Iraqi Journal of Information and Technology Systems
DOI:
ISSN 3079-580X (Online), Vol.1, No.1, P.P. 37-45
Received April 3, 2025 , accepted May 3, 2025

Developing an Optimized Model for Human Activity Recognition: A Performance Evaluation

Mohammad Khalaf Rahim Al-juaifari [1*]

[1] *College of Medicine, University of Kufa, Najaf, Iraq, Najaf, Iraq.*

Corresponding email: mohammad.aljuaifari@uokufa.edu.iq

Received April 3, 2025, Revised May 2, 2025, Accepted May 3, 2025, Published May 20, 2025

Abstract: Anticipating and recognizing human activities has significant implications across various domains, including surveillance, human-computer interaction, autonomous navigation, and robotics. The ability to classify and predict human actions precisely is crucial for enhancing system efficiency in these fields. This study proposes an optimized model that employs machine learning techniques to improve multi-label classification for Human Activity Recognition (HAR). By integrating a confusion matrix with an Extra Tree Classifier, the model enhances activity recognition and the prediction of subsequent actions. To facilitate a comparative analysis between deep learning and machine learning approaches, the UCI-HAR dataset, consisting of six interactive attributes, is utilized. The Multiclass-Classification SVM model achieves an accuracy of 0.9778, whereas the Artificial Neural Network (ANN) model, optimized with ADADELTA, attains an accuracy of 0.9984. A comprehensive evaluation of existing recognition methodologies confirms the proposed model's robustness and efficiency

Due to their high accuracy, low response time, and adaptability to real-time sensor inputs, the proposed models hold potential for implementation in wearable health monitoring, elderly fall detection systems, and intelligent home security solutions.

Keyword: Action Recognition, Deep Learning (DL), Machine Learning (ML), Wearable Sensors, Activities of Daily Living (ADL)

1. INTRODUCTION

Various ways were developed to solve the trend problems from traditional [1] to use statistics to solve the health issue [2] to using Artificial intelligence such as deep learning in [3]. Newborns mortality in [4] creates procedures that limit these problems in the future. The use of the Naïve Bayes algorithm and its inclusion in an electronic system to monitor maternal health by generating alerts that indicate risks and patient health monitoring. Human Action Recognition & Prediction encompasses four key dimensions: a) Action Recognition b) Action Prediction c) Action localization and detection [5] and d) Motion Trajectory Predictions [6], employing deep learning techniques.

The paper's structure begins with an introductory section, elucidating objectives and various facets of human activity recognition and prediction. A comprehensive literature review highlights relevant studies employing diverse methods and datasets. The methodology segment encompasses data pre-processing, employing the UCI-HAR dataset, The proposed model's architecture is outlined, encompassing its operational steps, while the evaluation phase substantiates its efficiency.

HAR and HAP encompass fields such as signal processing, AI, and robotics. However, these areas differ in terms of their objectives and applications.

HAR refers to the technology that involves extracting and examining data related to human movement and performance, acquired from various sources like acoustic signals, movement, or vibrations. The primary focus of HAR is to identify and categorize different human activities, such as walking, running, cycling, standing, and more. The applications of human activity recognition are diverse, including fitness tracking, healthcare, user behaviour analyze, and smart systems.

In essence, HAR serves as an initial step in analyzing present activities, while HAP focuses on anticipating future activities, the relation between HAP and HAR is shown clearly in the following Figure 1.

