JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Modeling of Human Activity Recognition using Crow Search Algorithm with Hybrid Deep Learning Approach

L. Maria Anthony Kumar[1*], Dr. S. Murugan[2], A. Therasa Alphonsa[3]
[1]Research Scholar, Department of Computer and Information Science, Annamalai University, Annamalai Nagar, Tamil Nadu, India.
[2]Assistant Professor, Dr. M.G.R. Government Arts and Science College for Women, Villupuram.
[3]Assistant Professor, PG and Research Department of Chemistry, Government Arts College, C-Mutlur, Chidambaram.

Abstract

Human Activity Recognition (HAR) is a vital task in the domain of computer vision and pattern recognition, with applications classifying healthcare to human-computer interaction. It contains the classification and detection of many human activities depending on visual input or sensor data. HAR has accomplished vital attention because of its potential in several domains comprising fitness tracking, surveillance, human-computer interaction, and healthcare. The deep learning (DL) structures excel in learning difficult features and temporal patterns in raw sensor data, allowing the formation of sophisticated methods that accurately identify a widespread of human activities. This study presents a novel approach to HAR by combining crow search algorithm with hybrid DL (HAR-CSAHDL) technique. The HDL technique involves Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. Additionally, to enhance the model's performance, hyperparameter tuning is employed using the CSA, an optimization technique simulated by the foraging behavior of crows. The HAR-CSAHDL architecture leverages the strengths of CNNs in feature extraction from raw sensor data and the sequential learning capabilities of LSTM networks to capture temporal dependencies in human activities. The performance of the HAR-CSAHDL approach is demonstrated through comprehensive experiments on benchmark HAR datasets. Comparative analyses with recent approaches highlight the superior performance of the CNN-LSTM hybrid model.

Keywords: Hybrid Deep Learning; Human Activity Recognition; Hyperparameter tuning; Crow Search Algorithm;

1. Introduction

Recently, Human Activity Recognition (HAR) has become an advanced research field, which can be applicable in different domains and increasing requirements for efficient services and home automation for elderly persons [1]. In these cases, AR in smart homes by exploiting simple and ubiquitous sensors captured heightened attention on the ambient intelligence field and allowed current technologies for enhancing the life quality of an individual within a home surroundings [2]. The crucial role of AR is to determine and recognize difficult as well as easier activities in real-time settings using sensor data. HAR is a significant research field because of its valuable contributions to human-centric field of research aimed to improve the quality of life by facilitating transportation, protecting smart cities and smart villages, and medical field [4], and supporting decision-makers responses effectively for improving the qualification of services.

HAR technique provides data with respect to the activity and performance of the subjects [5]. This is commonly achieved by receiving signals from smart sensors or smartphones and processing them by employing Machine Learning (ML) methods for identification [6]. HAR is utilized for continuously monitoring the patient with various diseases, locomotion, transportation, sports, and day-to-day living activities [7]. Fig. 1 depicts the general process of HAR. As smart sensors are the major sources of new data, pattern recognition techniques and ML approach determined an excessive contribution for establishing applications of smarter sensors. These techniques have been several algorithms appropriate for various fields. The data processing with ML methods comprised large data types such as velocity, variety, and volume; data techniques namely supervised and unsupervised approaches and utilizing effective methodologies [8], which adapts the data features. As data is generated by numerous sources with accurate data types, it is extremely important for adopting or applying techniques that control the data features [9]. Also, determining the optimum data algorithm that can be appropriate for the data, is one of the major stages to recognize patterns and accurately analyzing sensor data [10].

Helmi et al. [11] incorporate the DL and SI models for building a strong HAR model by employing portable sensor info. A lightweight model is constructed by using the Residual CNN and BiGRU (RCNN-BiGRU) model for the process of feature extraction. For optimum feature set selection, a novel feature selection model can employed depending on the Marine Predator Algorithm (MPA). Also, three binary variants such as MPAs, MPAs10, and MPAv are developed. In [12], a novel method by employing CNN model is presented with changing kernel dimensions together with BiLSTM method for capturing features at several resolutions. This method achieves the effectual extraction of temporal and spatial features from sensor data by employing conventional BiLSTM and CNN models. Additionally, the model is examined by using UCI and WISDM datasets.

Janardhanan and Umamaheswari [13] presented a DL-NN model by utilizing Depthwise Separable Convolution (DSC) with BLSTM (DSC-BLSTM) methods that follow three stages such as DSCBLSTM, Video data preparation, and Feature Extraction using Depthwise Separable CNN models. The redeeming aspects of this model consists of a DSC convolution. Surek et al. [14] introduce a DL technique for evaluating and mapping the present state of human activities in green, blue, and red videos. The self-DIstillation with NO labels (DINO) can be employed for enhancing the ResNet and ViT potential.

