AN EXTRA TREE ENSEMBLE OPTIMIZATION-BASED DEEP LEARNING FRAMEWORK FOR HUMAN ACTIVITY RECOGNITION

2024 5th International Conference on Innovative Trends in Information Technology (ICITIIT) | 979-8-3503-8681-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICITIIT61487.2024.10580296

Dr.Monika Arya
Dept. of CSE
CMR Engineering College
Hyderabad, Telangana,India
arya.monika0804@gmail.com

Dr. Hanumat Sastry G
School of Computer Science
UPES
Dehradun, Uttarakhand, India
sastrygh2000@yahoo.com

Akhilesh Gaidhane
Lecturer, Mathematics
SAGES M.G.Ward
Kondagaon, Chhattisgarh, India
akhil.engg2102@gmail.com

Dr. Anand Motwani
School of Computing Science & Engineering
VIT Bhopal University
Sehore MP,India
anand.motwani@vitbhopal.ac.in

***Abstract*— Human activity recognition (HAR) is a burgeoning field of study due to its real-life applications in the medical field, the e-health system, and elder care or care of physically impaired people in a smart healthcare environment. Using sensors built into wearable devices, such as smartphones, HAR provides an opportunity to identify human behavior and better understand an individual's health. In the past years, conventional machine learning techniques have made significant progress in HAR. However, these methods significantly rely on conventional feature extraction, which may hinder the effectiveness of the generalization model. The other challenges these methods face for HAR are performance degradation of many models with the increase in the number of activities, lack of capability of models to capture different layers of activities, high dimensionality, overfitting, and optimization problems. These challenges attracted numerous people to explore this field. The development in deep learning techniques has addressed most of these problems by automatically extracting discriminative features from raw input sequences obtained by multimodal sensing devices to acknowledge human activities accurately. This paper proposes an extra tree ensemble optimization-based deep learning framework (DELETO) for HAR. The model shows better results than recent and conventional techniques, with an accuracy of 98.7% and a computation time of 2.533 seconds.**

Keywords— human activity recognition, multiclass classification, deep learning, feature extraction, Data stream, machine learning.

I. INTRODUCTION

HAR can be applied in healthcare to build a smart environment and provide personalized support. A multiclass classifier for HAR is trained to recognize multiple activities like walking, running, sitting, etc. Multiclass classification is classifying or assigning an instance to only one class out of a set of three or more previously defined output classes. The machine learning (ML) technique has been traditionally used in the multiclass classification problem where the number of parameters was relatively smaller. However, the problems increased with the number of features, and consequently, the hyperspace grew exponentially with the increase in parameters. Selecting the best features can reduce dimensionality and increase accuracy [1]. Thus, conceivable feature selection methods like wrappers have improved accuracy by selecting the relevant feature subset, thereby enhancing the classifier's performance by reducing dimensionality. However, the wrapper approach is computationally more intensive, as the feature evaluation is done internally by the inductive algorithm in an iterative manner and effectively captures relationships among multiple features. Thus, feature selection is the first and most crucial step in building the model. Feature selection identifies the characteristics that add most to the interesting output. Feature selection thus has the following benefits:

1. Feature selection can reduce redundant data and, hence, reduce overfitting.
2. By selecting relevant features, model accuracy improves with fewer misleading data.
3. It also decreases the time for training and reduces the algorithm's complexity.

The proposed model utilizes the extra tree ensemble technique for selecting relevant features. This approach breaks nodes into arbitrary feature sets to create numerous decision trees. The decision trees take samples without replacing them, and the nodes are split randomly rather than the optimal splits, leading to a random subset of the features present at each node in the tree. The objectives of this paper are summarized as follows: -

Objective 1: To study the background and relevant literature.
Objective 2: To propose a methodology with an extra tree ensemble technique for selecting relevant features.
Objective 3: To implement the proposed model and evaluate the performance with parameters like accuracy, precision, recall, F1 Score, and computation time.
Objective 4: To compare the performance of the proposed model with existing techniques.

