SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V12I1P113

Volume 12 Issue 1, 142-150, January 2025

Original Article

Human Activity Recognition Using Chaotic Logistic Map Guided Grey Wolf Optimization with Decision Tree

M. Janaki[1*], Sarojini Balakrishnan[1], S.N. Geethalakshmi[1]

[1]*Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Tamilnadu, India.*
[*]*Department of BCA, New Horizon College, Karnataka, India.*

[*]*Corresponding Author :19phcsp008@avinuty.ac.in*

Received: 10 November 2024 Revised: 12 December 2024 Accepted: 13 January 2025 Published: 25 January 2025

***Abstract** - Wearable sensors are essential for recognizing human activity in sports, healthcare, and smart surroundings applications. Robust classification models and effective feature selection directly affect recognition accuracy. This paper proposes a novel approach called Chaotic Logistic Map-based Grey Wolf Optimization with Decision Tree (CLM-GWO-DT) to predict human activity recognition. The proposed technique improves the GWO algorithm by using chaotic logistic maps to enhance its exploration and exploitation abilities. CLM-GWO is used to find the most informative features in raw sensor data, thereby reducing dimensionality and enhancing relevant patterns. A Decision Tree (DT) classifier is then applied to the retrieved data to ensure accurate and interpretable identification of human activity. The experiments employed two popular datasets: UCI Human Activity Recognition (HAR) and Wireless Sensor Data Mining (WISDM). The results indicate that the proposed model exceeds the performance of existing methods in the literature concerning accuracy, precision, recall, F-Score, and Matthews Correlation Coefficient (MCC).*

***Keywords** - Chaotic logistic map, Feature selection, Grey wolf optimization, Human activity recognition.*

1. Introduction

The quickly developing field of HAR uses information gathered from wearable sensors like accelerometers and gyroscopes to identify and categorize human actions, such as walking, jogging, and sitting. With the increasing integration of smart devices in daily life, HAR has gained significant consideration for its possible benefits in enabling intelligent systems in numerous domains. Applications of HAR include healthcare for monitoring patient recovery and detecting falls, fitness tracking to provide personalized activity insights [1], elder care for ensuring safety [2], and smart environments for adaptive automation. The ability to accurately recognize activities is essential for enhancing quality of life, improving safety, and enabling real-time decision-making in diverse scenarios. Feature selection is a critical step in HAR systems, aimed at identifying the most relevant features from high-dimensional sensor data while eliminating irrelevant or redundant ones. This procedure improves Machine Learning (ML) model performance and lowers computing complexity [3, 4]. Optimization techniques are essential to feature selection because they efficiently traverse the vast search area to find the best subsets of features [5]. Among the various optimization algorithms like Particle swarm optimization [6], Whale Optimization [7], and Grey Wolf Optimization (GWO) [8], GWO has demonstrated superior performance due to its effective balance between exploration and exploitation, influenced by grey wolf hunting tactics and leadership structures. In this work, we provide an improved CLM-GWO method for HAR feature selection. By using a chaotic logistic map, we avoid premature convergence and increase the diversity of options, making the optimization process more accurate and efficient overall. We use a DT classifier, which is renowned for its ease of use, interpretability, and excellent performance in activity identification tests, to categorize the chosen characteristics [9]. Based on a particular criterion, the technique determines the feature and threshold that result in the optimal split [10]. Due to its effectiveness in [11], tree-based ensemble techniques have been selected.The performance of the proposed CLM-GWO with DT is evaluated using benchmark datasets, including WISDM and UCI HAR.

These datasets provide extensive sensor data, making them ideal for validating the effectiveness of feature selection and classification approaches in HAR [12]. The suggested approach is a reliable solution for HAR applications since it reduces computing overhead and achieves improved classification accuracy, as shown by the experimental findings. By leveraging the strengths of chaotic systems, grey wolf optimization, and decision tree classification, this study contributes to advancing HAR systems' efficiency and reliability with potential applications in healthcare, fitness, and smart environments.

2. Literature Review

In the study by Garcia et al. [13], they introduced a dataset for HAR that works regardless of where the sensors are placed, how they are oriented, or who is using them. They tested the dataset using the Support Vector Machine (SVM) methodology, which produced an accuracy of 74.39%. Their model helped to make the recognition of human activities more practical and closer to real-life applications. Gyroscope and accelerometer data from smartphones are used in the study by Hayat, A., Dias, M. Bhuyan, B.P. Tomar, R. [14] to track the activities of senior citizens in different settings. For activity recognition, ML and DL methodologies that include k-NN, Random Forest, SVM, ANN, and LSTM were used. With 10-fold cross-validation, SVM offered 89.07% accuracy at a low computation time of 0.42 minutes, while LSTM reached the maximum accuracy of 95.04%. ButchiRaju et al. [15] investigated a medical sensor dataset that included a number of factors in a different investigation. With a remarkable accuracy of 94%, they demonstrated a smart heart disease prediction system using the LSTM-CNN architecture.