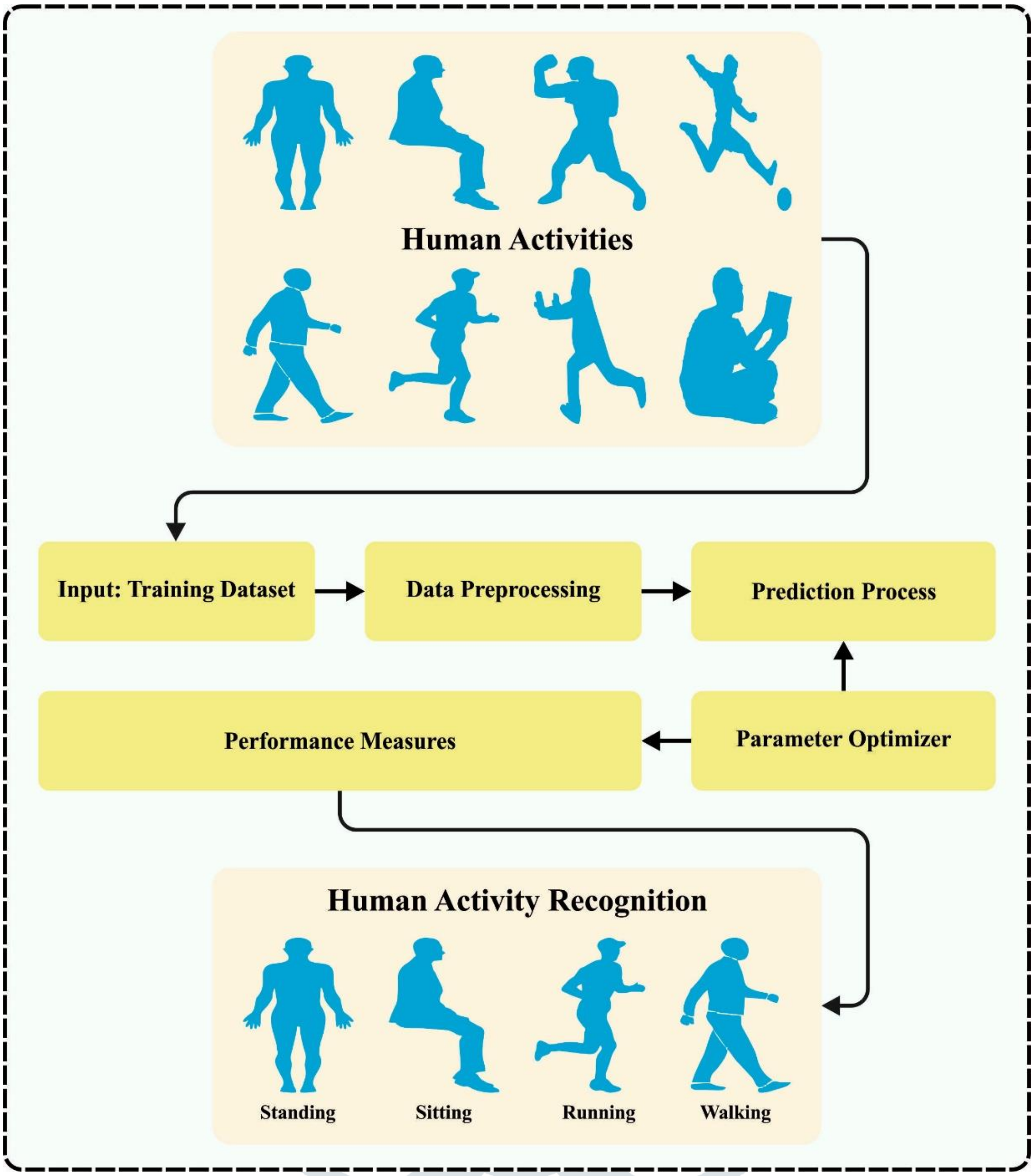

Fig. 1. General process of HAR

In [15], a new HAR scheme is presented depending on two optimization models namely Arithmetic Optimization Algorithm (AOA) and CNN. This CNN is enforced for extracting and learning aspects from input data, in which an altered AOA model, namely the Binary AOA (BAOA) method is utilized. Lastly, based on diverse actions, the SVM method is utilized for classifying the chosen factors. Nouriani et al. [16] present a technique incorporating DL computer vision classification and high-gain observer model. This presented non-sequential high-gain observer employs Lyapunov evaluation to precisely analyze the human subjects' attitude toward the chest by implementing dimensions from a single Inertial Measurement Unit (IMU). The signals put under processing by the observer are later altered into spectrograms for attaining imageries of the signal's frequency response.

This study presents a novel approach to HAR by combining crow search algorithm with hybrid DL (HAR-CSAHDL) technique. The HDL technique involves Convolutional Neural Networks (CNNs) and Long Short-

Term Memory (LSTM) networks. Additionally, to enhance the model's performance, hyperparameter tuning is employed using the CSA, an optimization technique inspired by the foraging behavior of crows. The HAR-CSAHDL architecture leverages the strengths of CNNs in feature extraction from raw sensor data and the sequential learning capabilities of LSTM networks to capture temporal dependencies in human activities. The performance of the HAR-CSAHDL approach is demonstrated through comprehensive experiments on benchmark HAR datasets.

2. The proposed method

In this study, we presented a novel HAR-CSAHDL technique. The proposed IMRFO-DAE methodology mostly detection the different classes of HARs. The main purpose of IMRFO-DAE approach contains two stages namely HDL based classification model and CSA based parameter tuning.

2.1. Activity Recognition using HDL model

Convolutional, output and multiple pooling layers are interconnected to the input layer in a CNN model [17].

Convolutional layer

The convolutional layer conducts a convolution function utilizing raw input information and convolution kernel for generating novel attribute values. The input signal needs to follow the organized matrix procedure because the model is made to feature extract from the image database. Compared to input matrix, the convolutional kernel is considered a small window which arranges coefficient into matrix. An attribute property named a convolved design is generated by the filter's assigned dimension element along with coefficient value. By applying multiple convolutional kernels to that input data, convoluted features can be generated, that are often useful than the essential characteristics of the original dataset. The convolutional layer serves as the basis of CNNs since they are where most of the computations were done.

$$P_m^{(a)} = \sigma\left(G_m^{(a)} + \sum_{n=1}^{s(a-1)} P_n^{(a-1)} * U_{m,n}^{(a)}\right) \tag{1}$$

In Eq. (1), the operator $\star$ represents the convolution function, σ shows the activation matrix, and $U_{m,n}^{(a)}$ denotes the filter connecting the n^{th} feature maps at $a-1$ layer with m^{th} feature maps in layer, the function a is used for increasing the nonlinearity.