The research is broken up into the following parts: In Section 2, the background of the identified problem in the HAR domain is studied and relevant literatures are reviewed to find appropriate and feasible solutions. In section 3, the Materials and methods used in the proposed work are discussed. In this section the details of the Dataset and all the related operations for data preparation are elaborated. Section 4 discusses the proposed methodology of the suggested framework. In this

Authorized licensed use limited to: South Eastern Technological University (SETU). Downloaded on August 18,2025 at 23:25:39 UTC from IEEE Xplore. Restrictions apply.

section the algorithm and flowchart of the proposed model is explained. Section 5 is experiments and results. This section discusses the experimental setup, Analysis of results and comparison of proposed model with existing techniques. The conclusion and recommendations for additional research are included in section 6 as well.

II. BACKGROUND STUDY AND LITERATURE REVIEW

Researchers have used several techniques, including Machine learning, Deep learning Ensemble learning to recognize human behaviors [2]. Considerable study has been done on this topic. In this paper, we give a thorough analysis of the research done between 2018 and 2022 in several fields of human activity recognition, with a particular emphasis on deep learning methods for HAR.

Authors in [3] employed the enhanced deep-learning technique for HAR in IoHT environments.

A. Chaudhary et al. [1] proposed a random forest classifier and instance filter method for multiclass disease classification. V. Panca et al. [2] suggested a feature selection method deployed on the repetitive feature removal technique of the support vector machine (SVM-RFE) principle. It improved multiclass classification by combining the results of multiple classifiers. In their work, A. Beygelzimer et al. [3] reduced k-class classification to binary classification using the filter tree method. W. La Cava et al. [4] suggested a new multiclass classification_method using genetic programming to learn multidimensional feature transformations. J. Zhang et al. [5] aggregated seven individual classifiers to form an ensemble machine learning technique. Class hierarchy can be generated using the confusion matrix of a multiclass classifier, as proposed by D. Silva-Palacios et al. [6]. B. Liu et al. [7] produced a better method using SVM and called it "multi-state-mapping" (MSM). L. Tang et al. [8] put forward a method for the *k*-class classification using a regular simplex support vector machine (RSSVM). N. Agarwal et al. [9] developed a strategy to break a multiclass problem into smaller binary problems by substituting the softmax layer with a collection of binary SVM classifiers. M. Ramírez-Corona et al. [10] utilized the prediction value and level of the node for selecting the path from the root of the hierarchy. K. Kowsari et al. [11] proposed ensemble classification using multimodal deep learning architectures for gender detection. A. K. Verma et al. [12] used machine learning methods to ensemble five different data mining methods. They generated an ensemble model for prediction. D. Ndirangu et al. [13] used an ensemble filter method to select relevant data sets features. In their research, J. Bi et al. [14] put forward a method for selecting the most appropriate classification algorithm from various classification algorithms suited to each individual subset. The method is called DECOC. F. Chan et al. [15] concluded that an ensemble of neural networks (EBNN) could detect novelties and classify future data sets. K. Bhowmick et al. [16] proposed a hybrid ensemble strategy in order to classify multiclass imbalanced data (HECMI). K. V. Kokilam et al. [17] proposed the intensified approach for the multiclass classification of imbalanced data. Authors [32] in their paper corresponded that the Data Stream classification algorithm must be designed for processing fast drifting features. Tavasoli et al. [18] proposed an adaptive algorithm for data stream classification updates.

A. Research gaps

HAR using traditional ML techniques is challenging because of the following reasons-

i. Traditional ML requires manual extraction of features, limiting its effectiveness based on the domain knowledge of an individual.
ii. These approaches often extract low-level features, limiting their ability to recognize complex activities beyond basic physical or postural movements.
iii. Traditional ML does not effectively leverage the temporal correlations between input samples.
iv. Analyzing unstructured data is difficult for many machine learning algorithms.