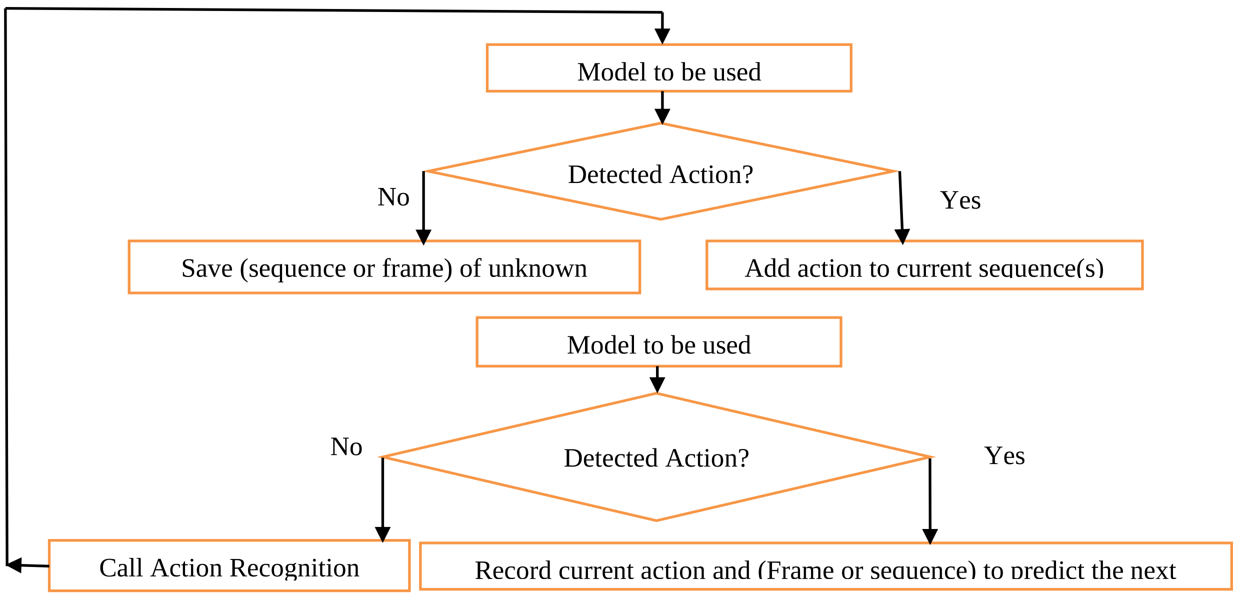

Figure 1- Activity Prediction Vs Recognition

The experimentation component delves into data models, presenting the results of various methods with their accuracies. Ambiguous equations are clarified, ensuring reader comprehension. Concluding this work is a section encapsulating its contributions while highlighting the significance of:

1- DL proposed an approach to pre-processing by DWT (db4) with a TCNN model to predict step(s) validated by a confusion matrix.

2- Implementation of Artificial Neural Network (ANN) by enhanced optimizer.

3- Machine learning (ML) Multi-class with almost all classification algorithms. Integration classifier for activity recognition.

All three contributions can be used as recognized activities and predicted by calculating the confusion matrix as in DL's proposed approach to this work. The contributions extend to fostering a deeper understanding of efficient human activity recognition and prediction techniques, noteworthy strides in refining action prediction methodologies, and comparative analyzing results to introduce optimal methods.

2. RELATED WORK

Various model demonstrates significant performance improvements, outperforming state-of-the-art methods and underscoring their effectiveness in intelligent sensor systems.

There are many applications and methods in the domain of human-computer interaction, and early approaches [7] utilized computer vision for human motion prediction. The author in [8] An automated CNN and dataset UCI-HAR performs an average accuracy of 98.5 ($\mp$1.1) with hyperparameters nb_filters nb_blocks Alpha.

The dataset's description [9] is based on the recordings of 30 study participants between the ages of 19 and 48 who performed efficient UCI-ADL.

Traditional HAR approaches utilize machine learning algorithms for classification, such as support vector machines (SVMs) [10], deep learning approach using LSTM for human activity prediction [11], random forest [12], k-nearest neighbors (KNN) [13], and hidden Markov models (HMM) [14]. However, these methods have certain limitations in deep feature extraction, especially for complex action recognition tasks. Their performance is often restricted and requires domain knowledge and manual feature extraction.

In recent years, with the rise of deep learning technologies, deep neural networks have excelled in the HAR field and gradually become the mainstream approach [15], evaluated by accuracy as follows: CNN-LSTM model in [16] is 92.13, the Bi-dir-LSTM model in [17] is 92.67, LSTM-CNN model [18] is 95.78, and CNN-GRU model in [19] is 96.20, Moreover, Huafeng, G. et al [20] proposed a method to reduce data bias and enhance the accuracy of human activity recognition by employing a CNN and a bidirectional LSTM to automatically extract features from raw sensor data, combined with a self-attention mechanism to identify key time point relationships. The integration of self-attention mechanisms within deep learning frameworks notably boosts recognition accuracy.

3. METHODOLOGY

This section contains data description, details about the proposed model, and equations. Preprocessing (as will be stated in Algorithm 1) extract frequency using WDT involves signal decomposition.