Pooling layer

Generally, the pooling layer is put after the convolution layer. The pooling layer should streamline the output data. Using the information from all the feature maps in the convolution layer, the pooling layer generates compressed feature maps. Maximum and average pooling are the two commonly used techniques. In these layers, filters of size NxN were chosen.

$$\overline{a} = \frac{1}{L}\Sigma_{(m,n)\in G}\, a_{m,n} \tag{2}$$

$$a_{\max} = \max_{(m,n)\in G}\left(a_{m,n}\right) \quad (3)$$

From the expression, $a_{m,n}$ denotes the amount of all the pixels in region, G and L show the area's pixel count.

Dense layer

In the dense layer, LSTM method is utilized. RNN especially, LSTM-NN, has the capability to learn over time through feedback connection. Through cyclic connection on the HL, this technique produces short-term memory and gathers information from it. Also, it collects information via time sequences and series. LSTM unit includes a memory cell and the three gating mechanisms namely input, forget, and output. This enables to select of the data requires that "remembered," and the data needs that "forgotten" which leads to controlled data flow and training for long-term dependency.

$$a_u = \sigma(P_u m_s + R_u n_{s-1} + t_u) \quad (4)$$

In Eq. (4), R and P denote the weight matrix, m_s shows the input, σ indicates the sigmoid function, and t refers to the bias term vector.

Output layer

The neuron of output layer, also represented as the FC layer, depends entirely on the area of prior layer. This data became a 1D matrix. The overall amount of FC layer in the model might differ.

$$a_m^s = \sum_n u_f^{s-1} v_n^{s-1} \quad (5)$$

In Eq. (5), s shows the amount of layers, $\mathrm{v}_\mathrm{n}^\mathrm{s}$ denotes the values in the output layer and a_m^s indicates the value of activation function in the resultant layer. m and n show the neuron count, u_f^{s-1} refers to the HL weight, and v_n^{s-1} shows the input neuron.

2.2. Parameter tuning using CSA model

In this work, hyperparameter tuning is employed using the CSA model. CSA is developed based on the concept that crows store food in hiding place and brings it back if it can be need it [18]. In this study, crow characterizes a solution (location) to the optimization problems. Assume that the location of i^{th} crow at t^{th} iteration is formulated by: $X_i^t = \left(x_{i,1}^t, x_{i,2}^t, \dots, x_{i,D}^t\right)$, where $i = 1,2,\dots,N, t = 1,2,\dots,T_{\max}$, and $T_{\max}$ shows the maximal iteration count N is the size of crow population and D is the size of solution space of the optimizer problems. Every individual crow has memory and remembers their hiding places of food, i.e., the better location. Consider M_i^t as crow i hiding places at iteration t. During the foraging, crow moves in the environment and finds best hiding places (food sources).

The major stages of CSA are defined below:

Step1: Arbitrarily initialize a crow P population, and take P as an M primary memory (hiding places) of crows.

Step2: Crow i updates the location by randomly choosing j crow and following it. The location can be produced using the following expression:

$$X_i^{t+1} = \begin{cases} X_i^t + r_i \times fl_i^t \times (M_j^t - X_i^t), & r_j \geq AP_j^t \\ a\ random\ position, & otherwise \end{cases}, \tag{6}$$

In Eq. (6), r_i and r_j denotes the randomly generated number within [0,1], fl_i^t shows the flight length of the i crow at t^{th} iteration, and AP_j^t indicates the awareness probability of being followed by j crow at t iteration.

Step3: Compute the fitness of i crow based on the newest location, and upgrade the memory using Eq. (7):

$$M_i^{t+1} = \begin{cases} X_i^{t+1}, & f(X_i^{t+1})\ \ is\ better\ than\ f(M_i^t) \\ M_i^t, & otherwise \end{cases} \tag{7}$$

In Eq. (7), f shows the fitness value.

Step4: Repeat steps 2-3 for the crow until the ending criteria is reached.

3. Experimental validation

The HAR-CSAHDL approach is examined on two databases namely UCI HAR database and USC HAD database. The UCI HAR database includes 10299 instances with six classes. In addition, the USC HAD database comprises 420 samples with six classes.

Table 1 Recognition outcome of HAR-CSAHDL approach on UCI HAR database

UCI HAR Database					
Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{Score}	**MCC**
Training Phase (70%)					
Walking	96.84	90.62	98.10	90.62	88.72
Walking Upstairs (WU)	95.78	83.01	98.06	85.61	83.19
Walking Downstairs (WD)	95.67	83.75	97.60	84.34	81.83
Sitting (Si)	96.73	89.74	98.13	90.19	88.23
Standing (St)	96.62	94.40	97.10	90.98	88.99
Lying (Sl)	95.37	87.77	97.18	87.96	85.09
Average	**96.17**	**88.21**	**97.70**	**88.28**	**86.01**
Testing Phase (30%)					
Walking	96.83	91.32	97.91	90.43	88.54
Walking Upstairs (WU)	95.21	80.66	97.72	83.22	80.48
Walking Downstairs (WD)	95.18	78.66	97.66	80.97	78.26
Sitting (Si)	96.38	90.32	97.74	90.16	87.94
Standing (St)	95.95	93.37	96.58	90.01	87.57
Lying (Sl)	96.25	88.99	97.83	89.47	87.19
Average	**95.97**	**87.22**	**97.57**	**87.38**	**85.00**