Inspired by these difficulties, the current study employs a deep learning approach for human activity recognition (HAR). This deep learning framework automates the extraction of features and optimizes them for the intended purpose. It can identify key features necessary for recognizing complex human activities. Unlike traditional methods, there is no need to manually extract features in advance, thus saving time. Deep learning is capable of analyzing both structured and unstructured data from diverse sources. The deep-learning-based framework is further optimized using the extra tree ensemble technique. Optimization is done to reduce losses and provide more accurate results. In the proposed framework, the cross-entropy loss function is applied for optimization. Recent works are compared to the DELETO framework's accuracy and calculation time for recognizing various human activities. The proposed framework shows improved performance concerning accuracy and computation time.

III. MATERIALS AND METHODS

In this section the dataset used for human activity recognition in the proposed approach is discussed in detail along with the data preparation methods used.

A. Dataset Description

The UCI repository data set for human activity recognition was used for experimental purposes. The data set was built from the recordings of 30 people between 19 and 48 years of age. Various activities were captured using a wearable device equipped with built-in inertial sensors. The dataset comprised 10,299 instances with 563 attributes, containing no missing values and categorized into six classes. Table 1 provides detailed specifications of the training data and activities.

TABLE I SPECIFICATIONS OF INSTANCES OF EACH CLASS

Activity	No. of instances
LAYING DOWN	1,407
STANDING	1,374
SITTING	1,286
WALKING	1,226
WALKING_UPSTAIRS	1,073
WALKING_DOWNSTAIRS	986

Authorized licensed use limited to: South Eastern Technological University (SETU). Downloaded on August 18,2025 at 23:25:39 UTC from IEEE Xplore. Restrictions apply.

Figure 1 shows the percentage of various activities in the data set

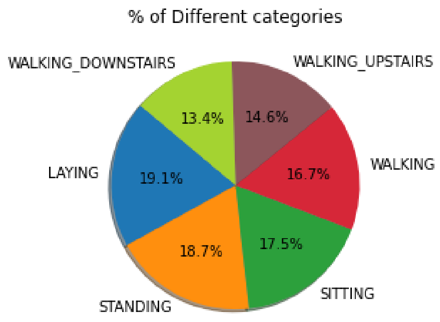

Fig .1. Percentage of activities in the data

Figure 2 shows the Scatter plot for different activities

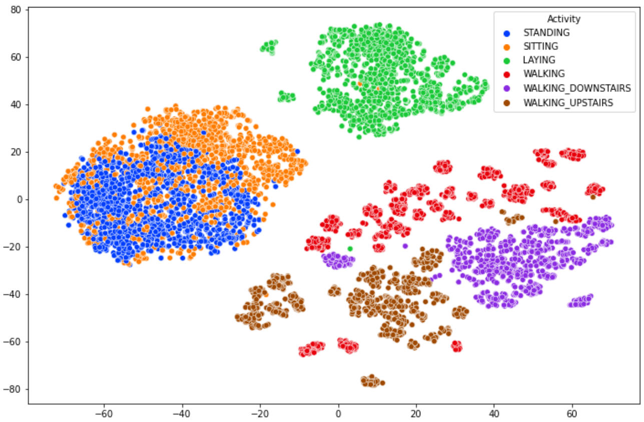

Fig .2. Scatter plot for different activities

B. Data Acquisition

The UCI repository was used to obtain the dataset. It contains 10299 instances containing 561 attributes of the subject with no missing values and a total of 6 classes. Each instance is associated with an activity label and an identifier. The obtained dataset is randomly divided into two sets, where 70% of volunteers were selected to generate training data and 30% to generate test data.

C. Data preparation

In this phase, the available data instances are first divided for training the model and testing the developed model in the ratio of 80:20. Following data splitting, data preprocessing deals with garbage values, missing values, followed by converting nominal values to numerical values. New features were generated, and the relationship among features is defined in the feature engineering step. Normalization is done in the next step to handle the variations in the data scales, followed by data transformation, where new features are generated from existing features.

Down sampling accelerometer observations to fractions of a second, using simple outlier detection and removal methods, removing activities with relatively fewer observations, rebalancing the activities by oversampling under-represented activities or under sampling over-represented activities in a training dataset, standardizing and normalizing data are some of the preparation schemes used during modelling.

