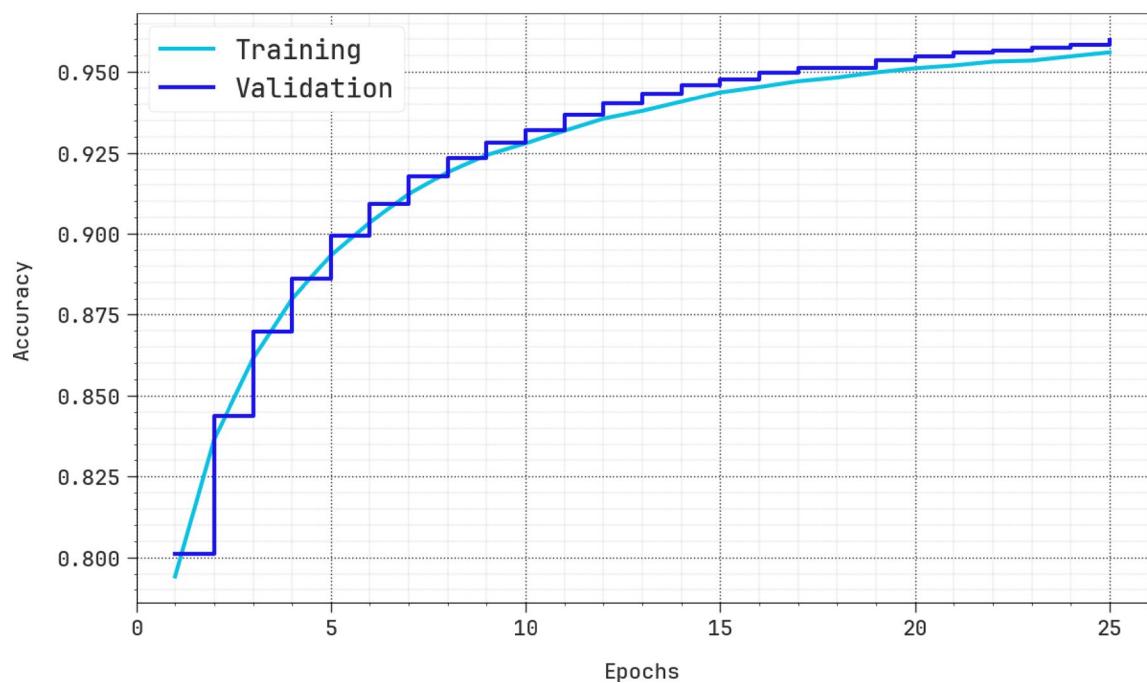


Fig. 7. $Accu_y$ curve of the DATTMoa-HARD approach.