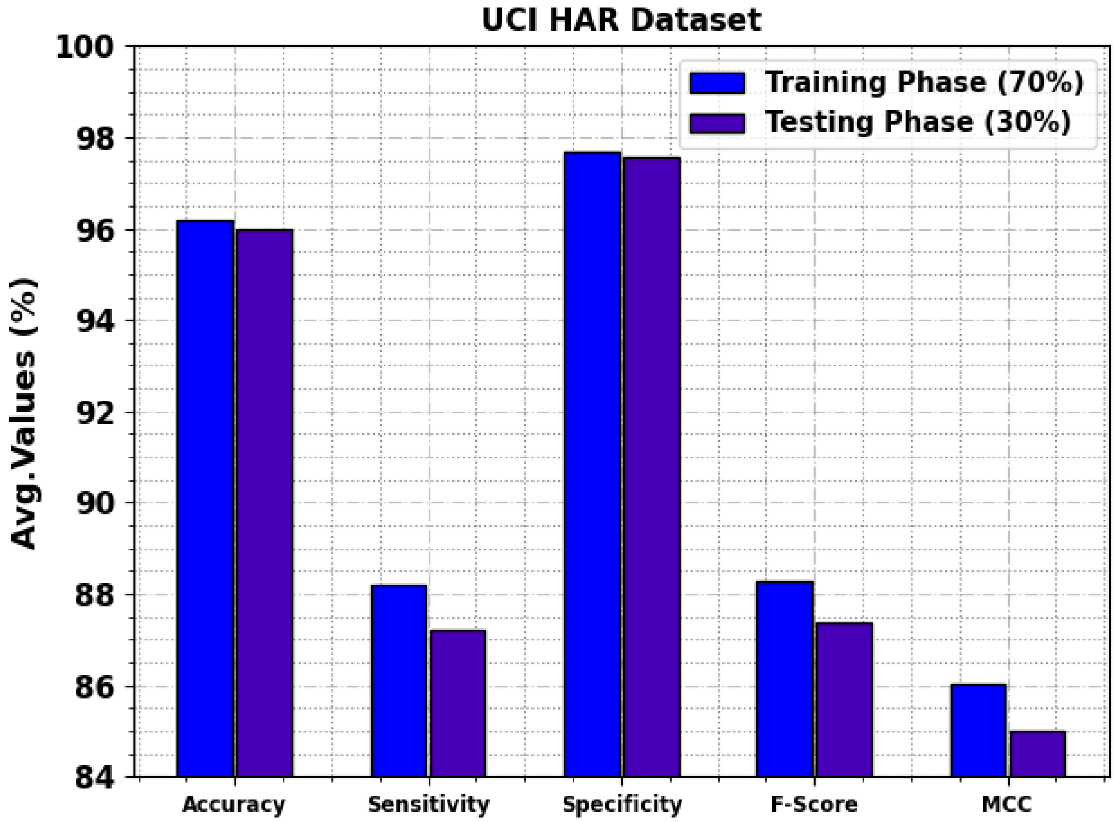

Fig. 2. Average of HAR-CSAHDL approach on UCI HAR database

The recognition results of the HAR-CSAHDL approach are examined on the UCI HAR database as depicted in Table 1 and Fig. 2. The results indicate that the HAR-CSAHDL approach reaches effectual outcomes on six classes. On 70% TRP, the HAR-CSAHDL approach resulted in average $accu_y$ of 96.17%, $sens_y$ of 88.21%, $spec_y$ of 97.70%, F_{score} of 88.28%, and MCC of 86.01%. In addition, on 30% TSP, the HAR-CSAHDL method resulted in average $accu_y$ of 95.97%, $sens_y$ of 87.22%, $spec_y$ of 97.57%, F_{score} of 87.38%, and MCC of 85%.

The recognition outcome of the HAR-CSAHDL system is examined on the USC HAD database as shown in Table 2 and Fig. 3. The result indicates that the HAR-CSAHDL approach reaches effectual outcomes on six classes. On 70% TRP, the HAR-CSAHDL system resulted in average $accu_y$ of 92.40%, $sens_y$ of 77.23%, $spec_y$ of 95.44%, F_{score} of 77.18%, and MCC of 72.73%. Furthermore, on 30% TSP, the HAR-CSAHDL methodology resulted in average $accu_y$ of 96.03%, $sens_y$ of 88.85%, $spec_y$ of 97.62%, F_{score} of 88.31%, and MCC of 86.13%.

Table 2 Recognition outcome of HAR-CSAHDL approach on USC HAD database

USC HAD Database					
Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{Score}	**MCC**
Training Phase (70%)					
Walking Left (WL)	94.90	86.54	96.69	85.71	82.61
Walking Downstairs (WD)	90.48	78.00	93.03	73.58	67.96
Running Forward (RF)	92.86	78.72	95.55	77.89	73.64

Standing (St)	93.20	79.55	95.60	77.78	73.79
Sleeping (Sl)	93.20	72.55	97.53	78.72	75.09
Elevating Up (EU)	89.80	68.00	94.26	69.39	63.29
Average	**92.40**	**77.23**	**95.44**	**77.18**	**72.73**
Testing Phase (30%)					
Walking Left (WL)	98.41	100.00	98.15	94.74	93.99
Walking Downstairs (WD)	97.62	90.00	99.06	92.31	90.94
Running Forward (RF)	94.44	86.96	96.12	85.11	81.72
Standing (St)	94.44	76.92	99.00	85.11	82.44
Sleeping (Sl)	95.24	84.21	97.20	84.21	81.41
Elevating Up (EU)	96.03	95.00	96.23	88.37	86.30
Average	**96.03**	**88.85**	**97.62**	**88.31**	**86.13**