IV. PROPOSED METHODOLOGY

This paper suggests the DELETO framework for recognizing various human body activities. The framework's methodology consisted of two main steps: learning and prediction.

In the learning phase, the available data instances are first divided for training and testing the developed DELETO framework in the ratio of 80:20. Followed by data splitting, data preprocessing is done to deal with garbage values, missing values, and so on, followed by converting nominal values to numerical values. New features were generated, and the feature engineering step defines the relationship among features. Normalization is done in the next step, followed by data transformation, where new features are generated from existing features to handle the variations in the data scales. Once the data is ready, model training and optimization are done using the proposed framework. In this step, the deep neural network, consisting of five layers, is trained using the training data. The first and last layers are the input and output layers, respectively, with three hidden layers. For multiclass classification, leaky ReLU was applied in the input and hidden layers (*the leaky ReLU has an advantage in that it overcomes the "dying ReLU" problem; thus, it takes less time to train a model*), and Softmax was applied in the output layer. *T*he mathematical expression for leaky ReLU is given in Equation 1:

$$f(x)=0.01x, x<0 \quad and\ f(x)=x, x \geqslant 0 \tag{1}$$

Figure 3 shows the deep learning model's input, output, and hidden layers.

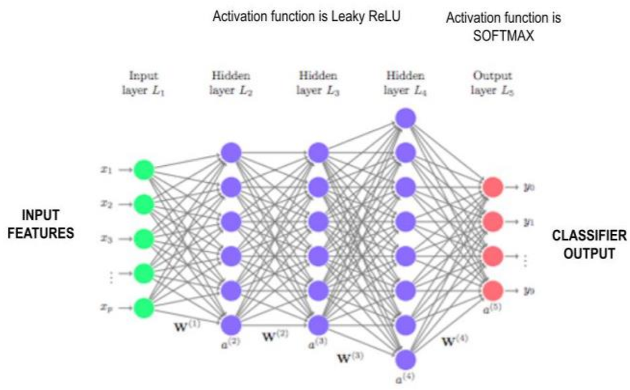

Fig .3. Layers of Deep learning model

$$E(W) = -\frac{1}{m}\sum_{i=1}^{m} y_i \, log(\gamma) + (1-y_i)\, log(1-\gamma) \tag{2}$$

Finally, the model was optimized using the extra tree ensemble-based feature selection technique. The optimization goal was that the model performed well and gave accurate predictions. The extra tree classifier built multiple trees by making bootstrap "false," which meant it sampled without replacement. Then the classifier chose the optimal split point for each of the k randomly chosen features at every node, which meant that the algorithm selected a split point randomly. The existing extra tree classifier chose a subset "S" and an attribute "a" for random split. The value of "a" was chosen randomly after finding the maximum and minimum values in "S" [amax, amin] [4]. The basis of the

Authorized licensed use limited to: South Eastern Technological University (SETU). Downloaded on August 18,2025 at 23:25:39 UTC from IEEE Xplore. Restrictions apply.

decision was the value of information gain. Feature selection is made by arranging features in descending order of their Gini importance and selecting the top k feature. The equation for Gini importance is given in Equation 3-

$$Gini\ (P) = \sum_{i=1}^{n} pi\ (1 - pi) \tag{3}$$

Table 2 shows the feature ranking of 25 essential features.

TABLE II FEATURE RANKING AND FEATURE IMPORTANCE

Feature rank	Feature number	Feature importance	Feature rank	Feature number	Feature importance
1	feature0	0.064192	14	feature 11	(0.029509)
2	feature 9	(0.059762)	15	feature 12	(0.029431)
3	feature 6	(0.059146)	16	feature 19	(0.029420)
4	feature 8	(0.058976)	17	feature 20	(0.029261)
5	feature 2	(0.058384)	18	feature 13	(0.029221)
6	feature 7	(0.054218)	19	feature 23	(0.029114)
7	feature 1	(0.054100)	20	feature 17	(0.029010)
8	feature 4	(0.052680)	21	feature 21	(0.029003)
9	feature 5	(0.050127)	22	feature 24	(0.028991)
10	feature 3	(0.047777)	23	feature 15	(0.028979)
11	feature 22	(0.031353)	24	feature 10	(0.028808)
12	feature 14	(0.030111)	25	feature 18	(0.028425)
13	feature 16	(0.030002)			

Figure 4 shows the graphic depicting the feature importance of all features.