This study highlights how important it is to include a variety of medical characteristics when using sophisticated architectural models to forecast health issues.Kun Xia et al. [16] used the WISDM dataset to apply a 3-CNN approach for human activity recognition, which produced a 92.30% accuracy rate. The significance of DL methods for precise and subtle activity identification from wearable sensor data is highlighted by their emphasis on Convolutional Neural Networks (CNNs). Federico Cruciani et al. [17] study on UCI-HAR utilized Convolutional Neural Networks (3-CNN) to learn features with a classification accuracy of 95.85%, demonstrating CNN's potential for HAR. Syed K. Bash et al. [18] introduced Smartphone-Based Neighborhood Component Analysis (NCA) with Feature selection, SVM and Dense Neural Network, reaching an excellent accuracy of 95.79%. This study stands out from other works due to the innovative combination of feature selection, SVM, and dense neural networks. It shows that the completeness of models can improve HAR accuracy. Laith Abualigah's [4] work investigates diverse feature selection techniques, focusing on Filter, Wrapper, and Embedded methods to identify optimal feature subsets for enhanced predictive accuracy. The study also reviews optimization algorithms and assesses various feature selection approaches on standard datasets.

3. Proposed Methodology

A HAR framework, including pre-processing, feature selection, and classification methods, is shown in Figure 1. The UCI HAR and WISDM datasets are used in this work. The datasets are pre-processed using data normalization techniques to standardize the input for consistent and efficient processing; noises and special characters are eliminated. Next, the feature selection process uses the Chaotic Logistic Map-Grey Wolf Algorithm (CLM-GWO). This step identifies the most relevant features from the high-dimensional sensor data, reducing redundancy and dimensionality while retaining critical information. The optimized features are then passed to a Decision Tree Classifier, which performs the classification of human activities such as walking, sitting, or standing. The GWO method is a metaheuristic optimization method originally developed based on the social organization and hunting techniques used by grey wolves in the wild. In it, the population is divided into alpha, beta, delta and omega mobs-mimicking the leadership systems and pack hunting technique.