Training and Validation Loss

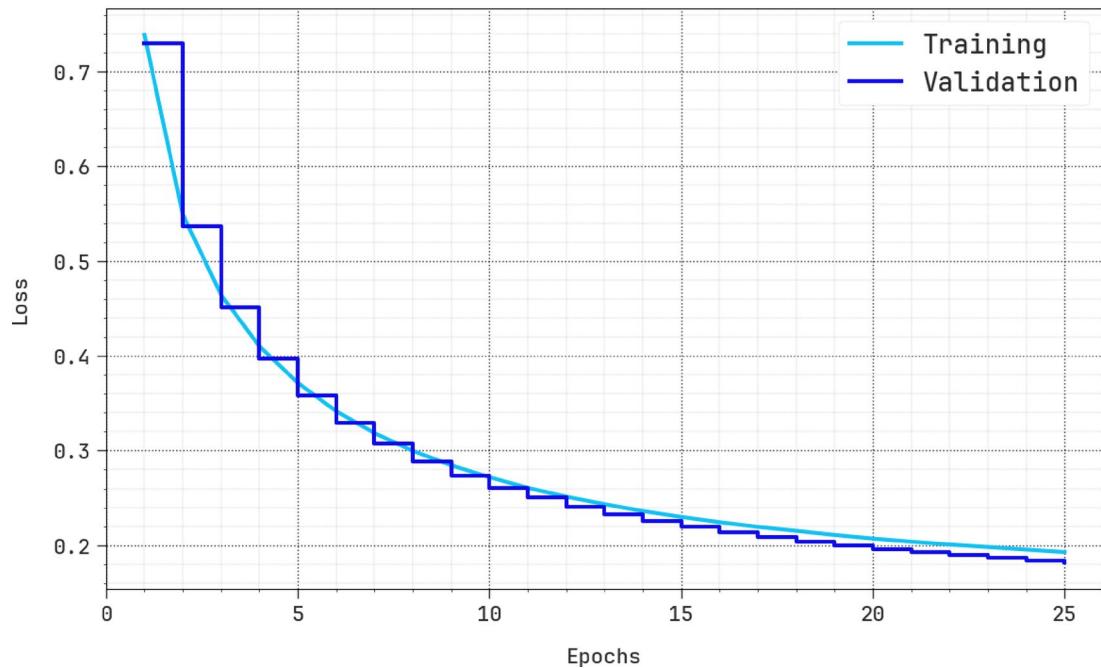
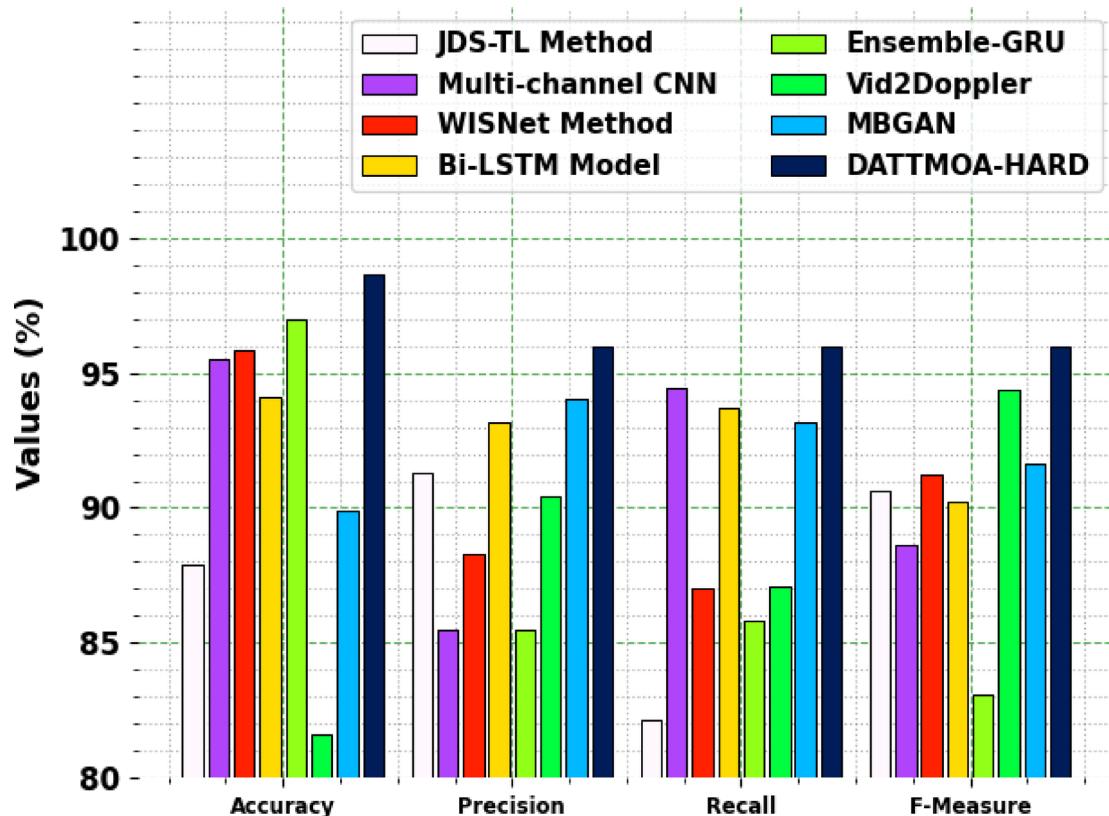


Fig. 8. Loss curve of the DATTMoa-HARD approach.

Classifier	<i>Accu_y</i>	<i>Prec_n</i>	<i>Recal_t</i>	<i>F_{Measure}</i>
JDS-TL Method	87.89	91.31	82.13	90.62
Multi-channel CNN	95.51	85.49	94.44	88.64
WISNet Method	95.82	88.29	87.05	91.25
Bi-LSTM Model	94.10	93.14	93.72	90.20
Ensemble-GRU	96.96	85.46	85.80	83.10
Vid2Doppler	81.64	90.46	87.12	94.39
MBGAN	89.90	94.02	93.17	91.63
DATTMOA-HARD	98.66	95.98	95.97	95.97

Table 3. Comparative results of the DATTMOA-HARD model with existing techniques^{33–35}.**Fig. 9.** Comparative analysis of the DATTMOA-HARD model with existing techniques.

a maximum $F_{measure}$ of 95.97%. At the same time, the JDS-TL, Multi-channel CNN, WISNet, Bi-LSTM, Ensemble-GRU, Vid2Doppler, and MBGAN models have reached a decrease $F_{measure}$ of 90.62%, 88.64%, 91.25%, 90.20%, 83.10%, 94.39%, and 91.63%, respectively.