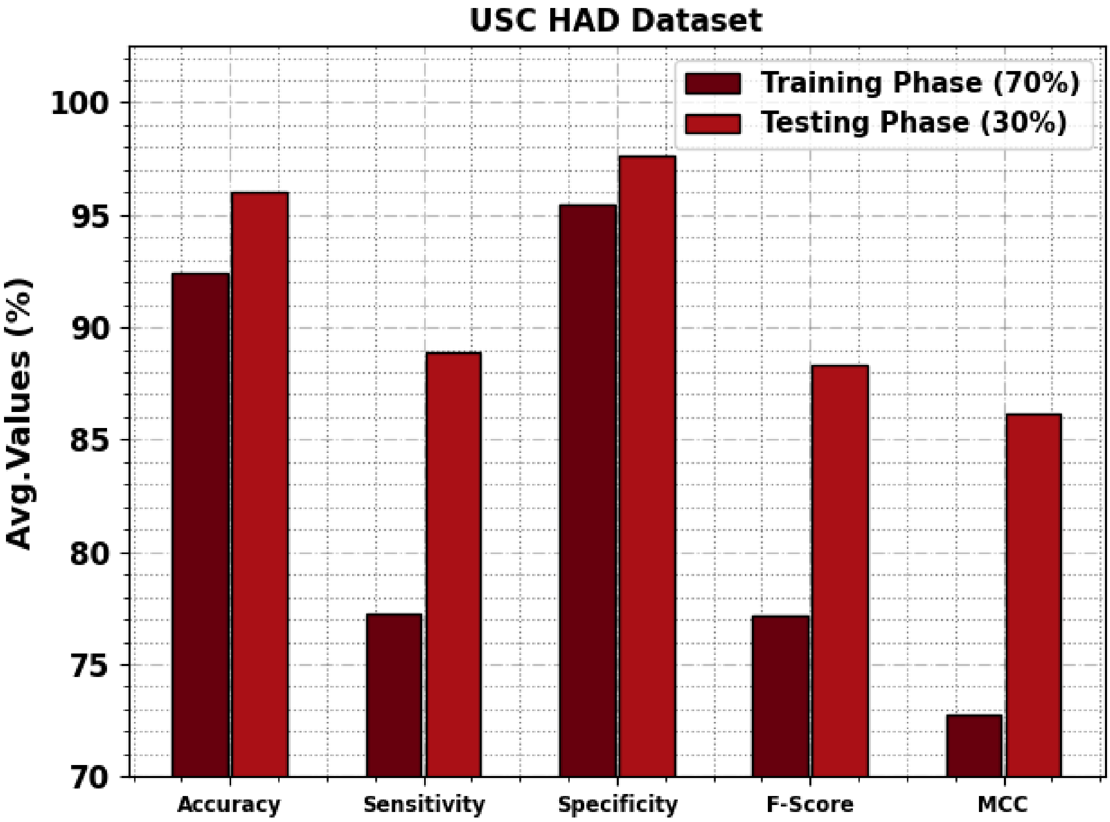

Fig. 3. Average of HAR-CSAHDL approach on USC HAD database

Fig. 4 illustrates the training accuracy TR_accu_y and VL_accu_y of the HAR-CSAHDL system on USC HAD database. The TL_accu_y is determined by the assessment of the HAR-CSAHDL method on TR database whereas the VL_accu_y is computed by evaluating the performance on a separate testing database. The results exhibit that TR_accu_y and VL_accu_y increase with an upsurge in epochs. Consequently, the performance of the HAR-CSAHDL technique gets improved on the TR and TS database with a rise in number of epochs.