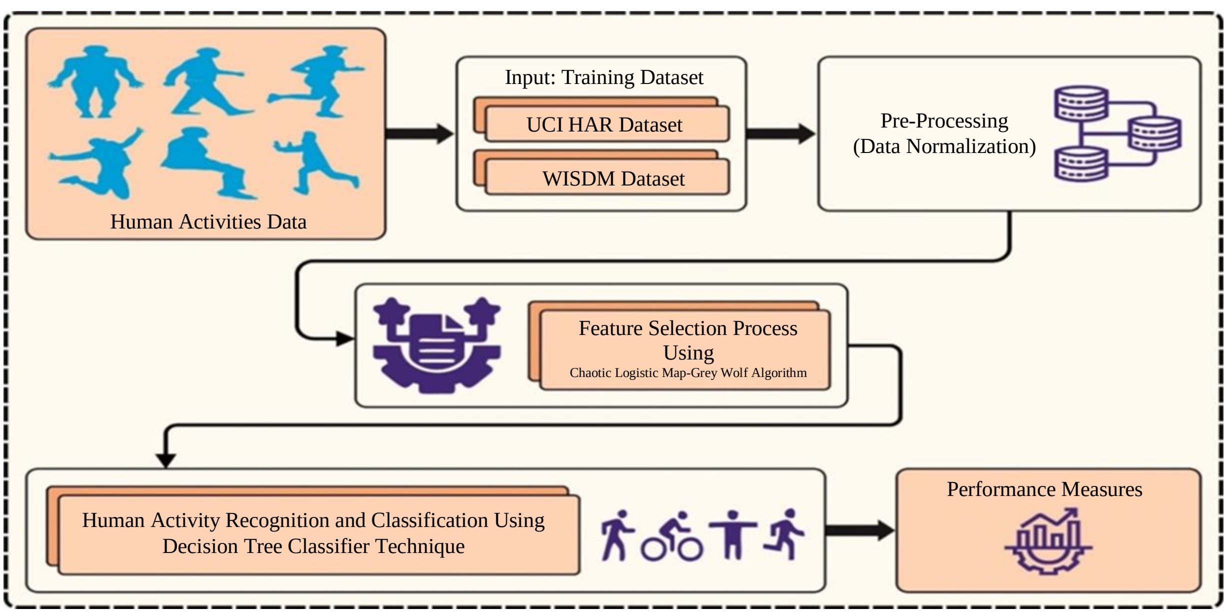

Fig. 1 Framework of proposed system

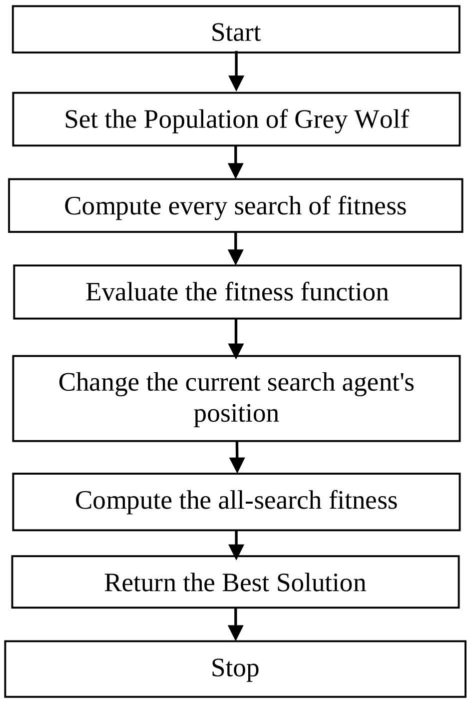

Fig. 2 Steps involved in GWO

GWO excels at balancing exploration (global search) and exploitation (local refinement), which makes it particularly effective in high-dimensional optimization problems like feature selection [19]. Particularly in nonlinearand complicated datasets, GWO is more resilient against local optima, has fewer parameters, and is computationally easier than techniques like “Whale Optimization Algorithm (WOA) or Particle Swarm Optimization (PSO)” [20, 21]. These advantages make it an excellent choice for reducing redundancy, improving model accuracy, and increasing computational efficiency. Figure 2 depicts the steps involved in GWO. When combined with chaotic systems, such as the logistic map, GWO’s performance is further enhanced. Chaos introduces dynamic randomness that enriches diversity in the population, preventing premature convergence and improving global search capability [22, 23].

The CGWO method is especially effective for feature selection, as it efficiently identifies the most relevant features while eliminating noise and irrelevant data by integrating CGWO with a DT classifier; the most optimized feature subset is utilized for classification. DT offer interpretable, rule-based models, and their collaboration with CGWO leads to high accuracy and reduced computational complexity, making them ideal for applications like HAR and other domains requiring precise, scalable, and interpretable solutions.In order to make the framework deliver high accuracy and resilience in HAR tasks, the methodology efficacy is finally evaluated through the use of performance metrics. This systematic approach enhances the scalability and applicability of HAR systems in real-world scenarios like healthcare and fitness monitoring.

4. Experimental Results and Discussions

The performance of the CLM-GWO-DT technique on WISDM and HAR was analyzed in this section.

4.1. WISDM Dataset Description

There are 1098209 samples in the WISDM database. Walking (38.6%) and standing (4.4%) are the most and least common activities, respectively. Furthermore, the study design of WISDM involved an experimental goal of 36 participants who were required to do specific routine tasks with the Android phone placed in the pocket of their pants.

When a motion sensor was integrated into cellphones, the sensor that was put into cellphones was an accelerometer that provided a 20-Hz sampling rate. Standing (Std), sitting (Sit), walking (Walk), going upstairs (Up), coming down (Down), and jogging (Jog) were the activities that were observed. For purposes of verifying the data collected, one person was assigned the role of overseeing this process to guarantee the most incredible quality of data collection. Table 1 describes the WISDM dataset.

Table 1. Description of the WISDM dataset

WISDM Dataset			
Activity	**Samples**	**Percentage (%)**	**For Experimental**
Walk	424400	38.6	5000
Jog	342177	31.2	5000
Up	122869	11.2	5000
Down	100427	9.1	5000
Sit	59939	5.5	5000
Std	48397	4.4	5000
Total Number of Instances			**30000**

4.1.1 Analysis of the WISDM Dataset Results

The CLM-GWO-DT technique has selected five features as discriminatory features from the available six features. The CLM-GWO-DT technique’s performance on the WISDM dataset is displayed in Table 2.

All six classes were correctly classified, as shown by the confusion matrix for 70% of the dataset used for training and 30% for testing in Figures 3(a) and 3(b). The classification performance bar charts for each class are displayed in Figures 3(c) and 3(d). Figure 4 shows a brief recognition result of the CLM-GWO-DT technique on the WISDM dataset.

Figure 5 displays a thorough PR analysis of the CLM-GWO-DT methodology on the WISDM dataset. Table 3 represents the Precision-Recall (PR) curve values for the CLM-GWO-DT technique when evaluated on the WISDM dataset.

Table 2. Prediction outcome of CLM-GWO-DT technique on the WISDM dataset

Class	Accuracy	Precision	Recall	Fscore	MCC
Training Phase (70%)					
Walk	92.69	79.41	75.78	77.55	73.21
Jogging	94.55	86.43	79.78	82.97	79.82
Upstairs	94.35	83.39	82.65	83.02	79.63
Downstairs	94.25	80.43	86.86	83.52	80.13
Sitting	93.89	83.12	79.20	81.11	77.50
Standing	95.10	82.11	90.16	85.94	83.12
Average	**94.14**	**82.48**	**82.41**	**82.35**	**78.90**
Testing Phase (30%)					
Walking	92.79	79.55	76.32	77.90	73.62
Jogging	94.53	86.08	80.24	83.06	79.87
Upstairs	94.66	84.81	82.47	83.62	80.44
Downstairs	94.33	80.30	86.72	83.39	80.06
Sitting	93.97	84.17	79.13	81.57	78.02
Standing	94.94	81.31	90.82	85.80	82.93
Average	**94.20**	**82.71**	**82.62**	**82.56**	**79.16**