A. Data

The dataset used "UCI-HAR is publicly available and de-identified"[9] due to a transfer learning rate of 50Hz which speculates human motion axes (x,y,z) with 3-axial linear acceleration and 3-axial angular velocity; more than 563 columns entries and volunteers generate 561-feature of attributes that produce (70%) of the training data and (30%) test data, also manually video recorded to label the data, the dataset contains the activity of (WALKING, WALKING-(UPSTAIRS, WALKING-DOWNSTAIRS, SITTING, STANDING, LAYING) carrying a waist-mounted smartphone (Samsung Galaxy S II) with embedded accelerometer and gyroscope (inertial sensors) to classify activities into one of the six activities achieved.

The dataset determines the human current actions by three-dimensional positioning sensors where data is pre-processed by noise filter (low-pass filter), source of data is an accelerometer and gyroscope sensor, time domain and frequency domain used to determine variable to get feature vector.

B. Proposed Model:

This paper proposes HAR and HAP based on the following techniques:

1) DL Proposed Model:

architectural aspects of the full map for our proposed methodology in Figure 2 below shows the recognition and prediction for human activity, loading data, and step(s) of the pre-processing agenda had been eliminated to focus on the most effective elements in conduct modelling, the following steps Transformer Aspects for Proposed DL Modelling: import libraries, load dataset, define feature matrix 'X' and target labels 'yvalues', apply wavelet transform to target labels, convert 'Class' column to numeric and drop rows with NaN values, separate target variable ('Class') from feature set 'X', Scale the feature matrix 'X', split the data into training and test sets, proposed a 1D-TCNN model: add LSTM, convolutional, max pooling, flatten, dense, dropout, and output layers, compile the model with appropriate loss function and optimizer, define early stopping criteria, and fit the model using training data, validate on test data, and include early stopping. The steps can be represented as algorithm 1 below:

Algorithm 1: DL proposed Approach

1: Let D represent the UCI HAR dataset.
2: Input: UCI-HAR dataset
3: Output: Predicted step
4: X, y ← ExtractFeatures(D)
5: ywavelet ← DWT(y, 'db4')
6: Plot(ywavelet)
7: D ← Preprocess(D)
8: X_train, X_test, y_train, y_test ← SplitData(X, y)
9: X_scaled ← Scale(X_train)
10: Model ← BuildModel()
11: Compile(Model)
12: EarlyStopping ← DefineEarlyStopping()
13: Fit(Model, X_train, y_train, X_test, y_test, callbacks = [EarlyStopping])

2) ANN Proposed Methods

ANN is used to resolve non-linear problems, proposed ANN model flows can be described as importing libraries, loading datasets, data pre-processing, dropping rows with missing or improperly formatted data, converting columns to numeric type if applicable, Separate features (X) and targeting (y) columns, label encoding, Encode categorical labels using LabelEncoder, Convert encoded labels to one-hot encoded vectors using to_categorical, Train-Test Split: split the dataset into training and testing sets using train_test_split, Feature Scaling: standardize the features using StandardScaler, Model Definition:
(Create a Sequential model, Add Dense layers with ReLU activation, Add Dropout layers to prevent overfitting, Final layer has SoftMax activation for multi-class classification), compile Model:
Compile the model using the Adam optimizer with a specified learning rate.
Use categorical cross-entropy as the loss function. Track accuracy as a metric. 2- Model Training, Fit the model to the training data using a validation split of 30%.
Train for 75 epochs (iteration steps) with a batch size of 64, Model Evaluation: Evaluate the trained model on the test data, Print test loss, and accuracy.
The proposed model described in Figure 3 used in this work is as follows:
Each node in a dense layer computes a weighted sum of all node values from the preceding layer and adds a bias term as in equation 1 below:

$$z[i]=W[i]a\,[i-1] + b[i] \qquad \text{……………………….…(1)}$$

optimizer can be represented mathematically as:
Compile model

$$\text{optimizer} = \text{Adam}(\text{learning\_rate}=0.001) \qquad \text{…….(2)}$$

Where we replace W(t) with W new and W(t-1) with W(old) in the original equation (2).
Adagrad power in using different learning rules in every hidden neuron, layer, iteration, and weight.
Equation 3 is Adagarad (Adaptive Gradient Algorithm) optimization with a small positive value ε to prevent dividing by zero in case l_t becomes zero.

$$\eta\_t\hat{}\backslash = \eta \,/ \sqrt{}\,(l\_t+ \varepsilon) \qquad \text{………………………..(3)}$$

ADADELTA solves the two problems of Adagrad; the first is preventing l_t to be increased so big due to increasing in every iteration, and the second is preventing η_t^\ from becoming small but not very small whenever Adagrad becomes deeper, to become as in equation 4 using W_avg (weighted average).

$$\eta\_t\hat{}\backslash=\partial l/ \sqrt{}\,(W\_avg+\varepsilon) \qquad \text{…………………….…(4)}$$

3) *ML algorithms*

The following steps show the steps to use to show results using ML algorithms.