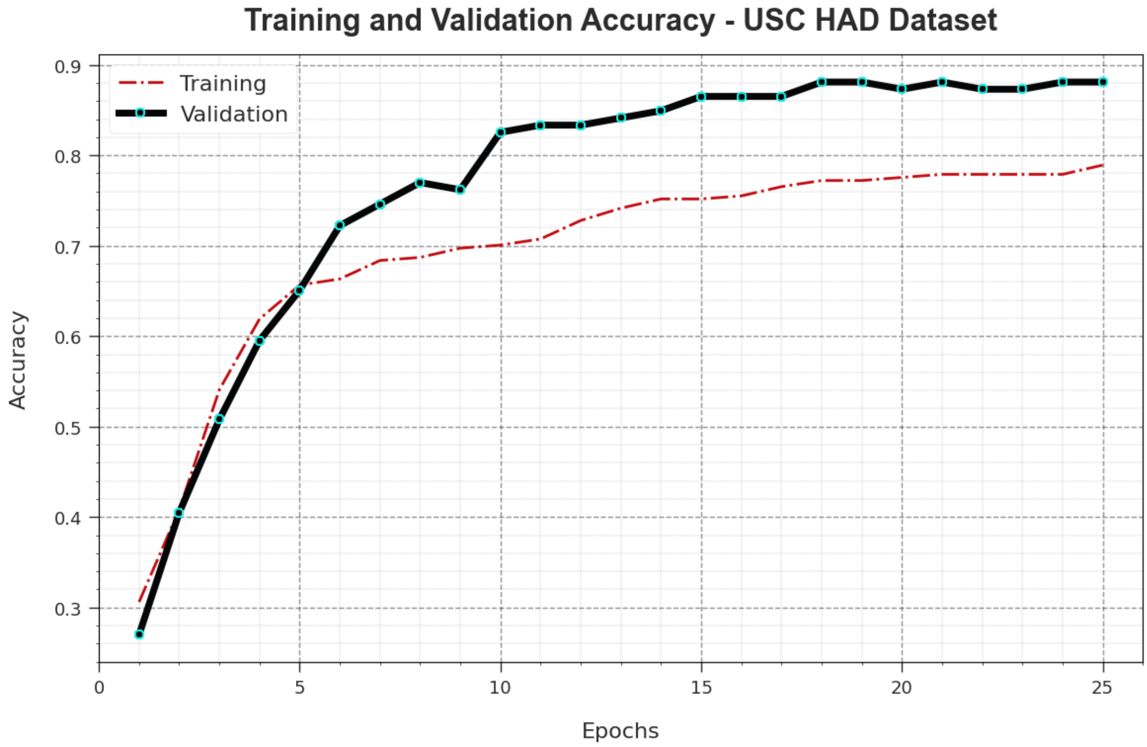

Fig. 4. $Accu_y$ curve of HAR-CSAHDL approach on USC HAD database

In Fig. 5, the TR_loss and VR_loss outcome of the HAR-CSAHDL technique on USC HAD database is shown. The TR_loss defines the error among the predictive performance and original values on the TR data. The VR_loss represents the measure of the performance of the HAR-CSAHDL technique on individual validation data. The results indicate that the TR_loss and VR_loss tend to decrease with rising epochs. It portrayed the enhanced performance of the HAR-CSAHDL method and its capability to generate accurate classification. The reduced value of TR_loss and VR_loss demonstrates the enhanced performance of the HAR-CSAHDL technique on capturing patterns and relationships.

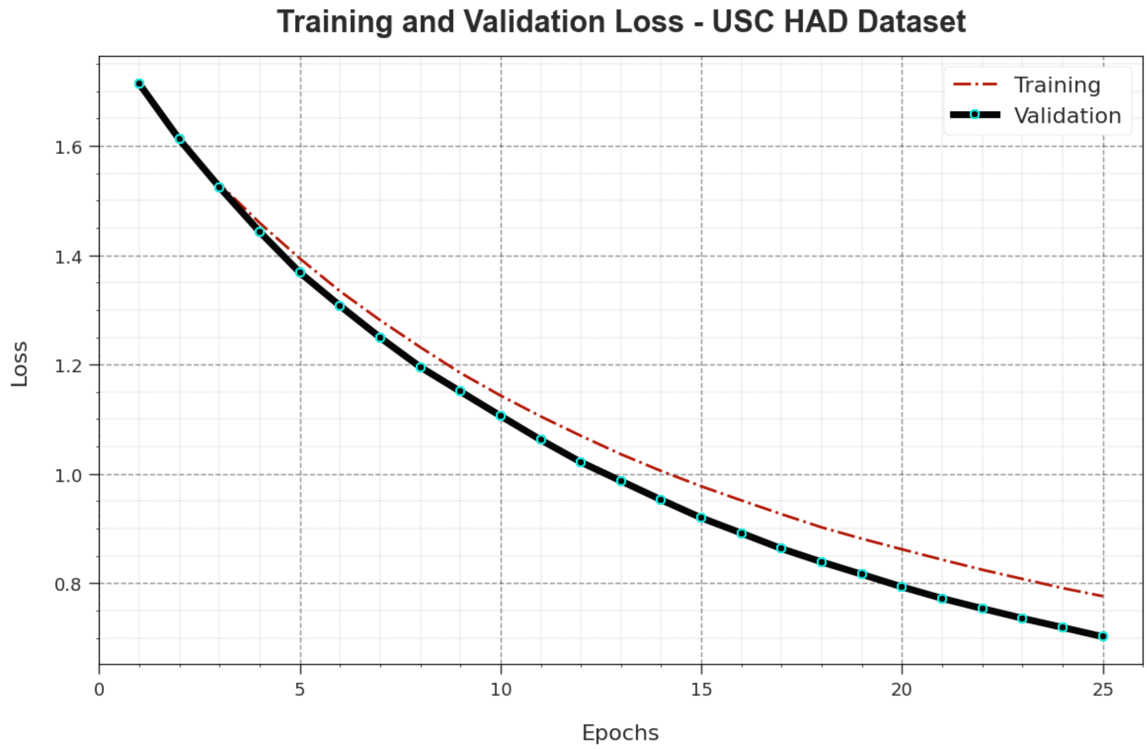

Fig. 5. Loss curve of HAR-CSAHDL approach on USC HAD database

The comparative statement of the HAR-CSAHDL approach with recent DL models is established in Table 3.

Table 3 $Accu_y$ outcome of HAR-CSAHDL approach with other systems on two databases

Methods	**UCI HAR Database**	**USC HAD Database**
CNN	89.456	85.264
LSTM	89.674	83.084
CNN-LSTM	87.339	87.414
Conv. LSTM	90.851	84.862
RHAR-EODELM	95.950	95.920
The Proposed Model	96.17	96.03

In Fig. 6, the activity recognition outcome of the HAR-CSAHDL system with existing methods on the UCI HAR database is reported. The outcome signified that the HAR-CSAHDL approach produces maximum $accu_y$ of 96.17%. In the meantime, the CNN, LSTM, CNN-LSTM, Conv. LSTM, and RHAR-EODELM models accomplish decreased $accu_y$ values of 89.456%, 89.674%, 87.339%, 90.851%, and 95.950% respectively.