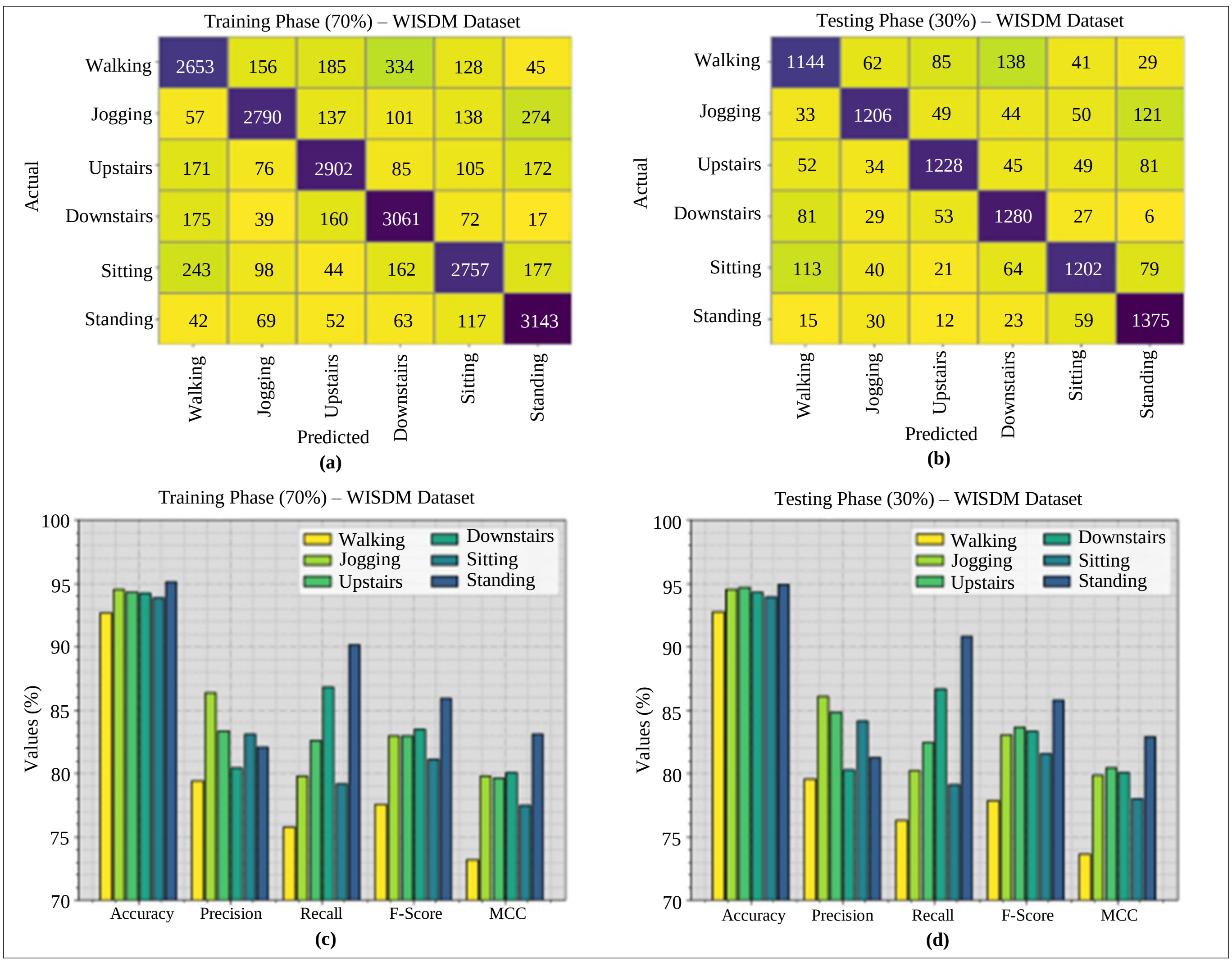

Fig. 3(a) and 3(b) Confusion matrices, (c) and (d) Classifier performance on training and testing.