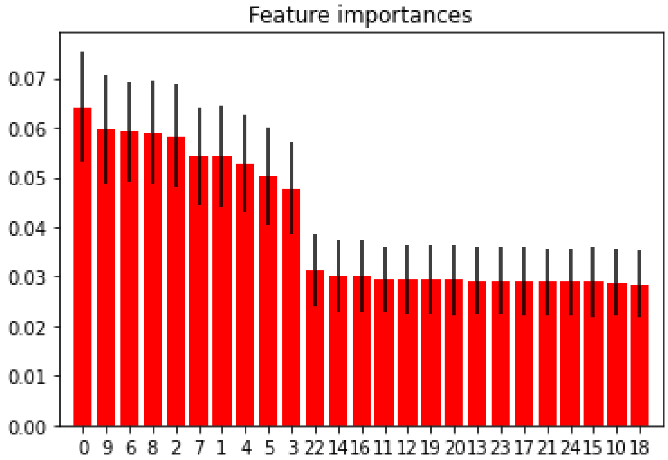

Fig .4. Feature Importance

All the feature subsets were ensembled to obtain the final feature subset. This approach was used to obtain the best features, procuring feature subsets from all stances. The steps of feature selection are shown in Figure 5. Finally, using the objective function, the model was optimized.

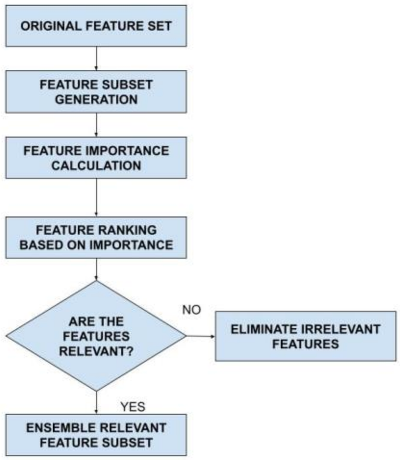

Fig .5. Flowchart for feature selection

The proposed DELETO algorithm is as follows-

Algorithm 1

Algorithm: An extra tree ensemble optimization based deep learning (DELETO)

Input: features x_n

Output: Label k for unseen sample x where $k \in (1 \ldots\ldots K)$

1: Transform $\boldsymbol{x_n}$ to tensor $\boldsymbol{T_{x_n}}$

2: At dense layer 1, feed $\boldsymbol{T_{x_n}}$

3: Calculate: $[\boldsymbol{T_{x_n}}, \boldsymbol{w_n}] = \boldsymbol{T_{x_i}} * \boldsymbol{w_i}$

4: Apply *leaky relu* activation function to output tensor $[T_{x_n}, w_n]$.

$$\boldsymbol{T_{d_n} = Leaky\,relu\,(W_n[T_{x_n}, w_n] + b_n)}$$

5: At dense layer 2, feed $\boldsymbol{T_{d_n}}$

6: Calculate: $[\boldsymbol{T_{d_n}}, \boldsymbol{w_m}] = \boldsymbol{T_{d_i}} * \boldsymbol{w_i}$

7: Apply *leaky relu* activation function to output tensor $[T_{d_n}, w_m]$.

$$\boldsymbol{T_{d_m} = relu\,(W_m[T_{d_n}, w_m] + b_m)}$$

8: At dense layer 3 (Output), feed $\boldsymbol{T_{d_m}}$

9: Now apply Softmax activation function σ to T_{d_m}

10: calculate the probability score for predicting the class of the transaction:

$$\hat{y} = \boldsymbol{\sigma\,(W_o[T_{d_m}] + b_f)}$$

11: Apply **Build_ETE( )**

12: calculate $\boldsymbol{\gamma = h(\hat{y}1, \hat{y}2, \ldots\ldots \hat{y}n)}$

13: Calculate objective function such as the error function: E(W)

$$\mathbf{E(W)} = -\frac{1}{m}\sum_{i=1}^{m} y_i \log(\gamma) + (1 - y_i)\log(1 - \gamma)$$

14. Apply optimization

Notations

$\boldsymbol{T_{x_n}}$, $\boldsymbol{T_{d_n}}$, and $\boldsymbol{T_{d_m}}$ are tensors applied to input, dense layer 1, and dense layer 2, respectively.