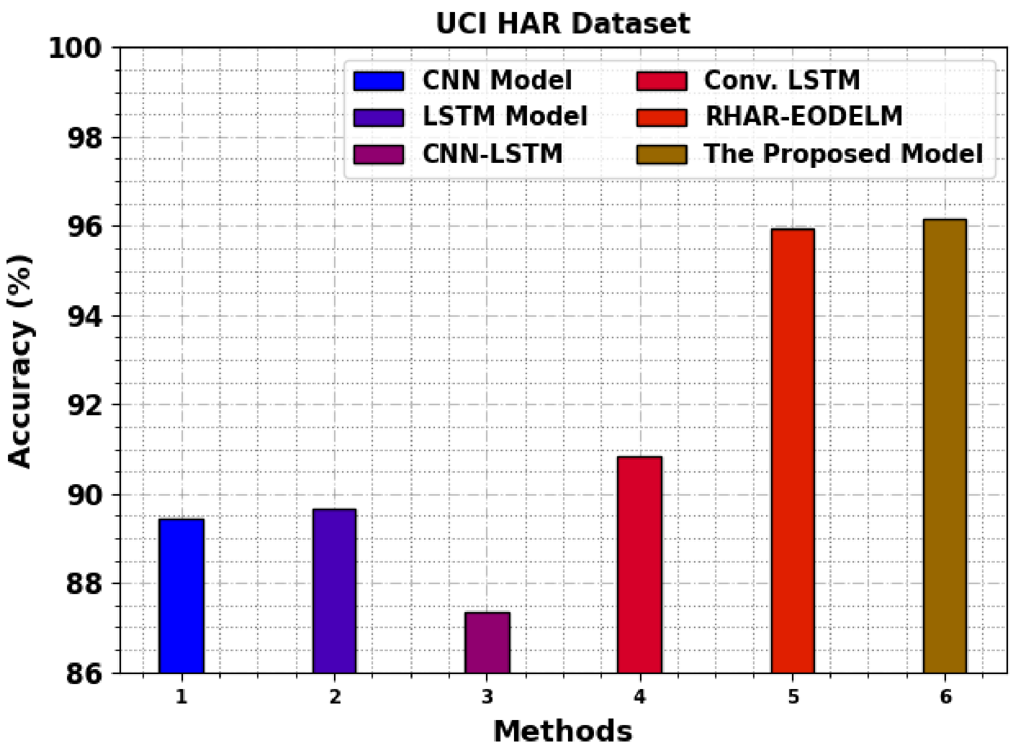

Fig. 6. $Accu_y$ outcome of HAR-CSAHDL approach on UCI HAR database

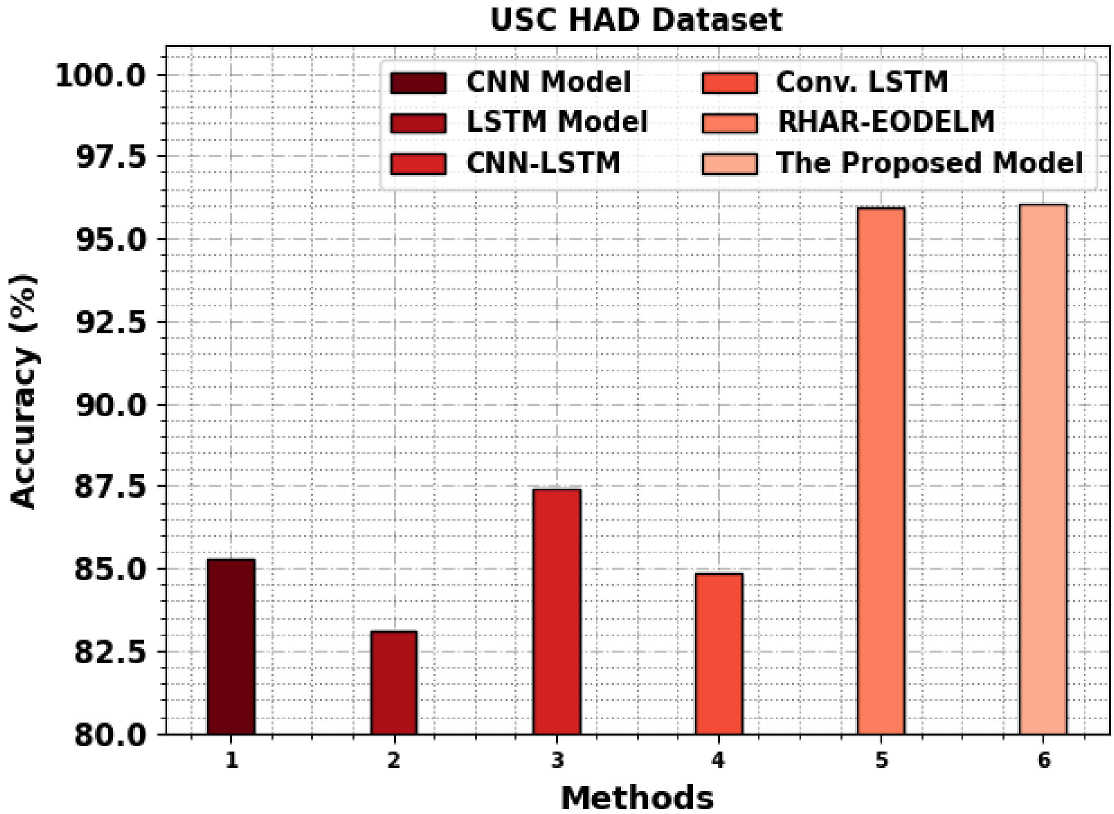

Fig. 7. $Accu_y$ outcome of HAR-CSAHDL approach on USC HAD database

In Fig. 7, the activity recognition outcome of the HAR-CSAHDL system with existing methods on the USC HAD database is reported. The result represented that the HAR-CSAHDL approach produces maximal $accu_y$ of 96.03%. Besides, the CNN, LSTM, CNN-LSTM, Conv. LSTM, and RHAR-EODELM approaches accomplish lesser $accu_y$ values of 85.264%, 83.084%, 87.414%, 84.862%, and 95.920% correspondingly.