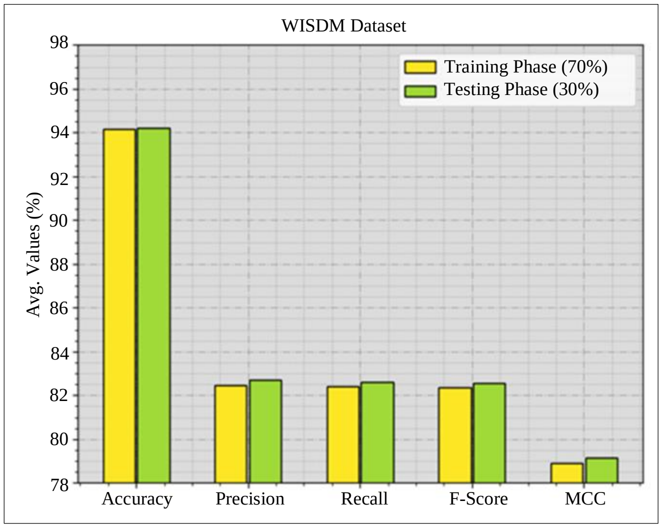

Fig. 4 Average of CLM-GWO-DT technique on WISDM dataset

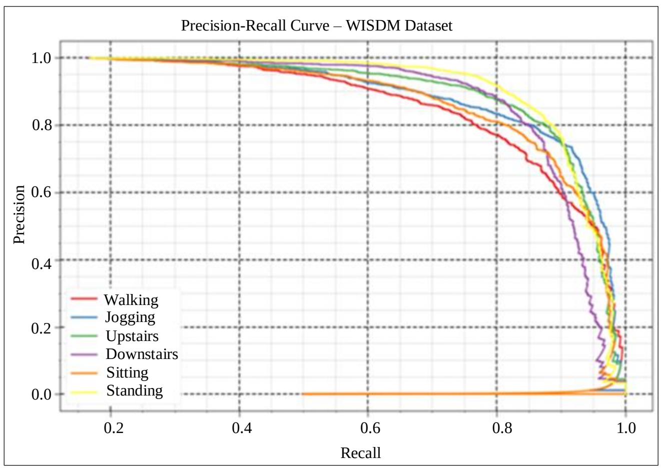

Fig. 5 PR curve of CLM-GWO-DT technique on WISDM dataset

Table 3. PR curve values of CLM-GWO-DT WISDM dataset

Activity	WISDM	
	Precision	Recall
Walking	0.9467	0.9386
Jogging	0.9666	0.9304
Upstairs	0.8499	0.9058
Downstairs	0.8799	0.8598
Sitting	1	1
Standing	1	1

4.2. UCI HAR Dataset Description

The UCI HAR dataset is a well-structured resource commonly utilized for worldwide machine learning and human activity recognition. It was created by gathering data from 30 volunteers who engaged in six physical activities: walking, walking upstairs, walking downstairs, sitting, standing, and lying down. Each participant carried a smartphone at their right waist, equipped with inertial sensors, an accelerometer, and a gyroscope. This setup continuously recorded sensor data at a fixed rate of 50 Hz for 15 seconds, resulting in a series of 3-axial linear acceleration and 3-axial angular velocity measurements. Following data collection, pre-processing took place to extract time-domain and frequency-domain features, including mean, standard deviation, Signal Magnitude Area (SMA), and Fast Fourier Transform (FFT) coefficients. The final dataset comprises 561 features associated with activities; therefore, it is appropriate for applying supervised learning algorithms. It has a training data set and a test data set in order to provide the tools for model assessment and check its effectiveness. The scientific community recognizes the UCI HAR dataset as a benchmark for evaluating different classification algorithms in activity recognition and wearable body sensor networks, as shown in Table 4.

Table 4. Description of the UCI-HAR dataset

UCI-HAR Dataset			
Activity	**Samples**	**Percentage (%)**	**For Experimental**
Walking	122091	16.30	5000
Upstairs	116707	15.60	5000
Downstairs	107961	14.40	5000
Sitting	126677	16.90	5000
Standing	138105	18.50	5000
Lying Down	136865	18.30	5000
Total Number of Instances			**30000**

4.2.1 Analysis of the UCI-HAR Dataset Results

Of the 561 features that are available, 403 have been selected by the CLM-GWO system. All six classes were correctly classified, as shown by the confusion matrix for 70% of the dataset used for training and 30% for testing in Figures 6(a) and 6(b). The classification performance bar charts for each class are displayed in Figures 6(c) and 6(d). Table 5 shows the proposed method’s classification performance on the UCI-HAR dataset.

Figure 7 shows a detailed recognition result of the CLM-GWO-DT method on the UCI-HAR dataset. A PR examination of the CLM-GWO-DT algorithm is depicted on the UCI HAR dataset in Figure 8. The outcome is defined as the CLM-GWO-DT algorithm leading to greater PR outcomes. Further, the CLM-GWO-DT system can achieve greater PR values on 6 class labels. When tested on the UCI HAR dataset, the PR curve values for the CLM-GWO-DT technique are displayed in Table 6.