Conclusion

In this manuscript, the DATTMOA-HARD technique is proposed. The main intention of the DATTMOA-HARD technique relies on improving HAR to assist disabled individuals. The DATTMOA-HARD model has data normalization, dimensionality reduction, HAR classification, and parameter selection to accomplish that. In the initial stage, the Z-score normalization converts input data into a beneficial format. Afterwards, BFA executed the feature selection method. Furthermore, the proposed DATTMOA-HARD model designs a DABiG technique for the classification process. Finally, the TDO-based hyperparameter selection is accomplished to enhance the detection outcomes of the DABiG model. The experimental evaluation of the DATTMOA-HARD approach is examined under the HAR dataset. The comparison analysis of the DATTMOA-HARD approach portrayed a superior accuracy value of 98.66% over existing methods.

Data availability

The data supporting this study's findings are openly available at <https://www.kaggle.com/datasets/die9origephit/human-activity-recognition>, reference number [32].

Received: 27 January 2025; Accepted: 16 July 2025

Published online: 29 September 2025

References

- Gupta, N. et al. Human activity recognition in artificial intelligence framework: a narrative review. *Artif. Intell. Rev.* **55** (6), 4755–4808 (2022).
- Zhang, P. et al. Semantics-guided neural networks for efficient skeleton-based human action recognition. In *proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 1112–1121). (2020).
- Qiu, S. et al. Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges. *Inform. Fusion.* **80**, 241–265 (2022).
- Xia, K., Huang, J. & Wang, H. LSTM-CNN architecture for human activity recognition. *IEEE Access* **8**, 56855–56866 (2020).
- Beddiar, D. R., Nini, B., Sabokrou, M. & Hadid, A. Vision-based human activity recognition: a survey. *Multimedia Tools Appl.* **79** (41), 30509–30555 (2020).
- Deepalechumi, N. & Mala, R. Leveraging variational autoencoder with Hippopotamus Optimizer-Based dimensionality reduction model for attention deficit hyperactivity disorder diagnosis data. *Journal Intell. Syst. & Internet Things*, **16**(1). (2025).
- Wang, Y., Cang, S. & Yu, H. A survey on wearable sensor modality centred human activity recognition in health care. *Expert Syst. Appl.* **137**, 167–190 (2019).
- Mekruksavanhich, S. & Jitpattanakul, A. Lstm networks using smartphone data for sensor-based human activity recognition in smart homes. *Sensors* **21**(5), 1636 (2021).
- Voicu, R. A., Dobre, C., Bajenaru, L. & Ciobanu, R. I. Human physical activity recognition using smartphone sensors. *Sensors* **19**(3), 458 (2019).
- Jalon Arias, E., Aguirre Paz, L. M. & Molina Chalacan, L. Multi-Sensor data fusion for accurate human activity recognition with deep learning. *Fusion: Pract. & Applications*, **13**(2). (2023).
- Khan, S. I. et al. Transition-aware human activity recognition using an ensemble deep learning framework. *Comput. Human Behav.* **162**, 108435 (2025).
- Duhayyim, M. A. Parameter-Tuned deep Learning-Enabled activity recognition for disabled people. *Computers Mater. & Continua*, **75**(3). (2023).
- Serpush, F., Menhaj, M. B., Masoumi, B. & Karasfi, B. Wearable sensors-based human activity recognition with deep convolutional neural network and fuzzy classification. *Wireless Pers. Commun.* **133**(2), 889–911 (2023).
- Dastbaravardeh, E., Askarpour, S., Saberi Anari, M. & Rezaee, K. Channel Attention-Based Approach with Autoencoder Network for Human Action Recognition in Low-Resolution Frames. *International Journal of Intelligent Systems*, **2024**(1), p.1052344. (2024).
- Cob-Parro, A. C., Losada-Gutiérrez, C., Marrón-Romera, M., Gardel-Vicente, A. & Bravo-Muñoz, I. A new framework for deep learning video based human action recognition on the edge. *Expert Syst. Appl.* **238**, 122220 (2024).
- Alotaibi, F. et al. Internet of Things-driven human activity recognition of elderly and disabled people using arithmetic optimization algorithm with LSTM autoencoder. *J. Disabil. Res.* **2** (3), 136–146 (2023).
- Jannat, M. K. A., Islam, M. S., Yang, S. H. & Liu, H. Efficient wi-fi-based human activity recognition using adaptive antenna elimination. *IEEE Access* <https://doi.org/10.1109/ACCESS.2023.3320069> (2023).