1- The first step after loading a dataset, using the classify for training data is to split it into train data and test data, fitting data and score then returning it to the classifier to complete the loop for the classifier.

2- Muli-class SVM classifier: The dataset used smartphones is a feature vector with high dimensional so SVM is a powerful method.

3- KNN classifier: KNN training data are examined depending on Euclidean distance to neighbor feature vectors and not subject to the data structure.

4- All algorithm descriptions are stated in Table 1.

TABLE 1- Key Principles of Each Algorithm's Description

Algorithm	Description
Logistic Regression	Models the probability of a binary outcome based on predictor variables using the logistic function.
Decision Trees	Recursively splits data based on attribute values to create homogeneous subsets concerning the target variable.
Random Forest	Constructs multiple decision trees during training and outputs the mode of classes or mean prediction of individual trees, reducing overfitting with bagging.
k-Nearest Neighbors	Assigns a data point to the majority class among its k nearest neighbors based on distance computation and majority voting.
Gradient	Builds a sequence of decision trees, with each correcting the errors of the previous one using gradient

Boosting Machines	descent optimization.
AdaBoost	Combines weak learners to create a strong classifier, adjusting weights at each iteration to focus on difficult instances and using weighted majority voting.
XGBoost	An optimized gradient boosting algorithm that builds decision trees sequentially, focusing on computational efficiency, regularization, and advanced optimization techniques.
Logistic Regression	models the probability of a binary outcome based on one or more predictor variables. It uses a logistic function to model the binary outcome and a decision boundary to separate classes.
Decision Trees	Decision trees recursively split the data into subsets based on the value of attributes, aiming to create homogeneous subsets concerning the target variable. Each node, the algorithm selects the attribute that best splits the data, based on criteria like information gain.
Random Forest	is an ensemble learning method that constructs multiple decision trees during training and outputs classification or regression. It uses bagging to train each tree on a random subset of data and randomly selects a subset of features at each split to reduce overfitting.

4. EXPERIMENTS RESULT

Methods performed on the dataset can be described by loading data first, next applying preprocessing to remove noise, then to the classifier to split it into train and test data with fitting and scoring then returning them to the classifier to complete the loop, and finally step doing multi-class SVM, KNN, and confusion matrix as shown in Table 1.

4.1. Experiments Results of ML

Figure 2 shows performance results for various ML algorithms applied with the most well-known algorithms:

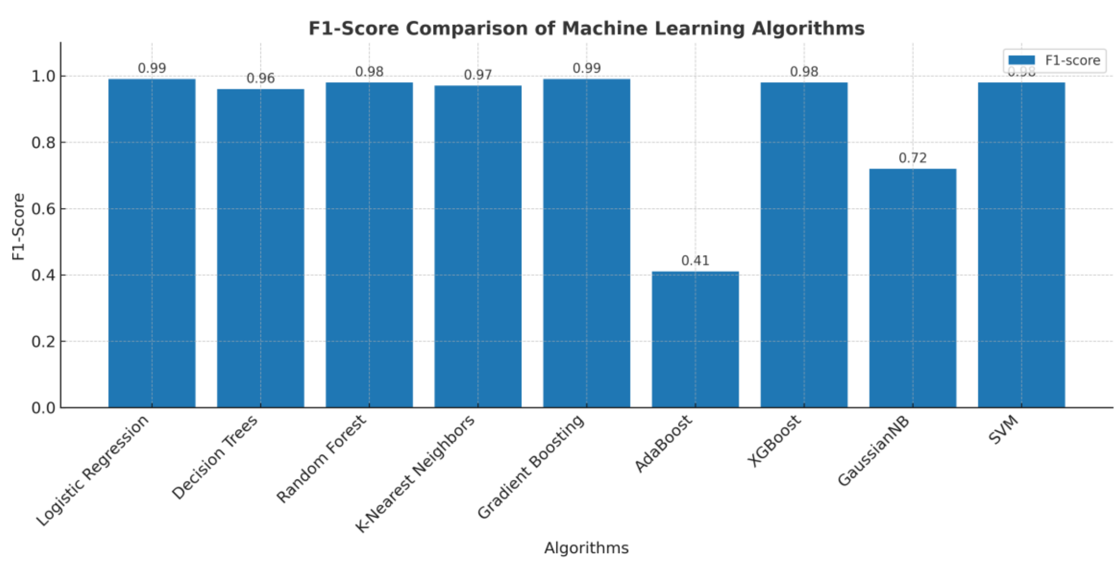

Figure 2- Performance Comparison of ML Algorithms

Table 2 shows the F1-score for the ML algorithm.

Table 2- F1 – Score for Each ML Algorithm

Algorithm	Logistic Regression	Decision Trees	Random Forest	k-Nearest Neighbors	Gradient Boosting	AdaBoost	XGBoost	Gaussian NB	SVM
F1-score	0.986	0.958	0.983	0.967	0.989	0.407	0.981	0.717	0.980

4.2. Experimental Results of ANN

Detailed analyze of the neural network model's performance during the last epoch of training, in which the neural network model achieved a training loss of 0.0273 and a training accuracy of 0.9906. These metrics indicate the model's proficiency in minimizing errors and accurately classifying training instances. Evaluation of the validation dataset yielded a validation loss of 0.0312 and a validation accuracy of 0.9901, demonstrating the model's generalization capability. Despite a slight increase in loss compared to training, the validation metrics affirm the model's robustness and reliability in making accurate predictions on unseen data.

Table 3- Experimental Results and Training Parameters

Parameters	**test_size=0.3**	**test_size=0.2**
Epoch	75	75
accuracy: val_accuracy:	0.9886 0.9877	0.9906 0.9901
loss: val_loss:	0.0340 0.0433	0.0273 0.0312
Test loss: Test accuracy:	0.038 0.989	Test loss: 0.0375 Test accuracy: 0.9885
F1 Score:	0.989	0.988
Confusion Matrix:	[[724 0 2 0 0 0] [2 636 0 0 0 0] [2 1 586 0 0 0] [0 0 0 701 16 0] [0 0 0 21 726 0] [0 0 0 0 0 762]]	[[506 1 1 0 0 0] [0 431 0 0 0 0] [2 1 384 0 0 0] [0 0 0 445 15 0] [0 0 0 12 479 0] [0 0 0 0 0 509]]

Table 4 shows proposed ANN involves a number of layers, units per layer, dropout rate, activation functions, optimizer (Adam, Adadelta), lr =0.001, Number of epochs =75=, batch size=64, and validation split =0.3. while in DL the number of LSTM layers, filter sizes, kernel size, pooling size, activation functions, dropout, optimizer, loss function, and early stopping parameters.

Table 4: ANN and DL Hyperparameters

Model	**Layers**	**Optimizer**	**Learning Rate**	**Epochs**	**Batch Size**	**Dropout**
ANN	3 Dense layers	Adadelta	0.001	75	64	0.3
DL	1 LSTM, 1 Conv1D, MaxPooling, Dense	Adam	0.001	50	64	0.25

4.3. Experimental Results of DL

Proposed Model results are discussed as follows:

4.3.1 Confusion Matrix:

The confusion matrix, an essential evaluation metric, quantifies the frequency of misclassifications among predicted classes during model testing. Employing for training, a confusion matrix was generated, reflecting the confusion rates among the four categories across 30% of the test dataset. calculated by epochs = 50 and time of training is 6s/epoch - 20ms/step. The confusion matrix is demonstrated in Figure 3 below to show the confusion matrix, the diagonal entries correspond to accurately predicted activity classes. Off-diagonal entries that are zero or close to zero suggest high classification accuracy, reflecting minimal errors in distinguishing between activities like sitting, standing, and walking.

	Predicted Activity					
True Activity	722	0	0	0	0	0
	0	665	1	0	0	0
	2	1	598	0	0	0
	0	0	0	710	10	1
	0	0	0	19	676	0
	0	0	0	0	0	774

Figure 3- Confusion Matrix of HAR

4.3.2 Comparison with other Work:

Table 5 contains the key considerations and results of other related work, the enhanced performance of the proposed ANN and TCNN models can be attributed to the use of advanced optimizers and effective preprocessing techniques such as wavelet transforms. Nonetheless, models like CNN-GRU and CNN-BiLSTM are often more effective at modeling temporal patterns in real-time scenarios or continuous sensor streams, although they typically require greater computational resources. Considerations and results of other related work.