$\boldsymbol{W_n}$, $\boldsymbol{W_m}$, and $\boldsymbol{W_o}$ are activations applied at dense layer 1, dense layer 2, and output, respectively.

$$\boldsymbol{x_i = \{x_1, x_2 \ldots\ldots x_n\}}$$

$$\boldsymbol{w_i = \{w_1, w_2 \ldots\ldots w_n\}}$$

$\boldsymbol{x_i}$ is feature; $\boldsymbol{w_i}$is weight; n is no. of features/weights.

$\boldsymbol{b_n}$, $\boldsymbol{b_m}$, and $\boldsymbol{b_f}$,are bias applied at dense layer 1, dense layer 2, and output layer, respectively.

Leaky ReLU: **Leaky** $\boldsymbol{relu\ f(x) = 0.01x, if\ x < 0\ \&\ x, if\ x > 0}$

$\hat{\boldsymbol{y}}$: Score calculated by deep learning

$\boldsymbol{\gamma}$: Score calculated by ensemble

$\boldsymbol{h}(.)$: Ensemble

Authorized licensed use limited to: South Eastern Technological University (SETU). Downloaded on August 18,2025 at 23:25:39 UTC from IEEE Xplore. Restrictions apply.

```
# method Build_ETE(ES)
Input: Data set DS
Output: Extra tree ensemble ETE = {t1, t2, t3 ............tm}
for l = 1 to m
    Create a tree ti = Build_ET (DS)
return ETE
```

```
# method Build_ET(DS)
Input: Data set DS
Output: tree t
return leaf node
        if |Ts| < nmin
otherwise
        Select random k features {f1, f2,..........fk} from all candidate features in DS;
        Generate k split points {s1, s2,............sk} by using Randsplit_procedure Si = (DS,fi), where i = 1 to k;
        Select best split point s* based on maximum score, where
            Max_score (s*,DS) = max i=1 to k for a sample DS and a split s where
                Score = 2*Info_gain (DS)/(Es(DS) + Ec(DS))
       Split DS in SL and SR based on score s*;
       Build tL and tR by calling Build_ET();
       Create a new tree t by using s*, tL and tR
       return t
```

```
#method Randsplit_procedure( F, FV, η, Θ, N)
Input: Data set DS with N samples
    F feature set where F = {f1, f2,..........fN-1}
    Fv feature vector in set F
Output: Split node index Sl
for every feature f calculate
    Rj+1 = (η*Rj+ Θ) mod N
return Split node index Sl
```

Notations

$0 < j < N$

R_0 is a positive number

η is a multiplier

Θ is a positive integer

N is sample size

V. EXPERIMENT AND RESULTS

A. Experimental Steup

The system for the experiment comprised Intel Core i5 (2.0 GHz) machines having 8 GB RAM and Python installed with essential libraries. The operating system environment used is Microsoft Windows 10 (64-bit).

B. Experimental Parameters

The performance of the present proposed work is evaluated with metrics like F-score, accuracy, and computational time. The metric accuracy measures all correctly identified classes and is used when all the classes are equally important. The leaky ReLU used in the deep learning framework minimized the computational time and faster classification.

C. Results

The proposed DELETO algorithm for human activity recognition is implemented and the results are summarized in Table 3

TABLE III RESULTS OF DELETO OVER HAR DATASET

S.No	Evaluation Parameter	Result
1	Precision	0.981
2	Recall	0.995
3	Prediction Accuracy	98.7
4	F1-Score	0.989
5	Categorical Accuracy	83.18

Figure 6 Comparison of the overall results of the proposed work.