- Sinha, K. P. & Kumar, P. Human activity recognition from uav videos using an optimized hybrid deep learning model. *Multimedia Tools Appl.* **83** (17), 51669–51698 (2024).
- Zhou, H., Zhang, X., Feng, Y., Zhang, T. & Xiong, L. Efficient human activity recognition on edge devices using DeepConv LSTM architectures. *Sci. Rep.* **15**(1), 13830. (2025).
- Sinha, K. P., Kumar, P. & Ghosh, R. Ensemble based feature extraction and deep learning classification model with depth vision. *Comput. Inform.* **42** (4), 965–992 (2023).
- Rafee, A. N. M., Zishan, M. A. O. & Noor, J. Decision Tree-Based Real-Time Personalized Human Activity Recognition and Fall Detection Utilizing Cost Effective Highly Resource-Constrained Microcontrollers. Available at SSRN 5146354. (2025).
- Chen, J., Xu, X., Wang, T., Jeon, G. & Camacho, D. An IoT framework with multi-modal frequency fusion for wifi-based coarse and fine activity recognition. *IEEE Internet Things Journal* (2024).
- Zhou, H., Le, H. T., Zhang, S., Phung, S. L. & Alici, G. Hand gesture recognition from surface electromyography signals with graph convolutional network and attention mechanisms. *IEEE Sens. Journal* (2025).
- Sinha, K. P. & Kumar, P. Human activity recognition from UAV videos using a novel DMLC-CNN model. *Image Vis. Comput.* **134**, 104674 (2023).
- Ghous, G., Najam, S. & Jalal, A. February. Human Emotion Recognition from EEG-brain Signals using Enhanced Machine Learning Method. In *2025 6th International Conference on Advancements in Computational Sciences (ICACS)* (pp. 1–7). IEEE. (2025).
- Bollampally, A. et al. Optimizing edge computing for activity recognition: A bidirectional LSTM approach on the PAMAP2 dataset. *Eng. Technol. Appl. Sci. Res.* **14** (6), 18086–18093 (2024).
- Zhou, L., Liu, X., Guan, X. & Cheng, Y. CSSA-YOLO: Cross-scale spatiotemporal attention network for fine-grained behavior recognition in classroom environments. *Sensors* **25**(10), 3132 (2025).
- Henderi, H., Wahyuningsih, T. & Rahwanto, E. Comparison of min-max normalization and z-score normalization in the K-nearest neighbor (kNN) algorithm to test the accuracy of types of breast cancer. *Int. J. Inf. Inform. Syst.* **4**(1), 13–20 (2021).
- Jin, Z. et al. A data-driven framework for lithium-ion battery RUL using LSTM and XGBoost with feature selection via binary firefly algorithm. *Energy* **314**, 134229 (2025).
- Kaleeswari, P. et al. R and DABiG: breath pattern classification using the hybrid deep learning with optimal feature selection. *Technology Health Care*, p.09287329241303368. (2025).
- Dong, J. Optimizing curriculum for students: A machine learning approach to time management analysis. *J. Artif. Intell. Syst. Modelling*, **1** (03), 36–51 (2024).
- <https://www.kaggle.com/datasets/die9origephit/human-activity-recognition>
- Sharen, H. et al. WISNet: A deep neural network based human activity recognition system. *Expert Systems with Applications*, **258**, p.124999. (2024).
- Dentamaro, V., Gattulli, V., Impedovo, D. & Manca, F. Human activity recognition with smartphone-integrated sensors: A survey. *Expert Syst. Applications*, p.123143. (2024).
- Waqr, S. & Pätzold, M. A simulation-based framework for the design of human activity recognition systems using radar sensors. *IEEE Internet Things Journal* (2023).

Acknowledgements

The authors extend their appreciation to the King Salman center For Disability Research for funding this work through Research Group no KSRG-2024- 380.

Author contributions

Conceptualization: Hend Khalid Alkahtani, Gouse Pasha Mohammed and Radwa Marzouk, Data curation and Formal analysis: Hend Khalid Alkahtani, Gouse Pasha Mohammed Investigation and Methodology: Hend Khalid Alkahtani, Gouse Pasha Mohammed Project administration and Resources: Supervision; Hend Khalid Alkahtani, Writing—original draft: Hend Khalid Alkahtani, Gouse Pasha Mohammed and Radwa Marzouk Validation and Visualization: Hend Khalid Alkahtani, Gouse Pasha Mohammed and Radwa Marzouk Writing—review and editing, Hend Khalid Alkahtani, Gouse Pasha Mohammed and Radwa Marzouk All authors have read and agreed to the published version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.K.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025