Table 5- Comparison of Results with other Methods

Paper	**Model**	**Explanation**	**Performance (%)**
[21] 2021	1-D CNN + Attention	The integration of the attention mechanism augmented the model's focus on pertinent aspects of input data. The impressive accuracy, precision, recall, and F1 score collectively indicate the attention mechanism's efficacy.	Accuracy = 98.18, Precision = 97.12, Recall = 97.29, F1 Score = 97.20
[22] 2021	CNN-GRU	Exploiting CNN's feature extraction and GRU's sequence modelling prowess.	Accuracy =96.20
[23] 2021	CNN-GRU CNN-LSTM CNN-Bi-LSTM	These hybridized variants, blending CNN with diverse recurrent layers, displayed no table accuracy levels. Such performance signifies the effective processing of sensor data through these amalgamated configurations.	Accuracy =94.58 Accuracy =94.80 Accuracy =96.37
[24] 2020	CNN	Renowned for its capacity to capture spatial patterns, CNN's accuracy underscores its adeptness in distinguishing between distinct activities.	Accuracy =92.71
[25] 2018	Res-LSTM	The amalgamation of ResNet's skip connections and LSTM's memory capabilities played out positively in recognizing activities, by the achieved accuracy.	Accuracy =91.60
[26] 2022	ELA model	ELA's focus on attentive aspects at the event level illustrated its capability to capture pivotal events within the activity recognition context.	Accuracy =97.70%
2022 [37]	CNN-GRU49	This specific adaptation of CNN-GRU with an augmented hidden layer count showcased a compelling accuracy.	Accuracy =96.71%
[28] 2023	DDF	DDF's dynamic fusion of data from diverse sources proved effective, as reflected in the heightened accuracy, thus enhancing activity recognition.	Accuracy: 97.81%
[29] 2023	Bi-HAR	No. of Parameters is 15,017,152	Accuracy 97.89 F1-Score 96.47
[30] 2023	ResNet+HC	No. of Parameters is 0.42 M 1	Accuracy 97.01
[31] 2023	GRU-INC	No. of Parameters is 666,112	Accuracy 96.27 F1-Score 96.26
[32] 2023	DMEFAM	No. of Parameters is 1.6 M 1	Accuracy 96.00 F1-Score 95.80
[33] 2024	RMFSN	No. of Parameters is 239,846	Accuracy 98.13 F1-Score 98.21
ANN proposed Model	2- Dense layers	The proposed model's commendable performance metrics underscore its viability as a robust contender for activity recognition.	Test loss: 0.041 Test accuracy 0.987 F score: 0.9705 Val_loss: 0.0078
DL proposed Model	1D DL model	Pre-processing time and complexity	Epoch 50, 305/305 - 6s - loss: 0.0216 accuracy: 0.9915 - val_loss: 0.0448 val_accuracy 0.9919

Results clearly show how to reduce the number of features used in the data set with pre-processing before using the model with various methods of ML algorithms, …..etc, and evaluation confusion matrix. Evaluation of the model has approved the efficiency of our models with the same techniques and/or the same datasets.

5. CONCLUSION

After implementing several ML and DL methods for human activity recognition, DL ANN with the proposed optimizer significantly improves performance. It outperforms state-of-the-art methods, by exploring and dealing with various approaches of ML. On the other hand, in tunning Adam optimizers a parameter gets updated by changing the effectiveness of parameters during training to optimize the accuracy rather than using traditional Adam optimizers in almost related works. In future work, proposed methods will be applied to similar sensor-based data sources.

A notable limitation lies in the preprocessing stage using wavelet transforms (DWT), which, if not carefully configured, may lead to the loss of critical information or introduce biases specific to the dataset. Additionally, the model selection process prioritized accuracy without employing cross-validation across different sensor setups, which could restrict the model's ability to generalize effectively to other datasets or hardware platforms.