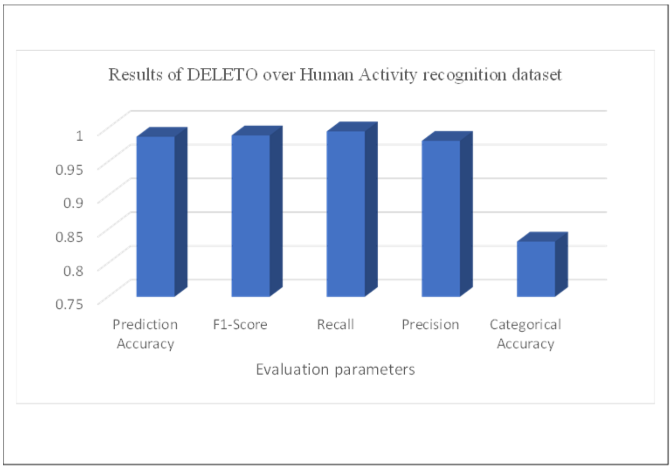

Fig .6. Results of DELETO over Human Activity recognition dataset

D. Comparison and Analysis

The proposed work is compared with state of art algorithms like MLP, SVM, RF, and LR as well as recent existing techniques for human activity recognition.

Table 4 compares the Accuracy, F1 scores, Computation time and Categorical prediction accuracy of the traditional classifiers with the proposed DELETO model.

TABLE IV COMPARISON WITH TRADITIONAL APPROACHES

Classifier	Accuracy	Computation time, (secs)	Categorical Prediction accuracy	F1-score
MLP	80.6	6.609	68.25	0.973
SVM	92.9	17.615	72.13	0.955
CNN	85.4	6.608	54.28	0.948
RF	91.9	15.126	52.62	0.98
LR	95.8	5.359	62.73	0.983
DELETO	98.7	2.533	83.18	0.989

Based on the findings, it can be concluded that the proposed model performed significantly better than the traditional approaches.

Table 5 compares the DELETO framework with related recent works.

Authorized licensed use limited to: South Eastern Technological University (SETU). Downloaded on August 18,2025 at 23:25:39 UTC from IEEE Xplore. Restrictions apply.

TABLE V COMPARISON WITH RECENT WORK

Title	Methodology	Accuracy (%)	Precsion	Recall	F1-score	Computation time (in Sec)
Deep RNN for HAR [5]	LSTM-based unidirectional, bidirectional, and cascaded architectures.	96.7	0.94	0.91	0.92	50.48
Ensemble ELM for HAR using smartphone sensors. [6]	Initialization of the input weights of base ELMs using Gaussian random projection.	97.3	0.93	0.94	0.93	44.32
Deep CNN for complex HAR using smartphones [7]	Building the DCNN classifier by adding a data compression module	95.27	0.94	0.92	0.95	49.3
DELETO (Proposed)		97.38	0.977	0.977	0.96	28.63

VI. CONCLUSION AND FUTURE SCOPE

Human activity recognition (HAR) is unfolding many research aspects due to its real-life applications in healthcare, the e-health system, and eldercare. HAR can be applied to build a smart environment and provide personalized support. ML techniques for HAR poses many challenges like manual extraction of relevant features, extraction of low-level features, hard to analyze unstructured data etc. The proposed DELETO model for multiclass classification of HAR was trained to recognize multiple activities like walking, running, and sitting. The extra tree forest feature selection technique selected the best features among all available features. The deep learning framework utilized leaky ReLU as the activation function in dense layers. The leaky ReLU had an advantage in overcoming the "dying ReLU" problem; thus, it took less time to train the DELETO model. Finally, the performance of the DELTO model was evaluated using metrics like F-score, accuracy, and categorical prediction accuracy. The DELETO model was also compared with conventional and recent works. The comparison showed a better result with approximately 1.5% improvement in accuracy and 2.8 % improvement in categorical accuracy.