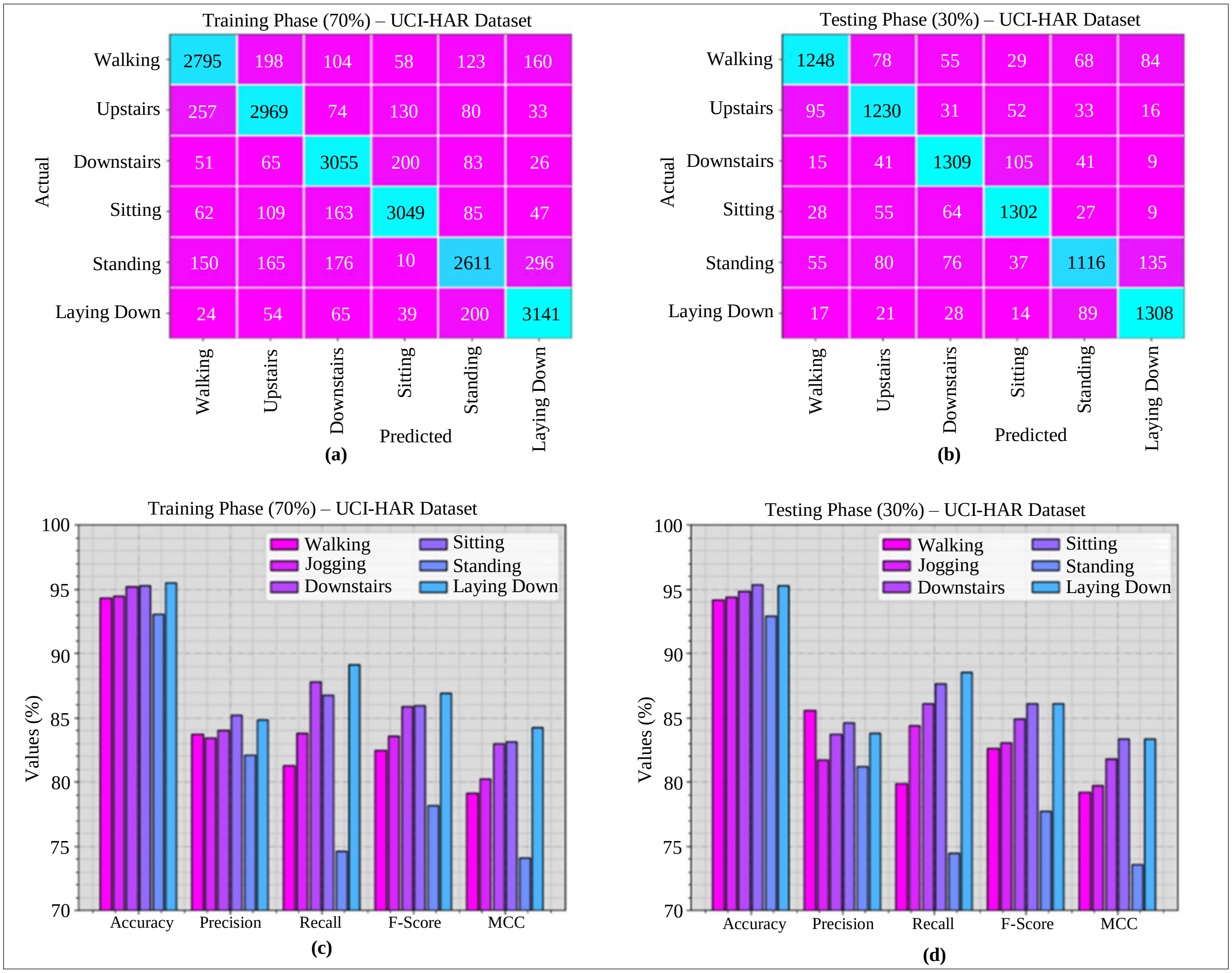

Fig. 6(a) and (b) Confusion matrices, (c) and (d) Classifier performance on training and testing dataset.

Table 5. Classification performance of CLM-GWO-DT technique on the UCI-HAR dataset

Class	Accuracy	Precision	Recall	FScore	MCC
Training Phase (70%)					
Walking	94.35	83.71	81.30	82.48	79.13
Upstairs	94.45	83.40	83.80	83.60	80.26
Downstairs	95.20	84.00	87.79	85.85	82.99
Sitting	95.26	85.19	86.74	85.96	83.11
Standing	93.04	82.06	74.58	78.14	74.13
Laying Down	95.50	84.82	89.16	86.94	84.26
Average	**94.63**	**83.86**	**83.89**	**83.83**	**80.65**
Testing Phase (30%)					
Walking	94.18	85.60	79.90	82.65	79.22
Upstairs	94.42	81.73	84.42	83.05	79.73
Downstairs	94.83	83.75	86.12	84.92	81.81
Sitting	95.33	84.60	87.68	86.11	83.33
Standing	92.88	81.22	74.45	77.69	73.56
Laying Down	95.31	83.79	88.56	86.11	83.34
Average	**94.49**	**83.45**	**83.52**	**83.42**	**80.16**