Future research will focus on utilizing multi-modal datasets that incorporate inertial sensor data alongside audio or video inputs to enhance recognition accuracy. Moreover, fusing information from diverse sources such as smartphones and wearable devices may improve the model's adaptability and enable more tailored applications in healthcare monitoring

DATA AVAILABILITY STATEMENT

The data used are included in the manuscript

CONFLICT OF INTEREST

The authors declare no conflict of interest in influencing the work in this paper

6. REFERENCES

[1] M. K. Rahim, "Unified patient information system for Iraqi hospitals," *Int. J. Comput. Appl.*, vol. 131, no. 4, pp. 11–14, Dec. 2015.
[2] M. K. R. Al-juaifari, A. Abdulmohson, and R. H. A. Al_Sagheer, "Iraqi newborns mortality prediction using naive Bayes classifier," *Int. J. Res. Eng. Innov. (IJREI)*, vol. 5, no. 2, pp. 123–127, 2020.
[3] M. K. R. AL-JUAIFARI and A. A. ATHARI, "Future human activity prediction using wavelet and LSTM," *J. Duhok Univ.*, vol. 26, no. 2, pp. 541–550, 2023.
[4] M. K. R. Al-juaifari, A. Abdulmohson, and R. H. A. Al_Sagheer, "Iraqi newborns mortality prediction using naive Bayes classifier," *Int. J. Res. Eng. Innov. (IJREI)*, vol. 5, no. 2, pp. 123–127, 2020.
[5] Y.-W. Chao, S. Vijayanarasimhan, B. Seybold, D. A. Ross, J. Deng, and R. Sukthankar, "Rethinking the Faster R-CNN architecture for temporal action localization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, 2018, pp. 1130–1139, doi: 10.1109/CVPR.2018.00124.
[6] N. Lee *et al.*, "DESIRE: Distant future prediction in dynamic scenes with interacting agents," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, 2017, pp. 2165–2174, doi: 10.1109/CVPR.2017.233.
[7] J. Martinez, M. Black, and J. Romero, "On human motion prediction using recurrent neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, 2017, pp. 4674–4683, doi: 10.1109/CVPR.2017.497.
[8] W. N. Ismail, H. A. Alsalamah, M. M. Hassan, and E. Mohamed, "AUTO-HAR: An adaptive human activity recognition framework using an automated CNN architecture design," *Heliyon*, vol. 9, no. 2, e13636, 2023.
[9] E. Bulbul, A. Cetin, and I. A. Dogru, "Human activity recognition using smartphones," in *Proc. 2nd Int. Symp. Multidiscip. Stud. Innov. Technol. (ISMSIT)*, Ankara, Turkey, 2018, pp. 1–6, doi: 10.1109/ISMSIT.2018.8567275.
[10] N. Ahmed, J. I. Rafiq, and M. R. Islam, "Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model," *Sensors*, vol. 20, 317, 2020. doi: 10.3390/s20010317.
[11] M. K. R. Al-juaifari and A. Athari, "A novel framework for future human activity prediction using sensor-based data," *Int. J. Intell. Eng. Syst.*, vol. 16, no. 6, 2023.
[12] S. Balli, E. A. Sağbaş, and M. Peker, "Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm," *Meas. Control*, vol. 52, no. 1–2, pp. 37–45, 2019, doi: 10.1177/0020294018813692.
[13] A. Tharwat, H. Mahdi, M. Elhoseny, and A. E. Hassanien, "Recognizing human activity in mobile crowdsensing environment using optimized k-NN algorithm," *Expert Syst. Appl.*, vol. 107, pp. 32–44, 2018, doi: 10.1016/j.eswa.2018.04.017.
[14] M. Biagi, L. Carnevali, M. Paolieri, F. Patara, and E. Vicario, "A continuous-time model-based approach for activity recognition in pervasive environments," *IEEE Trans. Hum.-Mach. Syst.*, vol. 49, no. 4, pp. 293–303, Aug.