4. Conclusion

In this study, we presented a novel HAR-CSAHDL methodology. The HDL technique involves CNNs and LSTM networks. Additionally, to enhance the model's performance, hyperparameter tuning is employed using the CSA, an optimization technique simulated by the foraging behavior of crows. The HAR-CSAHDL architecture leverages the strengths of CNNs in feature extraction from raw sensor data and the sequential learning capabilities of LSTM networks to capture temporal dependencies in human activities. The performance of the HAR-CSAHDL approach is demonstrated through comprehensive experiments on benchmark HAR datasets.

References

[1] Saha, A., Rajak, S., Saha, J. and Chowdhury, C., 2022. A Survey of Machine Learning and Meta-heuristics Approaches for Sensor-based Human Activity Recognition Systems. *Journal of Ambient Intelligence and Humanized Computing*, pp.1-28.

[2] Kaur, M., Kaur, G., Sharma, P.K., Jolfaei, A. and Singh, D., 2020. Binary cuckoo search metaheuristic-based supercomputing framework for human behavior analysis in smart home. *The Journal of*

Supercomputing, *76*(4), pp.2479-2502.

[3] Basak, H., Kundu, R., Singh, P.K., Ijaz, M.F., Woźniak, M. and Sarkar, R., 2022. A union of deep learning and swarm-based optimization for 3D human action recognition. *Scientific Reports*, *12*(1), pp.1-17.

[4] Nweke, H.F., Teh, Y.W., Mujtaba, G., Alo, U.R. and Al-garadi, M.A., 2019. Multi-sensor fusion based on multiple classifier systems for human activity identification. *Human-centric Computing and Information Sciences*, *9*(1), pp.1-44.

[5] Erdaş, Ç.B. and Güney, S., 2021. Human activity recognition by using different deep learning approaches for wearable sensors. *Neural Processing Letters*, *53*(3), pp.1795-1809.

[6] Li, T., Fong, S., Wong, K.K., Wu, Y., Yang, X.S. and Li, X., 2020. Fusing wearable and remote sensing data streams by fast incremental learning with swarm decision table for human activity recognition. *Information Fusion*, *60*, pp.41-64.

[7] Tuncer, T. and Ertam, F., 2021. Novel tent pooling based human activity recognition approach. *Multimedia Tools and Applications*, *80*(3), pp.4639-4653.

[8] Rosati, S., Balestra, G. and Knaflitz, M., 2018. Comparison of different sets of features for human activity recognition by wearable sensors. *Sensors*, *18*(12), p.4189.

[9] Tsokov, S., Lazarova, M. and Aleksieva-Petrova, A., 2021, June. Evolving 1d convolutional neural networks for human activity recognition. In *International Conference on Computer Systems and Technologies' 21* (pp. 49-54).

[10] Fang, H., Tang, P. and Si, H., 2020. Feature selections using minimal redundancy maximal relevance algorithm for human activity recognition in smart home environments. *Journal of Healthcare Engineering*, *2020*.

[11] Helmi, A.M., Al-qaness, M.A., Dahou, A. and Abd Elaziz, M., 2023. Human activity recognition using marine predators algorithm with deep learning. *Future Generation Computer Systems*, *142*, pp.340-350.

[12] Nafea, O., Abdul, W., Muhammad, G. and Alsulaiman, M., 2021. Sensor-based human activity recognition with spatio-temporal deep learning. *Sensors*, *21*(6), p.2141.

[13] Janardhanan, J. and Umamaheswari, S., 2022. Vision based Human Activity Recognition using Deep Neural Network Framework. *International Journal of Advanced Computer Science and Applications*, *13*(6).

[14] Surek, G.A.S., Seman, L.O., Stefenon, S.F., Mariani, V.C. and Coelho, L.D.S., 2023. Video-Based Human Activity Recognition Using Deep Learning Approaches. *Sensors*, *23*(14), p.6384.

[15] Dahou, A., Al-qaness, M.A., Abd Elaziz, M. and Helmi, A., 2022. Human activity recognition in IoHT

applications using arithmetic optimization algorithm and deep learning. *Measurement*, *199*, p.111445.

[16] Nouriani, A., McGovern, R. and Rajamani, R., 2023. Activity recognition using a combination of high gain observer and deep learning computer vision algorithms. *Intelligent Systems with Applications*, *18*, p.200213.

[17] Orosoo, M., Govindasamy, S., Bayarsaikhan, N., Rajkumari, Y., Fatma, G., Manikandan, R. and Bala, B.K., 2023. Performance analysis of a novel hybrid deep learning approach in classification of quality-related English text. *Measurement: Sensors*, p.100852.

[18] He, J., Peng, Z., Zhang, L., Zuo, L., Cui, D. and Li, Q., 2023. Enhanced crow search algorithm with multi-stage search integration for global optimization problems. *Soft Computing*, pp.1-31.