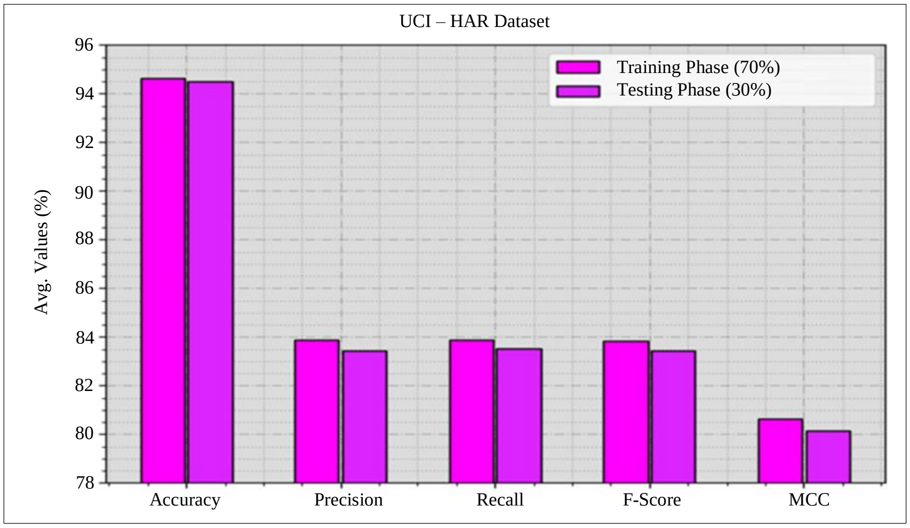

Fig. 7 Average of CLM-GWO-DT technique on UCI-HAR dataset

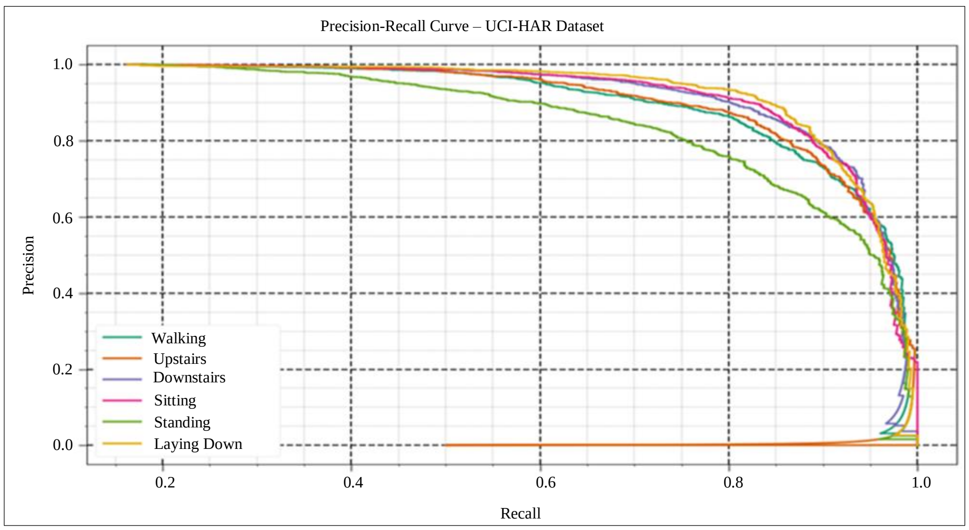

Fig. 8 PR curve of CLM-GWO-DT technique on UCI-HAR dataset

Table 6. PR curve values of CLM-GWO-DT technique on UCI HAR dataset

Activity	UCI HAR	
	Precision	Recall
Walking	0.9426	0.9147
Upstairs	0.8712	0.941
Downstairs	0.9353	0.8952
Sitting	0.908	0.9112
Standing	0.9083	0.9049
Laying Down	1	1

5. Comparison with Other Existing Methods in the Literature

Table 7 compares the deployed method with the Attention-Mechanism-Based DL Feature Combination method in the literature [24]. CLM-GWO-DT technique performs better than other approaches, with a maximal of 94.63% for UCI HAR and 94.20% for the WISDM dataset.

Table 7. Comparison of CLM-GWO-DT technique with attention mechanism-based DL feature combination

Model	Data	Accuracy(%)
Attention mechanism-based DL	WISDM UCI-HAR	93.89 93.48
CLM-GWO-DT	WISDM UCI-HAR	94.20 94.63

6. Conclusion

Therefore, this study proposes the Chaotic Logistic Map-based Grey Wolf Optimization with a Decision Tree (CLM-GWO-DT) framework for HAR using wearable sensor data. By incorporating chaotic logistic maps in the GWO methodology in the proposed method, the effort to find the right balance among the exploration and exploitation stages simplifies the task of selecting the right features from larger dimensions of sensor data. Upon operating with the DT classifier, these improved characteristics provide accurate, human-interpretable and efficient identification of human activities. UCI HAR & WISDM benchmark datasets used for the experimental evaluation demonstrate that the deployed CLM-GWO-DT surpasses the existing MCC, F-Score, recall, accuracy, and precision methodologies. In conclusion, these results corroborate the effectiveness and reliability of the presented approach for HAR applications. Besides improving the recognition accuracy, the values of the proposed CLM-GWO-DT framework decrease the computational cost; this is necessary to apply to use in reality, for instance, health care, sports science, and smart environments. In order to expand the scope and usefulness of the health of knowledge and intelligent systems, future research might build on this framework by integrating real-time systems, multimodal data sources, and more sophisticated classifiers.