2019, doi: 10.1109/THMS.2019.2903091.
[15] K. Hu, J. Jin, F. Zheng *et al.*, “Overview of behavior recognition based on deep learning,” *Artif. Intell. Rev.*, vol. 56, pp. 1833–1865, 2023, doi: 10.1007/s10462-022-10210-8.
[16] R. Mutegeki and D. S. Han, “A CNN-LSTM approach to human activity recognition,” in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Fukuoka, Japan, 2020, pp. 362–366, doi: 10.1109/ICAIIC48513.2020.9065078.
[17] F. Hernández, L. F. Suárez, J. Villamizar, and M. Altuve, “Human activity recognition on smartphones using a bidirectional LSTM network,” in *Proc. XXII Symp. Image Signal Process. Artif. Vis. (STSIVA)*, Bucaramanga, Colombia, 2019, pp. 1–5, doi: 10.1109/STSIVA.2019.8730249.
[18] K. Xia, J. Huang, and H. Wang, “LSTM-CNN architecture for human activity recognition,” *IEEE Access*, vol. 8, pp. 56855–56866, 2020, doi: 10.1109/ACCESS.2020.2982225.
[19] N. Dua, S. N. Singh, and V. B. Semwal, “Multi-input CNN-GRU based human activity recognition using wearable sensors,” *Computing*, vol. 103, pp. 1461–1478, 2021, doi: 10.1007/s00607-021-00928-8.
[20] H. Guo, X. Changcheng, and C. Shiqiang, “Wearable sensors for human activity recognition based on a self-attention CNN-BiLSTM model,” *Sensor Rev.*, ahead-of-print, 2023.
[21] Z. N. Khan and J. Ahmad, “Attention induced multi-head convolutional neural network for human activity recognition,” *Appl. Soft Comput.*, vol. 110, 107671, 2021, doi: 10.1016/j.asoc.2021.107671.
[22] N. Dua, S. N. Singh, and V. B. Semwal, “Multi-input CNN-GRU based human activity recognition using wearable sensors,” *Computing*, vol. 103, pp. 1461–1478, 2021, doi: 10.1007/s00607-021-00928-8.
[23] S. K. Challa, A. Kumar, and V. B. Semwal, “A multibranch CNN-BiLSTM model for human activity recognition using wearable sensor data,” *Vis. Comput.*, vol. 38, pp. 4095–4109, 2022, doi: 10.1007/s00371-021-02283-3.
[24] S. Wan, L. Qi, X. Xu *et al.*, “Deep learning models for real-time human activity recognition with smartphones,” *Mob. Netw. Appl.*, vol. 25, pp. 743–755, 2020, doi: 10.1007/s11036-019-01445-x.
[25] Y. Zhao, R. Yang, G. Chevalier, X. Xu, and Z. Zhang, “Deep residual bidir-LSTM for human activity recognition using wearable sensors,” *Math. Probl. Eng.*, vol. 2018, Art. no. 7316954, 2018.
[26] T.-H. Tan, J.-Y. Wu, S.-H. Liu, and M. Gochoo, “Human activity recognition using an ensemble learning algorithm with smartphone sensor data,” *Electronics*, vol. 11, no. 3, 322, 2022.
[27] C. Zhang, K. Cao, L. Lu, and T. Deng, “A multi-scale feature extraction fusion model for human activity recognition,” *Sci. Rep.*, vol. 12, no. 1, Art. no. 20620, 2022.
[28] Y. Zhang *et al.*, “Smartphone sensors-based human activity recognition using feature selection and deep decision fusion,” *IET Cyber-Phys. Syst. Theory Appl.*, vol. 8, no. 2, pp. 76–90, 2023.
[29] K. Venkatachalam, Z. Yang, P. Trojovský, N. Bacanin, M. Deveci, and W. Ding, “Bimodal HAR—An efficient approach to human activity analysis and recognition using bimodal hybrid classifiers,” *Inf. Sci.*, vol. 628, pp. 542–557, 2023.
[30] C. Han, L. Zhang, Y. Tang, W. Huang, F. Min, and J. He, “Human activity recognition using wearable sensors by heterogeneous convolutional neural networks,” *Expert Syst. Appl.*, vol. 198, 116764, 2022.
[31] P. Kumar, S. Chauhan, and L. K. Awasthi, “Human activity recognition (HAR) using deep learning: Review, methodologies, progress and future research directions,” *Arch. Comput. Methods Eng.*, vol. 31, no. 1, pp. 179–219, 2024.
[32] T. Mim, T. Rahman, M. Amatullah *et al.*, “GRU-INC: An inception-attention based approach using GRU for human activity recognition,” *Expert Syst. Appl.*, vol. 216, 119419, 2023.
[33] F. Zeng, M. Guo, L. Tan, F. Guo, and X. Liu, “Wearable sensor-based residual multifeature fusion shrinkage networks for human activity recognition,” *Sensors*, vol. 24, no. 3, 758, 2024.