REFERENCES

[1] Chaudhary, A., Kolhe, S., & Kamal, R. (2016). An improved random forest classifier for multiclass classification. *Information Processing in Agriculture*, *3*(4), 215–222.

[2] Panca, V., & Rustam, Z. (2017, July). Application of machine learning on brain cancer multiclass classification. In *AIP Conference Proceedings* (Vol. 1862, No. 1, p. 030133). AIP Publishing LLC.

[3] Beygelzimer, A., Langford, J., & Ravikumar, P. (2007). Multiclass classification with filter trees. *Preprint, June*, *2*.

[4] La Cava, W., Silva, S., Danai, K., Spector, L., Vanneschi, L., & Moore, J. H. (2019). Multidimensional genetic programming for multiclass classification. *Swarm and Evolutionary Computation*, *44*, 260–272.

[5] Zhang, J., Wang, Y., Sun, Y., & Li, G. (2020). Strength of ensemble learning in multiclass classification of rockburst intensity. *International Journal for Numerical and Analytical Methods in Geomechanics*, *44*(13), 1833–1853.

[6] Silva-Palacios, D., Ferri, C., & Ramírez-Quintana, M. J. (2017). Improving performance of multiclass classification by inducing class hierarchies. *Procedia Computer Science*, *108*, 1692–1701.

[7] Liu, B., Xiao, Y., & Cao, L. (2017). SVM-based multi-state-mapping approach for multiclass classification. *Knowledge-Based Systems*, *129*, 79–96.

[8] Tang, L., Tian, Y., & Pardalos, P. M. (2019). A novel perspective on multiclass classification: regular simplex support vector machine. *Information Sciences*, *480*, 324–338.

[9] Agarwal, N., Balasubramanian, V. N., & Jawahar, C. V. (2018). Improving multiclass classification by deep networks using DAGSVM and Triplet Loss. *Pattern Recognition Letters*, *112*, 184–190.

[10] Ramírez-Corona, M., Sucar, L. E., & Morales, E. F. (2016). Hierarchical multilabel classification based on path evaluation. *International Journal of Approximate Reasoning*, *68*, 179–193.

[11] Kowsari, K., Heidarysafa, M., Odukoya, T., Potter, P., Barnes, L. E., & Brown, D. E. (2020). Gender detection on social networks using ensemble deep learning. *arXiv preprint arXiv:2004.06518*.

[12] Verma, A. K., Pal, S., & Kumar, S. (2019). Classification of skin disease using ensemble data mining techniques. *Asian Pacific Journal of Cancer Prevention: APJCP*, *20*(6), 1887.

[13] Ndirangu, D., Mwangi, W., & Nderu, L. (2019). Using ensemble technique to improve multiclass classification.

[14] Bi, J., & Zhang, C. (2018). An empirical comparison on state-of-the-art multiclass imbalance learning algorithms and a new diversified ensemble learning scheme. *Knowledge-Based Systems*, *158*, 81–93.

[15] Chan, F. T., Wang, Z. X., Patnaik, S., Tiwari, M. K., Wang, X. P., & Ruan, J. H. (2020). Ensemble-learning based neural networks for novelty detection in multiclass systems. *Applied Soft Computing*, 106396.

[16] Bhowmick, K., Shah, U. B., Shah, M. Y., Parekh, P. A., & Narvekar, M. (2019). HECMI: Hybrid Ensemble Technique for Classification of Multiclass Imbalanced Data. In *Information Systems Design and Intelligent Applications* (pp. 109–118). Springer, Singapore.

[17] Kokilam, K. V., & Latha, D. P. P. (2020, December). Ensemble method to classify multi class with concept drift. In *Journal of Physics: Conference Series* (Vol. 1706, No. 1, p. 012151). IOP Publishing.

[18] Tavasoli, Hanane, B. John Oommen, and Anis Yazidi. "On utilizing weak estimators to achieve the online classification of data streams." *Engineering Applications of Artificial Intelligence* 86 (2019): 11-31.

Authorized licensed use limited to: South Eastern Technological University (SETU). Downloaded on August 18,2025 at 23:25:39 UTC from IEEE Xplore. Restrictions apply.