References

[1] Neha Gupta et al., "Human Activity Recognition in Artificial Intelligence Framework: A Narrative Review," *Artificial Intelligence Review*, vol. 55, pp. 4755-4808, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Ahatsham Hayat et al., "Human Activity Recognition for Elderly People Using Machine and Deep Learning Approaches," *Information*, vol. 13, no. 6, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Esther Omolara Abiodun et al., "A Systematic Review of Emerging Feature Selection Optimization Methods for Optimal Text Classification: The Present State and Prospective Opportunities," *Neural Computing and Applications*, vol. 33, pp. 15091-15118, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Laith Abualigah et al., "Optimization Algorithms to Solve Feature Selection Problem: A Review," *International Journal of Science and Advanced Information Technology*, vol. 8, no. 6, pp. 66-72, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] Shitharth Selvarajan, "A Comprehensive Study on Modern Optimization Techniques for Engineering Applications," *Artificial Intelligence Review*, vol. 57, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Ahmed G. Gad, "Particle Swarm Optimization Algorithm and Its Applications: A Systematic Review," *Archives of Computational Methods in Engineering*, vol. 29, pp. 2531-2561, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Quoc-Viet Pham, "Whale Optimization Algorithm with Applications to Resource Allocation in Wireless Networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4285-4297, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[8] Emmanuel Dada et al., "Application of Grey Wolf Optimization Algorithm: Recent Trends, Issues, and Possible Horizons," *Gazi University Journal of Science*, vol. 35, no. 2, pp. 485-504, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Samad Moslehi et al., "Application of Machine Learning Models Based on Decision Trees in Classifying the Factors Affecting Mortality of COVID-19 Patients in Hamadan," *BMC Medical Informatics and Decision Making*, vol. 22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Ibomoiye Domor Mienye, Yanxia Sun, and Zenghui Wang, "Prediction Performance of Improved Decision Tree-Based Algorithms: A Review," *Procedia Manufacturing*, vol. 35, pp. 698-703, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[11] Abbas Shah Syed et al., "Using Wearable Sensors for Human Activity Recognition in Logistics: A Comparison of Different Feature Sets and Machine Learning Algorithms," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 9, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Mohamed E. Issa et al., "Human Activity Recognition Based on Embedded Sensor Data Fusion for the Internet of Healthcare Things," *Healthcare*, vol. 10, no. 6, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Daniel Garcia-Gonzalez et al., "A Public Domain Dataset for Real-Life Human Activity Recognition Using Smartphone Sensors," *Sensors*, vol. 20, no. 8, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] Ahatsham Hayat et al., "Human Activity Recognition for Elderly People Using Machine and Deep Learning Approaches," *Information*, vol. 13, no. 6, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] K. ButchiRaju, "Smart Heart Disease Prediction System with IoT and Fog Computing Sectors Enabled by Cascaded Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1-23, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Kun Xia et al., "LSTM-CNN Architecture for Human Activity Recognition," *IEEE Access*, vol. 8, pp. 55000-55009, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Federico Cruciani et al., "Feature learning for Human Activity Recognition using Convolutional Neural Networks," *CCF Transactions on Pervasive Computing and Interaction*, vol. 2, pp. 18-32, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Syed K. Bashar, Abdullah Al Fahim, and Ki H. Chon, "Smartphone-Based Human Activity Recognition with Feature Selection and Dense Neural Network," *42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Montreal, QC, Canada, pp. 5888-5891, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Yihui Qiu, Xiaoxiao Yang, and Shuixuan Chen, "An Improved Grey Wolf Optimization Algorithm Solving Functional Optimization and Engineering Design Problems," *Scientific Reports*, vol. 14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Jhila Nasiri, and Farzin Modarres Khiyabani, "A Whale Optimization Algorithm (WOA) Approach for Clustering," *Cogent Mathematics and Statistics*, vol. 5, no. 1, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[21] Saman M. Almuft, Amar Yahya Zebari, and Herman Khalid Omer, "A Comparative Study of Particle Swarm Optimization and Genetic Algorithm," *Journal of Advanced Computer Science and Technology*, vol. 8, no. 2, pp. 40-45, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[22] S. Mokhnache et al., "A Combined Chaotic System for Speech Encryption," *Engineering, Technology and Applied Science Research*, vol. 12, no. 3, pp. 8578-8583, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Talha Umar, Mohammad Nadeem, and Faisal Anwer, "Chaos-Based Image Encryption Scheme to Secure Sensitive Multimedia Content in Cloud Storage," *Expert Systems with Applications*, vol. 257, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Morsheda Akter et al., "Human Activity Recognition Using Attention-Mechanism-Based Deep Learning Feature Combination," *Sensors*, vol. 23, no. 12, 2023. [CrossRef] [Google Scholar] [Publisher Link]